Graduate Game Theory

Muhamet Yildiz
MIT Economics Department

Spring 2022

Contents

I 14.126 Graduate Game Theory with Applications **1**

1 Basic Framework **3**

1.1 Extensive-form Representation 4
1.2 Normal-Form Representation 7
1.3 Informational Assumptions 11
1.4 Basic Concepts 13
1.4.1 Beliefs, Mixed Strategies, and Best Response 13
1.4.2 Perfect Recall 15
1.4.3 Mixed and Behavioral Strategies in Extensive-form Games 18
1.4.4 Continuity 20

2 Complete-Information Games in Normal Form **23**

2.1 Rationalizability 24
2.1.1 An Example 25
2.1.2 Rationalizability as the Outcome of Iterated Strict Dominance . . 27
2.1.3 Rationalizability—Fixed-point Definition 32
2.2 Nash Equilibrium 36
2.2.1 Existence of Nash Equilibrium 38
2.2.2 Upperhemicontinuity of Nash Equilibrium 39
2.2.3 Existence of Nash Equilibrium in Discontinuous Games 40
2.3 Correlated Equilibrium 42
2.4 Relation Between Nash Equilibrium, Correlated Equilibrium, and Rationalizability 52
2.5 Classical Applications 53

2.6 Notes on Literature . . . 61
2.7 Exercises . . . 62

3 Bayesian Games **67**
3.1 Basic Definitions . . . 67
3.2 Bayesian Nash Equilibrium . . . 72
3.2.1 Existence of Bayesian Nash Equilibrium . . . 74
3.2.2 Upperhemicontinuity of Bayesian Nash Equilibrium . . . 75
3.3 Rationalizability . . . 77
3.3.1 Ex-ante Rationalizability . . . 78
3.3.2 Interim Independent Rationalizability . . . 80
3.3.3 Interim Correlated Rationalizability . . . 81
3.3.4 Invariance to Representation of Belief-Hierarchies . . . 86
3.4 Correlated Equilibrium . . . 88
3.5 Bayesian Games in Extensive Form . . . 92
3.6 Notes on Literature . . . 93
3.7 Exercises . . . 94

4 Extensive-Form Games **101**
4.1 Sequential Equilibrium . . . 104
4.2 Equilibrium Refinements in Normal Form . . . 111
4.2.1 Perfect Equilibrium . . . 111
4.2.2 Proper Equilibrium . . . 115
4.2.3 Strategic Stability . . . 117
4.3 Backward Induction & Iterative Conditional Dominance . . . 122
4.4 Iterated Conditional Dominance in Bargaining . . . 125
4.5 Extensive-Form Rationalizability . . . 126
4.6 Forward Induction . . . 127
4.7 Exercises . . . 132

5 Supermodular Games **137**
5.1 Example . . . 138
5.2 Supermodular Optimization Problems . . . 139
5.2.1 Monotonicity Theorem . . . 140

5.2.2 Applications . . . 141
5.2.3 Extensions and Generalizations . . . 144
5.2.4 Monotonicity Theorem with Continuity and Completeness . . . 147
5.3 Supermodular Games . . . 148
5.3.1 Formulation . . . 148
5.3.2 Rationalizability and Equilibrium . . . 150
5.3.3 Comparative Statics . . . 153
5.3.4 Supermodular Bayesian Games . . . 155
5.4 Exercises . . . 158

6 Global Games **169**
6.1 Risk Dominance . . . 170
6.2 Example: A Partnership Game . . . 172
6.2.1 The case of uniform prior and small shock sizes . . . 177
6.2.2 The Case of Normal Distributions . . . 179
6.2.3 Model Uncertainty . . . 183
6.3 Global Supermodular Games . . . 187
6.4 Risk-dominant Equilibrium Selection in 2x2 Games . . . 192
6.5 Applications with Continuum of Players . . . 195
6.6 Dynamic Global Games . . . 199
6.6.1 Dynamic Contagion . . . 200
6.6.2 Dynamic Games of Regime Change . . . 205
6.6.3 Dynamics under Model Uncertainty . . . 205
6.6.4 Supermodularity with Dynamics . . . 205
6.7 Concluding Remarks . . . 205
6.8 Exercises . . . 205

7 Potential Games **217**
7.1 Ordinal Potential Games . . . 217
7.2 Potential Games . . . 220
7.3 Congestion Games . . . 224
7.4 Bayesian Potential Games . . . 226
7.5 Noise-independent Selection in Global Games . . . 227

II Advanced Topics **229**

8 Interactive Epistemology **231**

8.1 Standard Model of Knowledge—One-Person Case . . . 231
8.1.1 Knowledge Function . . . 232
8.1.2 Information Function . . . 233
8.1.3 Information Partition . . . 235
8.1.4 Knowledge Operator . . . 236
8.1.5 Knowledge Field . . . 239
8.1.6 Information Graph . . . 240
8.1.7 Beliefs and Certainty . . . 241
8.2 Informativeness and Lattice of Information Partitions . . . 243
8.3 Interactive Epistemology . . . 249
8.3.1 Epistemic Models . . . 249
8.3.2 Mutual Knowledge . . . 252
8.3.3 Common Knowledge . . . 255
8.3.4 Public Events . . . 257
8.3.5 Knowledge Field for Common Knowledge Operator . . . 258
8.3.6 Information Partition for Common Knowledge . . . 259
8.3.7 Information Graph for Common Knowledge . . . 261
8.4 Common Certainty . . . 263

9 Common-Prior Assumption **265**

9.1 Agreeing to Disagree . . . 267
9.2 No Trade Theorem . . . 272
9.2.1 Examples . . . 273
9.2.2 No Trade Theorem under Risk Neutrality . . . 278
9.2.3 No-Trade Theorem under Ex-ante Optimality . . . 280
9.2.4 No-Trade Theorem under Incentive Compatibility . . . 285
9.3 Exercises . . . 285

10 Epistemic Foundations of Solution Concepts **287**

10.1 Normal-Form Games . . . 287
10.1.1 Rationalizability . . . 287

10.1.2 Correlated Equilibrium 293
10.1.3 Nash Equilibrium 294
10.2 Bayesian Games 296
10.2.1 Interim Correlated Rationalizability 296

11 Incomplete Information and Higher-Order Beliefs 303
11.1 Examples 305
11.1.1 Higher-order Expectations with Normal Distributions 305
11.1.2 Cournot Duopoly 307
11.1.3 Cournot Oligopoly 309
11.1.4 Games with Linear Best Responses 310
11.1.5 E-mail Game 312
11.2 Model 315
11.3 Robustness to Incomplete Information 317
11.3.1 Upper-hemicontinuity of ICR 317
11.3.2 Structure Theorem 319
11.3.3 Common p-Belief and Strategic Topology 322
11.3.4 Ex-Ante Robustness 327
11.3.5 Common Belief Foundations of Global Games 328
11.4 Linear Algebra of Higher-Order Beliefs 330
11.5 Common-Prior Assumption and Higher-Order Beliefs 338
11.5.1 Limits of Higher-Order Expectations 338
11.5.2 Finite-order Implications of CPA 338
11.6 Notes on Literature 340
11.7 Exercises 342

A Mathematical Tools 1
A.1 Separating Hyperplane Theorem 1
A.2 Continuity of Correspondences 3
A.3 Berge's Maximum Theorem 4
A.4 Kakutani's Fixed-Point Theorem 7
A.5 Theory of Lattices and Supermodularity 7
A.5.1 Lattices 7

A.5.2 Strong Set Order and Sublattices 8
A.5.3 Functions on Lattices—Supermodularity 9
A.5.4 Increasing Differences and Supermodularity in Product Spaces . . 11
A.5.5 Order Topology and Continuity 14
A.5.6 Exercises 15

Part I

14.126 Graduate Game Theory with Applications

Chapter 1

Basic Framework

Game theory is the analysis of strategic situations in which there are multiple decision makers (i.e. players) and a decision maker's optimal action can depend on what the other players do. Such decision problems are formalized as *games*. A game can be formulated in two ways. First, one can describe the scenarios that players find it possible, describing who the players are, which player moves at any given point in time, what moves are available to him and what he knows when he makes his decision and how he feels about the possible outcomes. This is called *extensive-form representation*. A *strategy* of a player is a complete contingent plan that describes which move the player will play at any given contingency at which he makes a decision. Alternatively, one can simply describe the set of players, the set of strategies available to each player, and the payoffs. This is called *normal-form* or *strategic-form representation*. This section introduces these representations and some maintained assumptions formally.

Throughout the course the players are assumed to be Bayesian expected utility maximizers. First, every time they do not know something, they form a belief that is represented by a probability distribution. Second, they have a Von-Neumann and Morgenstern utility functions on the outcomes. When they need to make a choice, they choose an option that gives the highest expected utility under their beliefs.

Notation Throughout the course, for any given set X, $\Delta(X)$ denotes the set of probability distributions on X.[1] Given any probability distribution $p \in \Delta(X)$ and any

[1] Technical details: X is assumed to a topological space, and $\Delta(X)$ is the space of Borel probability measures on X, endowed with the weak topology.

measurable function f defined on X,

$$E_p[f] = \int f dp$$

denotes the expectation of f under p. The integral is in the summation form whenever X is finite or countable:

$$\int f dp = \sum_{x \in X} f(x) p(x).$$

Clearly, if X is a finite set with m members, $\Delta(X)$ is an $m-1$ dimensional simplex, embedded in $\mathbb{R}^m$. In that case, $\Delta(X)$ has the usual properties of $\mathbb{R}^m$. For example, a sequence (p_k) converges to p, denoted by $p_k \to p$, if $p_k(x) \to p(x)$ for each x. Of course, in that case, for any continuous and bounded function $f : X \to \mathbb{R}$,

$$\int f dp_k \to \int f dp.$$

The latter property also defines convergence of probability distributions for more general spaces X. That is, $p_k \to p$ if $\int f dp_k \to \int f dp$ for every continuous and bounded function $f : X \to \mathbb{R}$.

1.1 Extensive-form Representation

As illustrated in Figure 1.1, an extensive-form game consists of a tree that starts from an initial node and branches out towards terminal nodes, at which the players receive their payoffs. Along the way, there are decision nodes, at which the players make their moves. The players' information is represented by information sets, which are sets of decision nodes that the player cannot distinguish when he makes a decision. These concepts are formally introduced next.

Definition 1.1 *A directed graph is a pair $(X, <)$ where X is a set of nodes and $<$ is a binary relation; $x < y$ means x preceeds y (there is a directed path from x to y). A* tree *is a connected directed graph $(X, <)$ with an initial node ϕ that preceeds every $x \in X \setminus \{\phi\}$ and such that*

Transitivity $(x < y, y < z) \Rightarrow x < z$

Asymmetry $x < y \Rightarrow y \nless x$

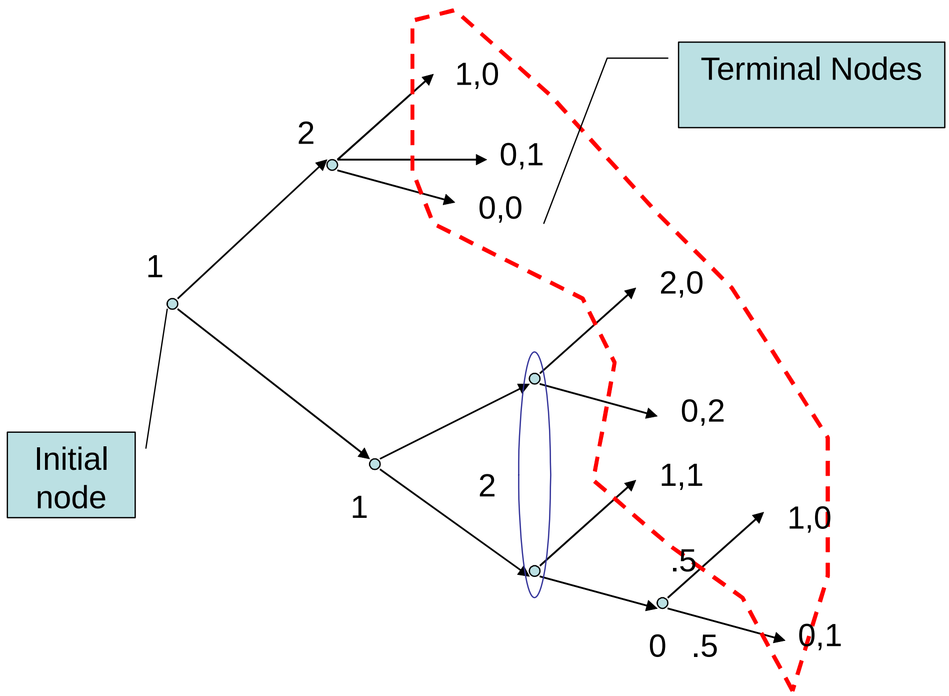

Figure 1.1: A game in extensive form

Arboresence $(x < z, y < z) \Rightarrow (x < y \text{ or } y < x)$.

It is the final property that makes a tree look like a tree. In a tree, each node is connected to the initial node through a unique path.

The nodes that are not followed by any other node are called *terminal*, and the set of all terminal nodes is denoted by Z. Each non-terminal node is connected to a set of immediate successors through an outgoing edge. These edges are labeled and called *moves* or *actions*. If two edges have the same label then they are the same move. The set of all moves (or actions) that are available at any given node x is denoted by $A(x)$. Since each node is connected to the initial node ϕ through a unique path, one can also represent any node x in a tree as the unique sequence $(a_0, a_1, \ldots)$ of actions that leads to x; the initial node ϕ corresponds to the empty sequence. For any node $x = (a_0, a_1, \ldots)$, define length $L(x)$ of x as the number of moves in the sequence $(a_0, a_1, \ldots)$. Observe that L can be finite or infinite. A tree has finite horizon if L is bounded; otherwise it has infinite horizon.

An extensive-form game is defined as follows.

Definition 1.2 (Extensive-form Game) *A game is a list* $G = (N, (X, <), H, \iota, p, u)$ *where*

- $N = \{1, \ldots, n\}$ *is the set of players,*
- $(X, <)$ *is a tree,*
- H *is a partition of non-terminal nodes* $X \backslash Z$ *such that,* $A(x) = A(y)$ *whenever* $x, y \in h \in H$*,*
- $\iota : H \to N \cup \{0\}$ *is a player mapping, where* $\iota(h) = 0$ *means that the nature moves at history* h*,*
- $p(\cdot|h) \in \Delta(A(h))$ *is a probability distribution on set* $A(h)$ *of available moves available at* h *for each history* h *with* $\iota(h) = 0$*, and*
- $u_i : Z \to \mathbb{R}^n$ *is the Von-Neumann and Morgenstern utility function of player* i*.*

Note that the partition H determines the players' information and their beliefs about those information. As in the definition above, we refer to h as a *history.* At any history h, the player $\iota(h)$ who moves at that history knows that he is at one of the nodes $x \in h$, but he cannot rule out any of the nodes x in h as the node he is at. When a history h is explicitly written as a set of nodes, it is also referred to as an *information set.* Here, the assumption that $A(x) = A(y)$ whenever $x, y \in h \in H$ ensures that a player knows what moves are available to him at any given history. Failure of such an assumption could lead to challenging (and perhaps interesting) philosophical issues.

The players' uncertainty is modeled by nature's moves when $\iota(h) = 0$ and the associated probability distributions $p(\cdot|h)$. Here, the formulation with a single common probability distribution $p(\cdot|h)$ corresponds to Common-Prior Assumption, which amounts to assuming that all the belief differences can be traced to informational differences. When the Common-Prior Assumption fails, one simply introduces n probability distributions $p_1(\cdot|h), \ldots, p_n(\cdot|h)$, one for each player. Also, the set of histories at which a player i moves is denoted by $H_i \equiv \iota^{-1}(i)$.

Finally, at any terminal node, the players receive their payoffs, and these payoffs are represented by the Von-Neumann and Morgenstern utility function $u_i : Z \to \mathbb{R}^n$ for each player i. That is, facing any two probability distributions p and q on the terminal nodes

Z, a player i prefers p to q if the expected value of u_i is higher under p; he chooses the option that maximizes the expected value of u_i. This will be the rationality assumption throughout: player i is said to be *rational* if u_i is his Von-Neumann and Morgenstern utility function in the above sense.

The following assumptions will be maintained throughout the course in theoretical results. First, there will be a finite number of players n. Second, the game tree will have "discrete time", in that any history can be indexed as a sequence of moves taken up to that point. The game can be of infinite horizon, in that some terminal histories may contain an infinite sequence of moves.

Finally, the players will have *perfect recall*. That is, a player does not forget what he has known or what he has done.

1.2 Normal-Form Representation

An extensive-form game can be represented in normal form as follows. In an extensive-form game, a *strategy* of a player i is any mapping

$$s_i : H_i \to \bigcup_{h \in H_i} A(h) \tag{1.1}$$

with

$$s_i(h) \in A(h) \qquad (\forall h \in H_i).$$

Note that the set of all strategies of a player i is

$$S_i = \prod_{h \in H_i} A(h). \tag{1.2}$$

Every strategy profile $s \in S = S_1 \times \cdots \times S_n$ induces a probability distribution $O(s)$ on Z, where the randomness comes from the nature's moves. The probability distribution $O(s) \in \Delta(Z)$ is called the *outcome* of s. The payoff of a player from s is the expected payoff under $O(s)$:

$$u_i(s) = \int u_i(z)\, dO(z|s). \tag{1.3}$$

The *normal-form representation* of the extensive-form game $G = (N, (X, <), H, \iota, p, u)$ is simply

$$(N, S, u)$$

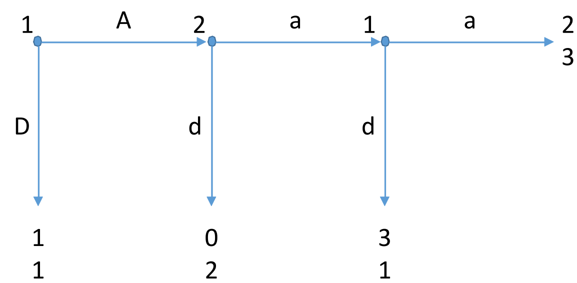

Figure 1.2: A Centipede Game

where N is the set of players, $S = S_1 \times \cdots \times S_n$ is the set of strategy profiles, $u = (u_1, \ldots, u_n)$ is the profile of payoff functions $u_i : S \to \mathbb{R}$. This is the representation of the game from the point of view of the players at the beginning of the game—i.e. at the *ex-ante* stage.

Example 1.1 *In the centipede game in Figure 1.2, Player 1 has four strategies, Aa, Ad, Da, and Dd where the first entry describes the move at the initial node while the second entry describes the move at the last one. Player 2 has only two strategies, a and d. The payoffs are as in the following table, where the first and the second entries are the payoffs of players 1 and 2, respectively,*

	a	d
Aa	$2,3$	$0,2$
Ad	$3,1$	$0,2$
Da	$1,1$	$1,1$
Dd	$1,1$	$1,1$

Two-player games with finite strategy sets will be defined through such a table in normal form.

Notation The following notational convention will be used throughout the course. Given any list $X_1, \ldots, X_n$ of sets with generic elements $x_1, \ldots, x_n$,

- $X = X_1 \times \cdots \times X_n$ with a generic element $x = (x_1, \ldots, x_n)$,
- $X_{-i} = \prod_{j \neq i} X_j$ with a generic element $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, and

- $(x_i', x_{-i}) = (x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$.

For example,

- $S = S_1 \times \cdots \times S_n$ is the set of strategy profiles $s = (s_1, \ldots, s_n)$,
- S_{-i} is the set of strategy profiles $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ for players other than player i, and
- $(s_i', s_{-i}) = (s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_n)$ is the strategy profile in which i plays s_i' and the others play s_{-i}.

Equivalent Strategies and Reduced-form Representation In some games there may be multiple strategies that lead to same outcome no matter what strategy other players play. For example, in the previous game, the strategies Da and Dd lead to same outcome for every strategy of Player 2. Such strategies differ only in what they prescribe at contingencies that are ruled out by those strategies. For example, under both strategies Da and Dd, Player 1 goes down at the initial node and the game ends. Hence, under those strategies the second decision node of Player 1 is never reached. At that node the strategies prescribe different actions, but that decision is not relevant for Player 1. Such strategies are called equivalent. Sometimes one can simplify the analysis by modeling equivalent strategies same, and reducing the set of strategies he considers. This is formalized next.

Definition 1.3 *Strategies s_i and s_i' are* equivalent *if*

$$O(s_i, s_{-i}) = O(s_i', s_{-i}) \qquad (\forall s_{-i}).$$

A *reduced-form representation* is a normal form representation in which one picks only one strategy from each equivalence class. For example, the normal-form game

	a	d
Aa	$2,3$	$0,2$
Ad	$3,1$	$0,2$
D	$1,1$	$1,1$

is a reduced-form representation of the centipede game above.

Agent-Normal Form The above normal-form definitions are *ex-ante* representations, in that they describe how players feel about the game at the beginning of the game. They make a contingent plan about how they will play the game and they evaluate those plans according to the expected payoffs they generate. Alternatively, one can represent a game closer to its extensive form, considering each decision problem separately. Under the alternative formulation, players take into account how they will play in the future but simply make the decision that they face at the moment. This is called agent-normal-form representation. This representation takes each decision node as a different player. Hence, the set of all players is simply H. The strategy set of each player h is simply $A(h)$. Finally, the payoff of each player $h \in H_i$ from a strategy profile, denoted by s, is simply $u_i(s)$, the ex-ante payoff defined in (1.3). The *agent-normal-form representation* of the extensive-form game $G = (N, (X, <), H, \iota, p, u)$ is simply

$$(H, A, u).$$

Since the payoffs of all "agents" $h \in H_i$ of player i are identical, one does not repeat them, writing one payoff for each player. The reason for taking a player's ex ante payoff as his payoff at a given history h is because his beliefs at the history is not given, and it is determined as part of the solution. (One could compute the player's belief on h induced by strategy profile s and compute the payoff of "agent" h from strategy profile s using the computed belief. Unfortunately, there are often information sets h that are reached with zero probability under s and one cannot apply this method to compute the payoff at h.)

Example 1.2 *In the centipede game of Figure 1.2, the agent-normal form representation has three players,* 1_1,1_2*, and* 2*, where* 1_1 *and* 1_2 *correspond to the first and the last information sets of player 1. The agent-normal-form game is*

	a	d
A	$2,3$	$0,2$
D	$1,1$	$1,1$

a

	a	d
A	$3,1$	$0,2$
D	$1,1$	$1,1$

d

where 1_1 *choose rows,* 2 *chooses columns, and* 1_2 *chooses matrices.*

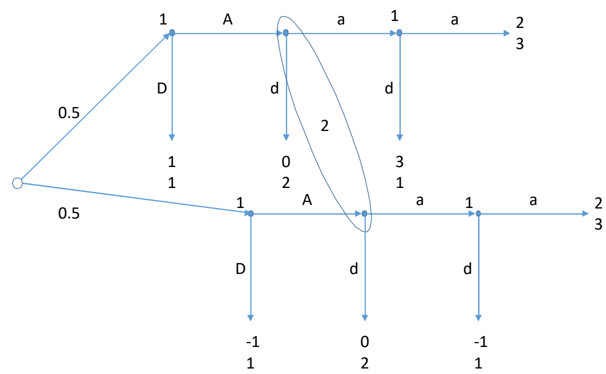

Figure 1.3: A centipede game with payoff uncertainty

1.3 Informational Assumptions

For any given extensive form game $G = (N, (X, <), H, \iota, p, u)$, everything in the description of the game is assumed to be *common knowledge* in the following sense. It is assumed that the set of players is N; at any given non-terminal history, player $\iota(h)$ makes a decision, and the set of available moves is $A(h)$; if $\iota(h)$ is 0, then the decision is made randomly according to $p(\cdot|h)$; otherwise player $\iota(h)$ knows that he is at one of the nodes x in h but does not know which one; the preferences of each player i on terminal nodes Z are represented by u_i. It is further assumed that every player knows all these, that every player knows that every player knows all these, and so on ad infinitum. Likewise, in a normal-form game $G = (N, S, u)$, it is assumed to be common knowledge that the set of players is N; the set of strategies available to i is S_i, and each player i is an expected utility maximizer with Von-Neumann and Morgenstern utility function u_i. Accordingly, these games are called games of *complete information.*

Note however that arbitrary information structures can be modeled within this framework by an appropriately chosen extensive-form game. A class of such games, namely Bayesian games, will be introduced later. Here, I will illustrate this point by an example. For example, consider the centipede game in Figure 1.2. In this game, it is assumed to be common knowledge that the players are 1 and 2; the players move alternatively and they can choose between exiting and staying, and the payoffs are as in the figure. What if Player 2 did not know the preferences of Player 1? Suppose she thinks that with probability 0.5 the payoffs are as in the figure, but with the remaining probability

Player 1 gets -1 whenever he exits; his payoffs are as before in other cases. One should also specify what Player 1 thinks about the beliefs of Player 2 and so on; let's assume that the above beliefs are common knowledge. This situation can be modeled by the game in Figure 1.3. At the beginning of the game, nature chooses between upper and lower branches with equal probabilities, modeling the uncertainty Player 2 faces. In the upper branch, the payoffs of Player 1 are as in the original game. They are as in the alternative scenario in the lower branch. Player 1 knows his payoffs. Accordingly, he knows which branch nature chose at every information set. Player 2 does not know the payoffs. Accordingly, she does not know which branch she is on when she makes her decision. Note that Player 2 knows that Player 1 knows his payoffs, and this is also reflected in this game. In order to model a situation in which Player 2 does not know whether Player 1 knows his payoffs, one can add two more branches that neither player can distinguish.

Since the new game in Figure 1.3 has a more nuanced extensive form, the normal-form representation contains more strategies for Player 1. He will have $2^4 = 16$ strategies. For example, one strategy could prescribe Dd on the upper branch and Aa on the lower branch. In the normal-form representation, it will be assumed that the players' payoff functions on such detailed strategy profiles are common knowledge. The payoff uncertainty is coded in the descriptions of the strategies. For example, suppose Player 1 plays the above strategy while Player 2 plays a. Then, the expected payoff of Player 1 will be $3/2$, as he gets 1 on the upper branch and 3 on the lower branch. In the normal form, this is common knowledge (and Player 2 knows the payoff function of Player 1), but Player 2 still does not know how Player 1 would feel about say exiting in the first chance. Likewise, although she knows the strategy of Player 1, she does not know what Player 1 does. She thinks that he will either exit at the beginning or go all the way, each with probability $1/2$, yielding the expected payoff of 2 for Player 2.

One could similarly introduce uncertainty about the set of moves available to other players or even uncertainty about the set of players that player may face. For example, in Figure 1.4, the game on the right panel depicts a situation in which Player 2 does not know whether Player 1 can exit the game. She thinks that he has that option with probability $1/2$ (as on the upper branch) and does not have that option with probability $1/2$ (as in the lower branch). The normal-form representation of this game

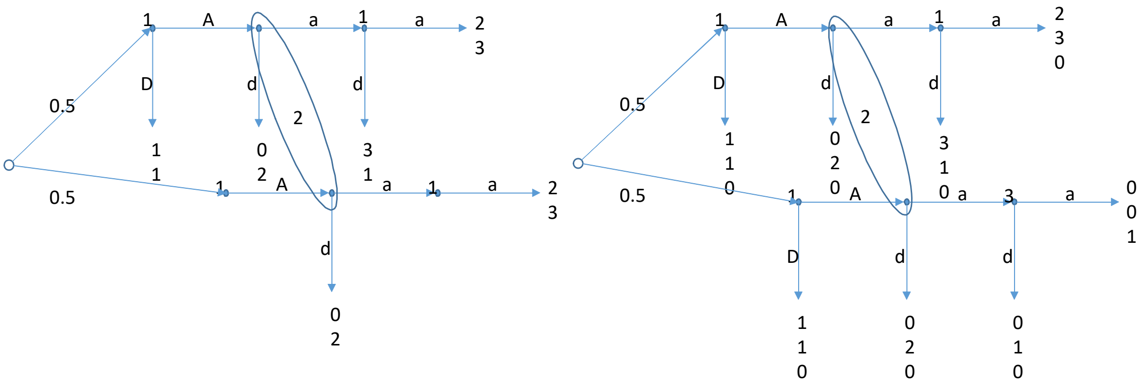

Figure 1.4: Centipede games with uncertainty about available moves (left) and the set of players (right)

has the same set of strategies as the original game but the payoffs take into account the possibility that Player 1 may not be able to exit; e.g., the expected payoffs from (Dd, a) are $0.5 \times 1 + 0.5 \times 2 = 1.5$ and $0.5 \times 1 + 0.5 \times 3 = 2$ for players 1 and 2, respectively. One can also be uncertain about the players she faces. For example, imagine that Player 1 is a supervisor and Player 2 is a subordinate. Player 2 does not know whether she will be working with Player 1 or there will be a new supervisor after remaining in. She thinks that the new supervisor would not terminate her but she does not want to work with the new supervisor. Player 1 knows whether he will be around, and all these are common knowledge. This case is modeled on the right panel of Figure 1.4, where the payoff is 0 when a player does not exist.

1.4 Basic Concepts

This section introduces the basic concepts such as beliefs, mixed strategies and best responses, and the related notational conventions throughout the course.

1.4.1 Beliefs, Mixed Strategies, and Best Response

Consider a normal form game $G = (N, S, u)$.

Definition 1.4 *For any player i, a* belief *(or a* conjecture*) of i about the other players'* *strategies is a probability distribution* μ_{-i} *on* $S_{-i} = \prod_{j \neq i} S_j$.

Remark 1.1 (Correlation) *This definition of belief* μ_{-i} *of player* i *allows correlation between the other players' strategies. For example, in a game of three players in which each player is to choose between Left and Right, Player 1 may believe that with probability 1/2 both of the other players will play Left and with probability 1/2 both players will play Right. Hence, viewed as mixed strategies, it may appear as though Players 2 and 3 use a common randomization device, contradicting the fact that Players 2 and 3 make their decisions independently. One may then find such a correlated belief unreasonable. This line of reasoning is based on mistakenly identifying a player's belief with other players' conscious randomization. For Player 1 to have such a correlated belief, he does not need to believe that the other players choose their decisions together. Indeed, he does not think that the other players are using randomization device. He thinks that each of the other players play a pure strategy that he does not know. He may assign correlated probabilities on the other players' strategies because he may assign positive probability to various theories and each of these theories may lead to a prediction about how the players play. For example, he may think that players play Left (as in the cars in England) or players play Right (as in the cars in France) without knowing which of the theories is correct.*

Definition 1.5 *The* expected payoff *from a strategy* s_i *against a belief* μ_{-i} *is*

$$u_i\left(s_i, \mu_{-i}\right) = \int u_i\left(s_i, s_{-i}\right) d\mu_{-i}\left(s_{-i}\right),$$

where the integral can be written as $\sum_{s_{-i} \in S_{-i}} u_i\left(s_i, s_{-i}\right) \mu_{-i}\left(s_{-i}\right)$ *when* S_{-i} *is finite.*

Definition 1.6 *For any player* i*, a strategy* s_i^* *is a* best response *to a belief* μ_{-i} *if*

$$u_i(s_i^*, \mu_{-i}) \geq u_i(s_i, \mu_{-i}), \forall s_i \in S_i.$$

The set of all best responses to a belief μ_{-i} *is denoted by* $B_i\left(\mu_{-i}\right)$.[2]

This uses the notion of a *weak best reply*, requiring that there is no other strategy that yields a strictly higher payoff against the belief. A notion of strict best reply would require that s^* yields a strictly higher expected payoff than any other strategy.

[2] For any set A of beliefs, I will also write $B_i\left(A\right) = \left\{s_i | s_i \in B_i\left(\mu\right), \mu \in A\right\}$ for the set of best responses to the beliefs in A.

Definition 1.7 *A* mixed strategy σ_i *of a player* i *is a probability distribution over the set* S_i *of his strategies.*

The expected payoff of a player from a mixed strategy profile $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is

$$u_i(\sigma) = \int u_i(s)\, d\sigma_1(s_1) \cdot \cdots \cdot d\sigma_n(s_n),$$

where the right-hand side is equal to $\sum_{s \in S} u_i(s)\, \sigma_1(s_1) \cdots \sigma_n(s_n)$ when S is finite. Here, it is assumed that $\sigma_1, \sigma_2, \ldots, \sigma_n$ are stochastically independent. For a correlated strategy profile $\sigma \in \Delta(S)$, the expected payoff is then

$$u_i(\sigma) = \int u_i(s)\, d\sigma(s).$$

In the correlated case, $\sigma(s)$ is not necessarily in the multiplicative form of $\sigma_1(s_1) \cdots \sigma_n(s_n)$. Similarly, I write

$$\begin{aligned} u_i(\sigma_i, s_{-i}) &= \int u_i(s)\, d\sigma_i(s_i), \\ u_i\left(s_i, \mu_{-i}\right) &= \int u_i(s)\, d\mu_{-i}(s_{-i}), \end{aligned}$$

and so on, where the integrals are again in summation form for finite strategy spaces.

1.4.2 Perfect Recall

The extensive-form representation above does not rule out the possibility that a player may forget what she has done or a piece of information she knew previously. Such forgetful players raise many important and interesting issues, but we will assume out these possibility in the remainder of the course. This section illustrates some of of the issues on a simple but interesting example and intrudces the formal concept of perfect recall that rules out possibility of forgetting.

Consider the one-person game in Figure 1.5, due to Piccione and Rubinstein (1997). This game is called the Absent-minded Driver. In th example, each decision node represents an exit on a highway where x means exit and a means continue (or go across). The driver's home is at the second exit, and she would like to go home. Unfortunately, there are no exit numbers and she forgets whether she passed an exit earlier. This is represented by the information set in the figure. Note that this example conforms all of

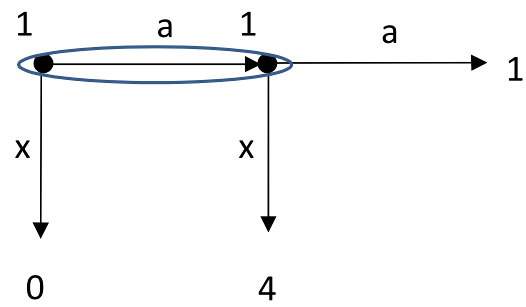

Figure 1.5: Paradox of Absent-minded Driver

the assumptions in the previous sections: the same player moves at both nodes and the set of available moves are the same. But since the second node comes after she chooses to continue, she does not know her own choice (i.e., she forgets by the time she arrives the second exit).

Her optimal action is to pass the first exit and exit at the second exit to go home, obtaining a payoff of 1. However, this plan is not possible as she cannot distinguish the two exits: she has to either choose x in both nodes, exiting at the first exit, or choose a in both nodes, going all the way through. Indeed, the normal-form representation of this game is $G = (N, S, u)$ where $N = \{1\}$, $S = \{a, x\}$, and

$$u(a) = 1 \text{ and } u(x) = 0.$$

In the normal-form representation, the payoff from any mixed strategy σ is

$$u(\sigma) = \sigma(a),$$

showing that her optimal strategy is to go across for sure, obtaining a payoff of zero.

The mixed strategy envisions a scenario in which the player randomizes between her strategies a and x, and if a comes up then she plays a at both nodes, and if x comes up, then she plays x at both nodes. In terms of behavior, her behavior at the two nodes are perfectly correlated. She mixes between a and x with probabilities $\sigma(a)$ and $\sigma(x)$, respectively, at the first node, but when she comes to the second node she plays a for sure. (Why?)

However, one can imagine that the player randomizes independently at the two nodes, randomizing between the moves at each node. If she randomizes deliberately, one could assume that the probability of $\sigma(x)$ is the same at both nodes (after all she could not distinguish the two nodes). But then her payoff from choosing $\sigma(x) = p$ leads to the following probability distribution on the outcomes:

$$\begin{aligned} \Pr(x) &= p \\ \Pr(ax) &= p(1-p) \\ \Pr(aa) &= (1-p)^2. \end{aligned}$$

Her expected payoff is

$$U(p) = 4p(1-p) + (1-p)^2 = 1 + 2p - 3p^2.$$

Her expected payoff is maximized at

$$p = 1/3.$$

In order to overcome her forgetfulness, she randomizes and exists with probability 1/3 to maximize her payoff. In contrasts she would randomize only if she were indifferent.

When a player is absent-minded, it is not clear at all that she will stick to her optimal strategy when it comes to executing her decision. Indeed, suppose that she can only choose a pure strategy. If she plans to choose x, then at the intersection, she will conclude that she must be at the first intersection because her plan calls for her exiting. Then, she would rather go across, deviating from her plan. If she plans to choose a, then she knows that each node is reached with probability 1 according to her plan. Hence, at her information set, she assigns probability $\Pr(n_1|I) = \frac{1}{1+1} = 1/2$ to the first node. Then, her payoff from exiting is $U(x) = \frac{1}{2} \times 0 + \frac{1}{2} \times 4 = 2$. Her payoff from going across is $U(a) = \frac{1}{2} \times 4 + \frac{1}{2} \times 1 = 5/2$; this is because, if she is at the first intersection, she knows that she will exit in the next exit and go home (according to her plan). Since $U(a) > U(x)$, she will once again deviate from her plan and continue. (This argument may not be as sound as it may appear; if she is not sticking with her plan, shy does she think that she will stick with it in the future?)

A remarkable result by Piccione and Rubinstein establishes that when players can randomize, the players will stick to their plans as long as they think that their past and future selves all stick to the plan. Indeed, in the absent-minded driver game above, under the optimal plan of exiting with probability 1/3, she will be indifferent between her actions at her information set. To see this, imagine that she is at an intersection. She wants to choose or randomize her move. She knows that at the initial node she would have gone across with probability 2/3. Hence, under her strategy, she will visit the first node for sure and she will visit the second node only with probability 2/3. Thus, knowning that she is at one of these nodes, she must assign probability

$$\Pr(n_2|p, I) = \frac{2/3}{1+2/3} = \frac{2}{5}$$

to the second node. Under her belief, her expected payoff from exiting is

$$U(x|p) = \frac{3}{5} \times 0 + \frac{2}{5} \times 4 = \frac{8}{5}.$$

Her expected payoff from going across is

$$U(a|p) = \frac{3}{5} \times (p \times 4 + (1-p) \times 1) + \frac{2}{5} \times 1 = \frac{8}{5},$$

where $p = 1/3$. She is indifferent and has no incentive to deviate.

In the rest of the course, we will assume that the players have *perfect recall.* That is, they do know the moves they make and they never forget any information that they had previously. Formally, the first requirement states that if $x < y$, then x and y cannot be in the same information set. This is because this information set corresponds to the player who moves at x and we require that she must know what she has done at x. The failure of this requirement is called absentmindedness (as in the above example). The second requirement is a bit more mouthfull: if x and y are in the same information set of a player i and there exists $x' < x$ at which player i moves, then there exists $y' < y$ such that x' and y' are in the same information set and player i makes the same move at x' on the path to x as the more she makes at y' on the path to y.

1.4.3 Mixed and Behavioral Strategies in Extensive-form Games

In an extensive form game, mixed strategies can be defined in two ways. First, one can consider mixing between the strategies in the normal form. This is what we call a *mixed strategy*:

$$\sigma_i \in \Delta(S_i).$$

Here, the player randomizes at the beginning of the game, but he then follows the realized deterministic strategy s throughout the game. Since we do not know the realized strategy, his behavior at any history h is random. Another way to define a mixed strategy would have been determining how a player mixes between the available actions at any given history. This is what we call a *behavioral strategy*:

$$\beta_i(\cdot|h) \in \Delta(A(h)), \qquad h \in H_i.$$

Here, the player randomizes between the available actions on the spot as he plays the game. Recall that in the absent-minded driver example in the previous section, some of the behavioral strategies could not be generated by any mixed strategy, and the optimal behavioral strategy was one of them.

Under perfect recall, these are equivalent ways of modeling randomized behavior. First, a mixed strategy σ determines a probability distribution at every history σ that it not precludes. Write

$$S_i(h) = \{s_i | \Pr(h|(s_i, s_{-i})) > 0 \text{ for some } s_{-i}\}$$

for the strategies s_i such that h is on the path of (s_i, s_{-i}) for some s_{-i}. When h is not precluded by σ_i in the sense that $\sigma_i(S_i(h)) > 0$, σ_i induces a unique probability distribution $\beta_{\sigma_i}(\cdot|h) \in \Delta(A(h))$ on the available actions $A(h)$, given by

$$\beta_{\sigma_i}(a|h) = \frac{\sigma_i(\{s_i \in S_i(h) | s_i(h) = a\})}{\sigma_i(S_i(h))}.$$

On the histories that are precluded by σ_i (with $\sigma_i(S_i(h)) = 0$), one can pick $\beta_{\sigma_i}(\cdot|h)$ arbitrarily. Here, β_{σ_i} is a behavioral strategy induced by mixed strategy σ_i. Here, the induced behavioral strategy is unique up to the arbitrariness on histories that are precluded by the mixed strategy. Second, conversely, every behavioral strategy is induced by a mixed strategy in this way. For example, given any behavioral strategy β_i, one can obtain a mixed strategy σ_{β_i} that induces the behavioral strategy by setting

$$\sigma_{\beta_i}(s_i) = \prod_{h \in H_i} \beta_i(s_i(h)|h) \tag{1.4}$$

at each s_i.

Example 1.3 *In the centipede game of Figure 1.2, consider the mixed strategy σ_i with $\sigma_i(Aa) = \sigma_i(Dd) = 1/2$. Strategy σ_i induces a behavioral strategy β_{σ_i} as follows. At history ϕ, both actions are equally likely, i.e.,*

$$\beta_{\sigma_i}(A|\phi) = \sigma_i(Aa) = 1/2.$$

At history $h = A\alpha$, Player 1 plays action a for sure:

$$\beta_{\sigma_i}(a|h) = \frac{\sigma_i(Aa)}{\sigma_i(Aa) + \sigma_i(Ad)} = 1.$$

Now, consider the behavioral strategy β_1 with $\beta_1(A|\phi) = \beta_1(D|\phi) = \beta_1(a|A\alpha) = \beta_1(d|A\alpha) = 1/2$. By (1.4), one can compute a mixed strategy σ_{β_1} that induces β_1 where $\sigma_{\beta_1}(Aa) = \sigma_{\beta_1}(Ad) = \sigma_{\beta_1}(Da) = \sigma_{\beta_1}(Dd) = 1/4$.

Exercise 1.1 *Find a game and two distinct mixed strategies that induce the same behavioral strategy.*

Kuhn (1953) shows that, under perfect recall, the mixed and behavioral strategies are equivalent. That is, any mixed strategy σ is equivalent to the behavioral strategy β_σ that it induces, in that they both induce the same probability distribution on the terminal histories. Conversely, for every behavioral strategy β, there exists a mixed staregy σ_β that induces β, and they are equivalent.

1.4.4 Continuity

The theoretical results in the course will often assume that utility functions are continuous and the strategy spaces are compact. This section introduces these concepts and assumptions. In some applications, these assumptions will fail. For example, the payoffs will be discontinuous in auctions, and it will be more convenient to allow any real number as strategies yielding a non-compact strategy space. Some of the theoretical results will allow these cases too.

Consider a normal form game $G = (N, S, u)$. It will be implicitly assumed that S is endowed by a metric or topology. In general, best response set will be non-empty if the strategy set is compact and the payoff function is continuous. Moreover, the best response can easily be empty if these conditions fail. For example, if a player is choosing how much money to have and he could choose any amount, his best response would be empty assuming he likes more money than less. This is because his strategy space is not compact although his utility is continuous. If his payoff function is discontinuous, the best response can also be empty even if the strategy set is discontinuous. For example, if the above player could select any amount of money up to a limit M, but he needs to pay a tax if he chooses M, then he will not have a best response. Therefore, the theoretical results typically make the following assumption:

Assumption 1.1 *The strategy space S is a compact metric space, and each utility function $u_i : S \to \mathbb{R}$ is continuous.*

A special case is of course finite games in which S is finite. Another important special case is when each S_i is a closed and bounded subset of $\mathbb{R}^m$ with continuous utility functions. Among other things, this assumption ensures that the expected utility is always well-defined, and the players always have a best response.

The next result, an immediate application of Berge's Maximum Theorem in the appendix, establishes the main existence and continuity properties of the best response correspondence B_i. (See Appendix A.2 for the relevant definitions and Appendix A.3 for the proof.)

Lemma 1.1 *Under Assumption 1.1, for each player i, the best response correspondence B_i is non-empty, compact-valued and upperhemicontinuous.*

The lemma has two parts. First, there exists a best response. Second, if a sequence of pairs $\left(s_i^m, \mu_{-i}^m\right)$ where $s_i^m \in B_i\left(\mu^m\right)$ converges to $\left(s_i, \mu_{-i}\right)$, then $s_i \in B_i\left(\mu_{-i}\right)$. That is, the limit of best responses to beliefs is a best response to the limiting belief. In other words, one cannot obtain a drastically different best response by perturbing the beliefs a little bit. All the new solutions will be near the original —although some of the best responses may disappear.

For an illustration, consider the following simple decision problem. There are two choices, Left and Right, and the payoff from Left is λ while the payoff from right is 0. Then, the best response correspondence is

$$B\left(\lambda\right) = \begin{cases} \{\text{Left}\} & \text{if } \lambda > 0, \\ \{\text{Left, Right}\} & \text{if } \lambda = 0, \\ \{\text{Right}\} & \text{if } \lambda < 0, \end{cases}$$

as a function of λ. For any given λ, if λ changes a little bit, no new solution appears (upperhemicontinuity). Also, note that, at $\lambda = 0$, some of the solutions disappear when λ varies slightly; for example, the solution Left disappears if λ increases. This is a failure of "lowerhemicontinuity". This illustrates what one can expect from solution concepts: upperhemicontinuity but not lowerhemicontinuity.

Extensive-Form Games The assumptions in extensive-form games are made to ensure that the normal-form representation satisfies the above assumption, i.e., the strategy sets are compact and the utility functions are continuous. It is assumed that, for each history h, the set $A(h)$ of available moves is a compact set (under some given topology). The set of strategies is then a compact set under the product topology on strategies. Defining continuity requires some care though.

Towards formulating the concept of continuity for utility functions, I first define a metric on the set of terminal nodes. For any two available actions $a \in A(h)$, $b \in A(h')$, let the distance between a and b be

$$d(a,b) = \begin{cases} d_h(a,b) & \text{if } A(h) = A(h') \\ 1 & \text{otherwise,} \end{cases}$$

where d_h is the metric on $A(h)$, with $d_h(a,b) \leq 1$ everywhere. Define a metric d on z by setting

$$d(z,z') = (1-\delta) \sum_{t=0}^{\min\{L(z),L(z')\}-1} \delta^t d(a_t,b_t)$$

at any $z = (a_0, a_1, \ldots)$ and $z' = (b_0, b_1, \ldots)$.

Assumption 1.2 *Under the metric d on Z, u is continuous. Moreover for each history h, $A(h)$ is compact.*

I will maintain Assumption 1.2 throughout. Observe that, since each $A(h)$ is compact, $S = \prod_{h \in H} A(h)$ is also compact (by Tychonoff's Theorem). Moreover, continuity of $u : Z \to \mathbb{R}^N$ implies that $u : S \to \mathbb{R}^N$ is also continuous.

When each $A(h)$ is discrete, the continuity of u in Assumption 1.2 reduces to the following well-known concept of continuity in (infinite-horizon) dynamic games. A game $G = (N, (X, <), H, \iota, p, u)$ is said to be *continuous at infinity* if for any $\varepsilon > 0$, there exists $\bar{t} < \infty$, such that

$$|u_i(z) - u_i(z')| < \varepsilon \tag{1.5}$$

for all $i \in N$ and for all terminal histories $z = (a_0, a_1, \ldots), z' = (b_0, b_1, \ldots) \in Z$ with $a_t = b_t$ for all $t \leq \bar{t}$. Note that finite-horizon games are continuous at infinity. When each $A(h)$ is finite, Assumption 1.2 reduces to assuming that u is continuous at infinity.

Chapter 2

Complete-Information Games in Normal Form

In this lecture, I will review the fundamental solution concepts of rationalizability and Nash equilibrium in complete information games. I will assume an introductory knowledge of these topics and present some main technical properties of these solution concepts. For rationalizability, I will show that it is equivalent to iterated elimination of strictly dominated strategies, using the Separating Hyperplane Theorem. For Nash equilibrium, I will show existence and upperhemicontinuity, using Kakutani's fixed point theorem and Berge's Maximum Theorem. I will show the continuity properties of rationalizability later for a more general setup. The above results will be used later in establishing existence and continuity properties of other solution concepts in seemingly more general setups. The three theorems mentioned above are the main technical tools used in economic theory, and my ulterior motive is to illustrate how to use these results. I present these results and the related concepts in the Appendix. It is advisable to study the appendix first.

Throughout this chapter, we fix a game $G = (N, S, u)$ in normal form where

- $N = \{1, \ldots, n\}$ is the set of players;

- $S = S_1 \times \cdots \times S_n$ is the set of all strategy profiles, so that S_i is the set of all strategies that are available to player i for each player $i \in N$; S is assumed to be a compact metric space, and

- for each player $i \in N$, $u_i : S \to \mathbb{R}$ is von Neumann-Morgenstern utility function of player i; it is assumed to be continuous.

Recall that each player i chooses a strategy that maximizes the expected value of u_i, where the expected values are computed with respect to his own beliefs. Player i is said to be *rational* if this is the case, i.e., his strategy maximizes the expected value of u_i (given his beliefs).

2.1 Rationalizability

The definition of a game (N, S, u) implicitly assumes that

1. the set of players is N, the set of available strategies to a player i is S_i, and the player i plays a strategy that maximizes the expected value of $u_i : S \to \mathbb{R}$ according to some belief, and that

2. each player knows 1, and that

3. each player knows 2, and that

   ...

n. each player knows $n - 1$

   ...

   ad infinitum.

That is, it is implicitly assumed to be *common knowledge* among the players that the game is $(N, S, u_1, \ldots, u_n)$ and that players are *rational* (i.e. they are expected utility maximizers). As a solution concept, rationalizability yields the strategies that are consistent with these assumptions, capturing what is implied by the model (i.e. the game). Other solution concepts impose further assumptions, usually on players' beliefs, to obtain sharper predictions. In this section, I will formally introduce rationalizability and present some of its applications. I will first illustrate the idea on a simple example, before presenting the formal theory.

2.1.1 An Example

Consider the following game.

1\2	L	R
T	$2, 0$	$-1, 1$
M	$0, 10$	$0, 0$
B	$-1, -6$	$2, 0$

(2.1)

A player is said to be rational if he plays a best response to a belief about the other players' strategies. What does rationality imply for this game?

Consider Player 1. Should he play T, or M, or B? A quick inspection of his payoffs reveals that his best play depends on what he thinks the other player does. Let's then write p for the probability he assigns to L (as Player 2's play), representing his belief about Player 2's strategy. His expected payoffs from playing T, M, and B are

$$\begin{aligned} U_T &= 2p - (1-p) = 3p - 1, \\ U_M &= 0, \\ U_B &= -p + 2(1-p) = 2 - 3p, \end{aligned}$$

respectively. These values are plotted in Figure 2.1 as a function of p. Clearly, U_T is the largest when $p > 1/2$, and U_B is the largest when $p < 1/2$. At $p = 1/2$, $U_T = U_B > 0$. Hence, if player 1 is rational, then he will play B when $p < 1/2$, D when $p > 1/2$, and B or D if $p = 1/2$.

Notice that, if Player 1 is rational, then he will never play M—no matter what he believes about the Player 2's play. Therefore, if we assume that Player 1 is rational (and that the game is as it is described above), then we can conclude that Player 1 will not play M. This is because M is a *strictly dominated strategy*. In particular, the mixed strategy that puts probability $1/2$ on T and probability $1/2$ on B yields a higher expected payoff than strategy M no matter what (pure) strategy Player 2 plays. A consequence of this is that M is never a weak best response to a belief p, a general fact that will be established momentarily.

What are the implications of the assumption that players know that the other players are also rational? Now, rationality of player 1 requires that he does not play M. For Player 2, her both actions can be a best reply. If she thinks that Player 1 is not likely

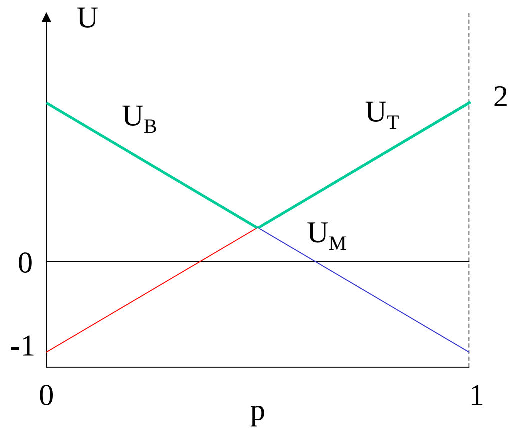

Figure 2.1: Expected payoffs in the example

to play M, then she must play R, and if she thinks that it is very likely that Player 1 will play M, then she must play L. Hence, rationality of player 2 does not put any restriction on her behavior. But, what if she thinks that it is very likely that player 1 is rational (and that his payoff are as in (2.1))? In that case, since a rational player 1 does not play M, she must assign very small probability for player 1 playing M. In fact, if she knows that player 1 is rational, then she must be sure that he will not play M. In that case, being rational, she must play R. In summary, *if player 2 is rational and she knows that player 1 is rational, then she must play R.*

Notice that we first eliminated all of the strategies that are never a (weak) best response (namely M), then taking the resulting game, we eliminated again all of the strategies that are never a best response (namely L). The resulting strategies are the strategies that are consistent with the assumption that players are rational and they know that the other players are rational.

As one imposes further assumptions about rationality, one keeps iteratively eliminating all strategies that are never a best response (if there remains any). Recall that rationality of player 1 requires him to play T or B, and knowledge of the fact that player 2 is also rational does not put any restriction on his behavior—as rationality itself does not restrict Player 2's behavior. Now, assume that Player 1 also knows (i) that Player 2 is rational and (ii) that Player 2 knows that Player 1 is rational (and that the game is as in (2.1)). Then, as the above analysis shows, Player 1 must know that Player 2 will play R. In that case, being rational, he must play B. Therefore, common knowledge of rationality implies that Player 1 plays B and Player 2 plays R.

2.1.2 Rationalizability as the Outcome of Iterated Strict Dominance

I will next apply these ideas more generally for a fixed game (N, S, u). First recall that a *belief* of a player i is a probability distribution μ_{-i} on $S_{-i} = \prod_{j \neq i} S_j$; a *mixed strategy* σ_i of a player i is a probability distribution over the set S_i of his own strategies, and $B_i\left(\mu_{-i}\right)$ is the set of best responses to the belief μ_{-i}, consisting of the strategies s_i^* with $u_i\left(s_i^*, \mu_{-i}\right) \geq u_i\left(s_i, \mu_{-i}\right)$ for every $s_i \in S_i$.

Depending on whether one allows correlated beliefs, there are two versions of Rationalizability. Following Remark 1.1, we allow correlated beliefs, leading to the correlated version of Rationalizability. Note that the original definitions of Bernheim (1985) and Pearce (1985) impose independence, and these concepts are identical in two player games.

Definition 2.1 *For any player i, playing a strategy s_i is said to be* rational *if and only if s_i is a best response to some belief μ_{-i}.*

Playing a strategy is not rational if and only if it is never a weak best reply. This idea of rationality is closely related to the following notion of dominance.

Definition 2.2 *A strategy s_i^** strictly dominates *s_i if*

$$u_i(s_i^*, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}.$$

Similarly, a mixed strategy σ_i strictly dominates *s_i if*

$$u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}.$$

That is, no matter what the other players play, playing s_i^* is strictly better than playing s_i for player i. In that case, if i is rational, he would never play the strictly dominated strategy s_i. That is, there is no belief under which he would play s_i, for s_i^* would always yield a higher expected payoff than s_i no matter what player i believes about the other players.[1]

[1] As a simple exercise, prove this statement.

Definition 2.3 *A strategy s_i is said to be* strictly dominated *if and only if there exists a pure or mixed strategy that strictly dominates s_i.*

Notice that, in the game in (2.1), neither of the pure strategies T, M, and B dominates any strategy. Nevertheless, M is dominated by the mixed strategy that σ_1 that puts probability 1/2 on each of T and B. For each p, the payoff from σ_1 is

$$U_{\sigma_1} = \frac{1}{2}(3p-1) + \frac{1}{2}(2-3p) = \frac{1}{2},$$

which is larger than 0, the payoff from M. Recall that in our example there is no belief (p) under which M is a best response. I will show next that this is indeed a general result.[2]

Theorem 2.1 (Pearce's Lemma) *Playing a strategy s_i is not rational for i (i.e. s_i is never a weak best response to a belief μ_{-i}) if and only if s_i is strictly dominated.*

Proof. I will first show that if s_i is not strictly dominated it is a weak best response to some belief. Let V be the set of continuous functions $v : S_{-i} \to \mathbb{R}$. For each mixed strategy σ_i, consider the function

$$u_i(\sigma_i, \cdot) : s_{-i} \mapsto (u_i(\sigma_i, s_{-i})),$$

and let A be the set of all such functions. For a visualization, note that when the strategy space is finite, $u_i(\sigma_i, \cdot)$ is simply the vector of expected payoffs from σ_i where the coordinates represent the pure strategies s_{-i} played by other players. Clearly, A is convex. Take any s_i that is not strictly dominated, and define

$$B = \{v \in V | v(s_{-i}) > u_i(s_i, s_{-i}) \ \forall s_{-i}\}.$$

Clearly, B is also convex and open. Moreover, since s_i is not strictly dominated, $A \cap B = \varnothing$. Hence, by the Separating-Hyperplane Theorem, there exists a continuous linear map $\lambda_{-i} : V \to \mathbb{R}$ such that

$$\lambda_{-i}(a) < \lambda_{-i}(b) \qquad \forall a \in A, b \in B. \tag{2.2}$$

[2] It may be useful to study the Separating-Hyperplane Theorem in the Appendix before studying the proof.

Since there is a sequence in B converging to $u_i(s_i,\cdot)$ and since λ_{-i} is continuous, inequality (2.2) implies that

$$\lambda_{-i}\left(u_i\left(\sigma_i,\cdot\right)\right) \leq \lambda_{-i}\left(u_i\left(s_i,\cdot\right)\right) \qquad \forall \sigma_i. \tag{2.3}$$

Moreover, since B is a comprehensive set (i.e., $x \in B$ whenever $x \leq y$ and $y \in B$), (2.2) implies that the linear mapping λ_{-i} is increasing. Hence, there exists a probability distribution $\mu_{-i} \in \Delta\left(S_{-i}\right)$ such that $\lambda_{-i}(v) = \int v d\mu_{-i}$. In particular, by (2.3),

$$\int u_i\left(\sigma_i, s_{-i}\right) d\mu_{-i}\left(s_{-i}\right) \leq \int u_i\left(s_i, s_{-i}\right) d\mu_{-i}\left(s_{-i}\right) \qquad \forall \sigma_i,$$

showing that s_i is a best response to the belief $\mu_{-i} \in \Delta\left(S_{-i}\right)$.

Conversely, take any s_i that is strictly dominated by σ_i:

$$u_i\left(\sigma_i, s_{-i}\right) > u_i\left(s_i, s_{-i}\right) \qquad \left(\forall s_{-i}\right).$$

Then, by monotonicity of expectations,

$$u_i\left(\sigma_i, \mu_{-i}\right) > u_i\left(s_i, \mu_{-i}\right) \qquad \left(\forall \mu_{-i}\right). \tag{2.4}$$

Now, suppose that s_i is a best response to some μ_{-i}:

$$u_i\left(s_i, \mu_{-i}\right) \geq u_i\left(s_i', \mu_{-i}\right) \qquad \left(\forall s_i'\right).$$

The monotonicity of expectation this time implies that

$$u_i\left(s_i, \mu_{-i}\right) \geq u_i\left(\sigma_i, \mu_{-i}\right),$$

contradicting (2.4). ■

Theorem 2.1 states that *if one assumes that players are rational (and that the game is as described), then one can conclude that no player plays a strategy that is strictly dominated (by some mixed or pure strategy), and this is all he can conclude.*

Write

$$S_i^1 = \left\{s_i \in S_i \mid s_i \text{ is not strictly dominated}\right\}.$$

By Theorem 2.1, S_i^1 is the set of all strategies that are best response to some belief, i.e.,

$$S_i^1 = B_i\left(\Delta\left(S_{-i}\right)\right).$$

Let us now explore the implications of the assumption that player i is rational and knows that the other players are rational. To this end, consider the strategies s_i that are best response to a belief μ_{-i} of i on S_{-i} such that $\mu_{-i}\left(S_{-i}^1\right) = 1$, where $S_{-i}^1 = \prod_{j \neq i} S_j$. That is, μ_{-i} puts zero probability for strategies that are not best response to any belief. Here, the first part (i.e. s_i is a best response to a belief μ_{-i}) corresponds to rationality of i and the second part (i.e. if $\mu_{-i}(s_{-i}) > 0$, then s_j is a best response to a belief μ_j) corresponds to the assumption that i knows that j is rational. By Theorem 2.1, each such s_j is not strictly dominated, i.e., $s_j \in S_j^1$. Hence, by another application of Theorem 2.1, s_i is not *strictly dominated given* S_{-i}^1, i.e., there does not exist a (possibly mixed) strategy σ_i such that

$$u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}) \qquad \forall s_{-i} \in S_{-i}^1.$$

Of course, by Theorem 2.1, the converse of the last statement is also true. Therefore, the set of strategies that are rationally played by player i knowing that the other players are also rational is

$$S_i^2 = \left\{ s_i \in S_i \mid s_i \text{ is not strictly dominated given } S_{-i}^1 \right\}.$$

By iterating this logic, one obtains the following iterative elimination procedure, called *iterative elimination of strictly-dominated strategies.*

Definition 2.4 (Iterative Elimination of Strictly-Dominated Strategies) *Set* $S^0 = S$, *and for any* $m > 0$ *and set*

$$S_i^m = \left\{ s_i \in S_i \mid s_i \textit{ is not strictly dominated given } S_{-i}^{m-1} \right\},$$

i.e., $s_i \in S_i^m$ *if and only if there does not exist any* σ_i *such that*

$$u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}) \qquad \forall s_{-i} \in S_{-i}^{m-1}.$$

Caution: Two points are crucial:

1. Only the *strictly* dominated strategies are eliminated. A weakly dominated strategy is not eliminated unless it is also strictly dominated. For example, no strategy is eliminated in

	L	R
T	1,1	0,0
B	0,0	0,0

although (T,L) is a dominant strategy equilibrium.

2. We do eliminate the strategies that are strictly dominated by mixed strategies (but not necessarily by pure strategies). For example, in the game in (2.1), we do eliminate M although neither T nor B dominates M.

Notice that when there are only finitely many strategies, this elimination process must stop at some m, i.e., there will be no dominated strategy to eliminate after a round. If elimination process never stops (in an infinite game), we eliminate indefinitely. The process is called *iterated elimination of strictly dominated strategies* or simply *iterated strict dominance.*

By Theorem 2.1, S^m corresponds to the set of strategies that are consistent with mth-order mutual knowledge of rationality, as formally stated next.

Theorem 2.2 *For every $m > 0$,*

$$S_i^m = B_i\left(\Delta\left(S_{-i}^{m-1}\right)\right).$$

Note that, for any m, a strategy s_i is in S_i^m if and only if it is rationally played by i in a situation in which (1) i is rational, (2) i knows that every player is rational, (3) i knows that everybody knows that every body is rational, and ... (m) i know that every body knows that ...everybody knows that everybody is rational. That is, s_i is a best response to a belief μ_{-i}^1 such that every s_j^1 in the support of μ_{-i}^1 is a best response to some belief μ_{-j}^2 such that every every s_k^2 in the support of μ_{-j}^2 is a best response to some belief μ_{-k}^3 ... up to order m. It is in that sense S^m is the set of strategy profiles that are consistent with mth-order mutual knowledge of rationality.

Rationalizability corresponds to the limit of the iterative elimination of strictly-dominated strategies.

Definition 2.5 (Rationalizability) *For any player i, a strategy is said to be* rationalizable *if $s_i \in S_i^\infty$ where*

$$S_i^\infty = \bigcap_{m\geq 0} S_i^m.$$

Rationalizability corresponds to the set of strategies that are rationally played in situations in which it is common knowledge that everybody is rational, as defined at the beginning of the lecture. When a strategy s_i is rationalizable, it can be justified/rationalized by an indefinite chain of beliefs μ_{-i} as above. On the other hand, if a strategy is not rationalizable, it must have been eliminated at some stage m, and such a strategy cannot be rationalized by a chain of beliefs longer than m. For the sake of completeness the following result states that there is always a rationalizable strategy. Theorem 10.2 formally establishes that rationalizability characterizes the strategies that are consistent with common knowledge of rationality.

Proposition 2.1 $S^{\infty} \neq \varnothing$.

Exercise 2.1 *Prove this proposition.*

Note that compactness and continuity assumptions in Assumption 1.1 are not superfluous. In a single player case, there may not be any optimal action when either of these conditions fail.

Note also that a strategy is eliminated even if it is only dominated by a previously eliminated strategies. That is, eliminated strategies stay around for elimination purposes. In practice one may want to ignore those strategies altogether, considering successive games with smaller strategies. This can be done when the game is finite but it may lead to wrong conclusions in infinite games.

Exercise 2.2 *For any game $G = (N, S, u)$, define $\hat{S}^0, \hat{S}^1, \ldots$ by $\hat{S}^0 = S$ and*

$$\hat{S}_i^m = \left\{ s_i \in \hat{S}_i^{m-1} | \ s_i \text{ is not strictly dominated in game } \left(N, \hat{S}^{m-1}, u\right) \right\}.$$

1. *Show that, if G is finite, then $\hat{S}^m = S^m$ for each m.*

2. *Find a game G in which $S^{\infty} = \varnothing$ but $\left|\hat{S}^1\right| = 1$, so that $\hat{S}^{\infty} = \hat{S}^1 = \hat{S}^2 = \cdots$ is non-empty. (Clearly Assumption 1.1 does not hold in G.)*

2.1.3 Rationalizability—Fixed-point Definition

There is also a fixed-point definition for rationalizability, which is useful in proving theorems, such as proving that rationalizability characterizes the strategies that are consistent with the common knowledge of rationality. Towards introducing fixed-point definition, I introduce a new formalism.

Definition 2.6 *A set* $Z = Z_1 \times \cdots \times Z_n \subseteq S$ *is said to* have best-response property *(or to be* closed-under rational behavior*) if for each* $i \in N$*,*

$$Z_i \subseteq B_i\left(\Delta\left(Z_{-i}\right)\right);$$

i.e., every $z_i \in Z_i$ *is a best response to some* $\mu \in \Delta\left(Z_{-i}\right)$*.*

To spell out the definition, suppose that it is common knowledge that a strategy profile in Z is played and this can be any strategy profile in Z. Best-response property states that one cannot refine this prediction any further by using rationality of the players and the common knowledge assumption. Note that if $Z = Z_1 \times \cdots \times Z_n$ and $Z' = Z_1' \times \cdots \times Z_n'$ have best-response property, so does $(Z_1 \cup Z_1') \times \cdots \times (Z_n \cup Z_n')$. Hence, the largest set Z with the best-response property exists. Rationalizability can also be defined as the largest Z with the best response property, as established in the following result. (Indeed, when this result fails, one must take the largest set with best response property as the set of rationalizable strategies, rather than the outcome of the iterated strict dominance.)

Theorem 2.3 *If* Z *has best-response property, then* $Z \subseteq S^\infty$*. Moreover, under Assumption 1.1,* S^∞ *has the best response property.*

Proof. For the first part, it suffices to show that $Z \subseteq S^m$ for every finite m. The statement is true for $m = 0$, by definition. Towards an induction, assume that $Z \subseteq S^{m-1}$. Then, for each i,

$$Z_i \subseteq B_i\left(\Delta\left(Z_{-i}\right)\right) \subseteq B_i\left(\Delta\left(S_{-i}^m\right)\right) = S_i^m,$$

completing the proof. (Here, the first inclusion is by definition of best-response property, the second inclusion is by the inductive hypothesis, and the equality is by Theorem 2.2.)

For the second part, take any $s_i \in S_i^\infty$. Since $s_i \in S_i^m$ for each m, there exists a sequence $\mu_m \in \Delta\left(S_{-i}\right)$ such that $s_i \in B_i\left(\mu_m\right)$ and $\mu_m\left(S_{-i}^{m-1}\right) = 1$. But since S_{-i} is a compact metric space, so is $\Delta\left(S_{-i}\right)$, and hence μ_m has a limit distribution $\mu \in \Delta\left(S_{-i}\right)$. Since B_i is upperhemicontinuous in μ (Lemma 1.1) and $s_i \in B_i\left(\mu_m\right)$ for each m, s_i is a best response to μ , i.e., $s_i \in B_i\left(\mu\right)$. One can also show that $\mu\left(S_{-i}^\infty\right) = 1$,[3] showing that S^∞ has the best-response property. ■

In the rest of this section, I will discuss rationalizability on two examples.

[3] For any m, since $\mu_k\left(S_{-i}^m\right) = 1$ for every $k > m$ and S_{-i}^m is closed, $\mu\left(S_{-i}^m\right) = 1$. Since S_{-i}^m converges monotonically to S_{-i}^∞, this further implies that $\mu\left(S_{-i}^\infty\right) = 1$.

Example: (Beauty Contest) Consider an n-player game in which each player i has strategies $x_i \in [0, 100]$, and payoff

$$u_i(x_1, \ldots, x_n) = -\left(x_i - \frac{2}{3}\frac{x_1 + \cdots x_n}{n}\right)^2.$$

Notice that, in this game, each player tries to play a strategy that is equal to two thirds of the average strategy, which is also affected by his own strategy. Each person is therefore interested guessing the other players' average strategies, which depends on the other players' estimate of the average strategy. Let's apply our procedure to this game.

First, since each strategy must be less than or equal to 100, the average cannot exceed 100, and hence any strategy $x_i > 200/3$ is strictly dominated by 200/3. Indeed, any strategy $x_i > x^1$ is strictly dominated by x^1 where[4]

$$x^1 = \frac{2(n-1)}{3n-2}100.$$

To show that $x_i > x^1$ is strictly dominated by x^1, we fix any $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n)$ and show that

$$u_i(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots x_n) < u_i\left(x_1, \ldots, x_{i-1}, x^1, x_{i+1}, \ldots x_n\right). \tag{2.5}$$

By taking the derivative of u_i with respect to x_i, we obtain

$$\frac{\partial u_i}{\partial x_i} = -2\left(1 - \frac{2}{3n}\right)\left(x_i - \frac{2}{3}\frac{x_1 + \cdots x_n}{n}\right).$$

Clearly, $\partial u_i / \partial x_i < 0$ if

$$\left(x_i - \frac{2}{3}\frac{x_1 + \cdots x_n}{n}\right) > 0,$$

which would be the case if

$$x_i > \frac{2}{3n-2}\sum_{j \neq i} x_j. \tag{2.6}$$

But since each $x_j \leq 100$, the sum $\sum_{j \neq i} x_j$ is less than or equal to $(n-1)100$. Hence, it suffices that

$$x_i > \frac{2}{3n-2}(n-1)100 = x^1.$$

Therefore, in the region $x_i > x^1$, u_i is a strictly decreasing function of x_i and (2.5) is satisfied. This shows that all the strategies $x_i > x^1$ are eliminated in the first round.

[4]Here x^1 is just a real number, where superscript 1 indicates that we are in Round 1.

On the other hand, each $x_i \leq x^1$ is a best response to some $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n)$ with

$$x_i = \frac{2}{3n-2} \sum_{j \neq i} x_j.$$

Therefore, at the end of the first round the set of surviving strategies is $[0, x^1]$.

Now, suppose that at the end of round m, the set of surviving strategies is $[0, x^m]$ for some number x^m. By repeating the same analysis above with x^m instead of 100, we can conclude that at the end of round $m+1$, the set of surviving strategies is $[0, x^{m+1}]$ where

$$x^{m+1} = \frac{2(n-1)}{3n-2} x^m.$$

The solution to this equation with $x^0 = 100$ is

$$x^m = \left[\frac{2(n-1)}{3n-2}\right]^m 100.$$

Therefore, for each m, at the end of round m, a strategy x_i survives if and only if

$$0 \leq x_i \leq \left[\frac{2(n-1)}{3n-2}\right]^m 100.$$

Since

$$\lim_{m \to \infty} \left[\frac{2(n-1)}{3n-2}\right]^m 100 = 0,$$

the only rationalizable strategy is $x_i = 0$.

Notice that the speed at which x^m goes to zero determines how fast we eliminate the strategies. If the elimination is slow (e.g. when $2(n-1)/(3n-2)$ is large), then many strategies are eliminated at very high iterations. In that case, predictions based on rationalizability heavily rely on strong assumptions about rationality, i.e., everybody knows that everybody knows that ... everybody is rational. For example, if n is small or the ratio $2/3$ is replaced by a small number, the elimination is fast and the predictions of rationalizability are more reliable. If the n is large or the ratio $2/3$ is replaced by a number close to 1, the elimination is slow and the predictions of rationalizability are less reliable. In particular, the predictions of rationalizability for this game is more robust in a small group than a larger group.

A general problem with rationalizability is that there are usually too many rationalizable strategies in economic models; the elimination process usually stops too early.

This limits the predictive power of the theory. For example, in the Matching Pennies game

	Head	Tail
Head	$-1, 1$	$1, -1$
Tail	$1, -1$	$-1, 1$

every strategy is rationalizable, and we cannot say what the players will do. By further requiring that the players' conjectures about the other players are know, Nash equilibrium attempts to refine rationalizability and sharpen the predictions.

2.2 Nash Equilibrium

In equilibrium, players' beliefs are identical to the mixed strategies of their opponents, and hence it is useful for equilibrium analysis to define the concept of best response to a strategy.

Definition 2.7 *For any player i, a strategy s_i^* is a* best response to a strategy profile s_{-i} *if*

$$u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}), \forall s_i \in S_i.$$

Similarly, a mixed strategy σ_i^* is a best response to a mixed strategy profile σ_{-i} *if*

$$u_i(\sigma_i^*, \sigma_{-i}) \geq u_i(s_i, \sigma_{-i}), \forall s_i \in S_i.$$

Note that, in the above definition, the mixed strategies are assumed to be stochastically independent:

$$\begin{aligned} u_i(\sigma_i^*, \sigma_{-i}) &= \sum_{s \in S} u_i(s)\, \sigma_i^*(s_i) \prod_{j \neq i} \sigma_j(s_j) \\ u_i(s_i, \sigma_{-i}) &= \sum_{s_{-i} \in S_{-i}} u_i(s) \prod_{j \neq i} \sigma_j(s_j). \end{aligned}$$

In the definition, I consider only the deviations by pure strategies. Indeed, since the payoffs are linear with respect to the probabilities, there exists a profitable deviation if and only if there exists a profitable deviation in pure strategies.

Definition 2.8 *A strategy profile* $(s_1^*, ..., s_n^*)$ *is a* Nash Equilibrium *if* s_i^* *is a best-response to* $s_{-i}^* = (s_1^*, ..., s_{i-1}^*, s_{i+1}^*, ..., s_n^*)$ *for each* i*. That is, for all* i*, we have that*

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \quad \forall s_i \in S_i.$$

Similarly, a mixed strategy profile $(\sigma_1^*, ..., \sigma_n^*)$ *is a* Nash Equilibrium *if* σ_i^* *is a best-response to* $\sigma_{-i}^* = (\sigma_1^*, ..., \sigma_{i-1}^*, \sigma_{i+1}^*, ..., \sigma_n^*)$ *for each* i*.*

In other words, no player would have an incentive to deviate, if he correctly guesses the other players' strategies. If we consider a strategy profile a social convention, then being a Nash equilibrium is tied to being self-enforcing, that is, nobody wants to deviate when they think that the others will follow the convention. A Nash equilibrium distribution is a distribution $\sigma^* = \sigma_1^* \times \cdots \times \sigma_n^* \in \Delta(S)$ induced by a mixed-strategy Nash equilibrium $\sigma_1^* \times \cdots \times \sigma_n^*$.

Rationalizability has a strong epistemic foundation: it characterizes the strategic implications of common knowledge of rationality, characterizing what is implied by the implicit assumptions made in the definition of a game. As we will see later, it has also strong evolutionary foundations: in any adaptive process the proportion of the players who play a non-rationalizable strategy will go to zero as the system evolves.

Nash equilibrium has weaker foundations. Here, I will mention two of these, as they provide alternative interpretations to a mixed strategy Nash equilibrium. First, Aumann and Brandenburger (1995) show that, in a two-player game, if it is mutually known that

- the conjecture of player 1 about player 2's strategy is $\psi_1 \in \Delta(S_2)$,
- the conjecture of player 2 about player 1's strategy is $\psi_2 \in \Delta(S_1)$,
- the players are rational,

then (ψ_2, ψ_1) is a Nash equilibrium.[5] Note that the conjecture ψ_j of player j about player i is a probability distribution on the strategy set of player i and hence can be formally considered as a mixed strategy of player i. Aumann and Brandenburger's result states that one can indeed interpret a Nash equilibrium as an equilibrium of conjectures, where

[5] The proof of this result is rather straightforward. If player i knows that player j is rational and has conjecture ψ_j, then every s_j that he assigns positive probability (i.e. $\psi_i(s_j) > 0$) is a best response to ψ_j. This is indeed an alternative definition of a Nash equilibrium.

one's conjecture is the other player's strategy. For more than two players, one needs much stronger assumptions to reach such an interpretation of a Nash equilibrium. After all, in order to interpret a player's mixed strategy as the other players' conjectures about him, all the other players must have the same conjecture about him. A second interpretation comes from evolution. If one interprets the mixed strategies as the proportion of players playing various strategies, all Nash equilibria are steady states of adaptive processes.

In the rest of this section, I will present the existence and continuity properties of Nash equilibrium. These results will be used later.

2.2.1 Existence of Nash Equilibrium

Under broad continuity assumptions for utility functions and compactness and convexity assumptions for strategy sets, one can easily show that a Nash equilibrium exists. Here, the continuity and compactness assumptions are indispensable because they are needed existence of a solution to optimization problems, which are Nash equilibria in single-player games. Convexity assumption is used for satisfying the conditions of fixed-point theorems, such as Kakutani's Fixed-Point Theorem. The latter fixed-point property is clearly relevant only for equilibrium in multi-player games. There are also fixed-point theorems for non-convex spaces, such a Tarski's Fixed-Point Theorem, and one can also use such theorems to establish existence of Nash equilibrium in non-convex games. Here, I will present a general existence theorem, building on Kakutani's Fixed-Point Theorem (see Appendix A.4).

Theorem 2.4 *Let $G = (N, S, u)$ be a game where each S_i is a convex, compact subset of a Euclidean space and each $u_i : S \to \mathbb{R}$ is continuous in s and quasi-concave in s_i. Then, there exists a Nash equilibrium $s^* \in S$ of game G.*

Proof. In the proof, I will construct a correspondence $F : S \rightrightarrows S$ that satisfies the conditions of Kakutani's Fixed-Point Theorem and whose fixed-points are all Nash equilibria of G. One can then conclude that F has a fixed point, which is a Nash equilibrium. Let $F : S \rightrightarrows S$ be the "best reply" correspondence:

$$F_i(s) = B_i(s_{-i}) \qquad (\forall s \in S, i \in N).$$

Since S is compact and the utility functions are continuous, by the Maximum Theorem (see Lemma 1.1), F is non-empty and has closed graph. Moreover, by quasi-concavity,

F is also convex-valued. Hence, F satisfies the conditions of Kakutani's fixed-point theorem. Therefore, F has a fixed point:

$$s^* \in F(s^*).$$

Since $s_i^* \in B_i\left(s_{-i}^*\right)$ for each i by definition of F, s^* is a Nash equilibrium. ■

For games with convex strategy sets and quasiconcave utility functions, Theorem 2.4 proves existence of a *pure* strategy Nash equilibrium. One can use this result to establish existence of equilibrium in classical economic models, such as the Cournot competition presented in the next section. Theorem 2.4 has another less obvious application:

Corollary 2.1 *Every finite game $G = (N, S, u)$ has a (possibly mixed) Nash equilibrium σ^*.*

Proof. Since S is finite, each $\Delta(S_i)$ is a simplex in a Euclidean space; in particular it is convex and compact. Moreover, $u_i(\sigma) = \sum_s u_i(s)\,\sigma_1(s_1)\cdot\cdots\cdot\sigma_n(s_n)$ is continuous in σ and linear in σ_i. Hence, $G' = (N, \Delta(S_1), \ldots, \Delta(S_n), u)$ satisfies the assumptions of Theorem 2.4. Therefore, there exists a Nash equilibrium $\sigma^* \in \Delta(S_1) \times \cdots \times \Delta(S_n)$, which is also a (mixed) Nash equilibrium of G. ■

While continuity and compactness assumptions are often made in theory (as they are used for existence of best response), the games in many applications, such as auctions, have discontinuous utility functions. (See Section 2.2.3 below for existence of equilibrium in games discontinuous utility functions.)

2.2.2 Upperhemicontinuity of Nash Equilibrium

The Maximum Theorem establishes that the best-response correspondence is upperhemicontinuous in parameters when the payoffs are continuous and the domain is compact. In that case, in optimization problems, the limits of the solutions is a solution to the optimization problem in the limit. One can then find *a* solution by considering approximate problems and taking the limit. There can be other solutions in the limit, and hence best response correspondence is not lowerhemicontinuous in general. Nash equilibrium (like many other solution concepts) inherits these properties of the best response correspondence. I will next establish this result.

Consider a compact metric space X of some payoff-relevant parameters. Fix a set N of players and set S of strategy profiles, where S is again a compact metric space. The utility function of each player i depends on x as well as s. That is, $u_i : S \times X \to \mathbb{R}$. The utility functions are assumed to be continuous (both in strategies and the parameters). It is also assumed that parameter value x is commonly known. Write $NE(x)$ and $PNE(x)$ for the sets of all Nash equilibria and all pure Nash equilibria, respectively, of game $(N, S, u(\cdot; x))$ in which it is common knowledge that the parameter value is x.

Theorem 2.5 *If S is a compact metric space and each u_i is continuous (in (s, x)), then the correspondences NE and PNE are compact-valued, and upperhemicontinuous.*

Proof. I will prove the result for NE; the result for PNE is more straightforward. Since the strategy space is compact, the above conditions are equivalent to closed-graph property, which I will prove. To this end, take any sequence $(x^m, \sigma^m) \to (x, \sigma)$ with $\sigma^m \in NE(x^m)$ for each m. Suppose that $\sigma \notin NE(x)$. Then,

$$u_i(s_i, \sigma_{-i}, x) > u_i(\sigma_i, \sigma_{-i}, x)$$

for some $i \in N$ and $s_i \in S_i$. But, since $u(\sigma, x)$ is continuous in (σ, x) (by Lemma A.1 in the Appendix) and $(x^m, \sigma^m) \to (x, \sigma)$, this implies that

$$u_i\left(s_i, \sigma_{-i}^m, x^m\right) > u_i\left(\sigma_i^m, \sigma_{-i}^m, x^m\right)$$

for some large m, showing that $\sigma^m \notin NE(x^m)$—a contradiction. ■

2.2.3 Existence of Nash Equilibrium in Discontinuous Games

The utility functions are discontinuous in many important games, such as first price auction and Bertrand competition. While one needs some form of compactness of strategy space and continuity of utility functions to ensure existence of best response, continuity assumption can be substantially relaxed for existence of equilibrium. In this section, I will present an existence theorem for discontinuous games that covers many important applications with discontinuous payoffs. The following is the key assumption that replaces the continuity assumption. (Here, each S_i is assumed to be a metric space. An open neighborhood of s_{-i} is an open subset of S_{-i} that contains s_{-i}.)

Definition 2.9 *A player i can* secure *a payoff v at strategy profile $s \in S$ if there exists $\bar{s}_i \in S_i$ such that*

$$u_i\left(\bar{s}_i, s'_{-i}\right) \geq v$$

for all s'_{-i} on some open neighborhood of s_{-i}.

That is, even if other players slightly deviate from their strategies, he can ensure a payoff of v by playing $\bar{s}_i$. For example, if u_i is continuous, he can secure $u_i(s) - \varepsilon$ for any positive ε by playing s_i. Similarly, in discrete games, each $\{s_{-i}\}$ is open, and hence player i can secure $u_i(s)$ by playing s_i. (The utility functions are vacuously continuous in such games.) We will use this definition for discontinuous games where one can secure $u_i(s) - \varepsilon$ by playing some other strategy.

Definition 2.10 *A game (N, S, u) is* better-reply secure *if for any s^* that is not a Nash equilibrium and for any sequence $s^n \to s^*$ with limit payoff $v^* = \lim_n u_i(s^n)$, some player i can secure $v_i^* + \varepsilon$ at s^* for some $\varepsilon > 0$.*

Note that, since s^* is not a Nash equilibrium, there exists a player i who will get more than $u_i(s^*)$ by deviating to some $\bar{s}_i$. If u_i is continuous, he would secure a payoff of $v_i^* + \varepsilon$ at s^* for some $\varepsilon > 0$ by playing $\bar{s}_i$. Hence, the better deviation payoff is guaranteed even if other players slightly deviate from s^*. The definition requires this without requiring continuity. This weaker assumption is satisfied in many applications, such as first-price auction. The next result establishes existence of a pure strategy equilibrium for such games.

Theorem 2.6 *If $G = (N, S, u)$ is better reply secure, S is compact, and u_i is quasi-concave in s_i for each i, then there exists a Nash equilibrium $s^* \in S$ of game G.*

This result is proven by approximating the utility functions u_i by continuous functions. Recall that, by upper-hemicontinuity, limits of equilibria of the approximated games would also be equilibria of the original game if the utility functions were all continuous–even at the limit. Without continuity, a limit of the equilibria of the approximations need not be an equilibrium of the original game. The better reply security condition ensures that it is indeed a Nash equilibrium. This is because, if it were not, a player would secure a better payoff by a deviation strategy, which will remain to be a profitable deviation when the approximation is sufficiently small. (See Bertrand competition below for an application of this result. There are also simpler sufficient conditions. As in Corollary 2.1, one can obtain an existence result in mixed strategies without requiring quasiconcavity.)

2.3 Correlated Equilibrium

We have so far considered two fundamental solution concepts: rationalizability and Nash equilibrium. These solution concepts can be viewed as two extreme benchmarks. At one extreme, rationalizability assumes only common knowledge of rationality, assumptions that are already made in the definition of the game. Under rationalizability, the players' beliefs are constrained only by support restrictions. Their beliefs can have no relation to what the other players are actually doing, and the beliefs of two players about a third one can be quite different. At the other extreme, Nash equilibrium assumes that players knows what the other players are doing; this assumption is made for the conjectures of the players in mixed strategy equilibria. In this section, we will consider a middle solution concept: correlated equilibrium. A correlated equilibrium is an ex-ante theory about how players play and rationalize their behavior. It restricts players' beliefs assuming that they come from a (possibly hypothetical) common prior.[6]

Definition 2.11 (Information Structure) *An* information *(or* belief*)* structure *is a list* $(\Omega, (I_i)_{i \in N}, (p_{i,\omega})_{i \in N, \omega \in \Omega})$ *where*

- Ω *is a (finite)* state space*,*
- I_i *is a partition of* Ω *for each* $i \in N$*, called* information partition *of* i*, each cell of* I_i *being called an* information set *of player* i*,*
- $p_{i,\omega}$ *is a probability distribution on* $I_i(\omega)$*, which is the information set of* i *that contains* ω*, representing the* belief *of* i*.*

For a given game $G = (N, S, u)$*, an epistemic model is a pair* $M = (\Omega, (I_j)_{j \in N}, P, \boldsymbol{s})$ *of an information structure* $(\Omega, (I_j)_{j \in N}, P)$ *and a mapping* $\boldsymbol{s} : \Omega \to S$ *such that each* s_i *is constant over each information set* $I_i(\omega)$*.*

[6] We will study the common-prior assumption in greater detal in Chapter 9.

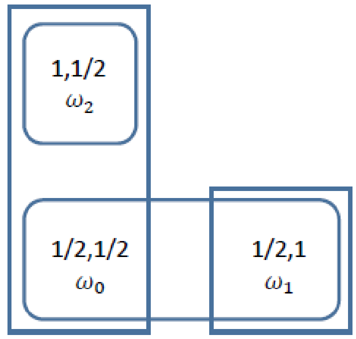

Figure 2.2: An information structure

Here, state ω summarizes all the relevant facts of the world. Note that only one of the states is the true state of the world; all the other states are hypothetical states needed to encode the players' beliefs. If the true state is ω, player i is informed that the true state is in $I_i(\omega)$, and he does not get any other information. Such an information structure arises if each player observes a state-dependent signal, where $I_i(\omega)$ is the set of states in which the value of the signal of player i is identical to the value of the signal at state ω. The mapping $\boldsymbol{s} : \Omega \to S$ gives Game Theoretical meaning to the abstract state space, by identifying the strategy profile $\boldsymbol{s}(\omega)$ played at each state ω. The mapping s_i is constant over each information set $I_i(\omega)$ in order to ensure that players know their own strategies.

Example 2.1 *For $N = \{1, 2\}$, consider the information structure in Figure 2.2. Here, the state space is $\Omega = \{\omega_0, \omega_1, \omega_2\}$. The information sets of players 1 and 2 are depicted by rounded and regular rectangles, respectively. For example, at ω_0, player 1 finds ω_1 possible and rules out ω_2. Similarly, player 2 rules out ω_1. At each state ω, the probabilities $p_{1,\omega}(\omega)$ and $p_{2,\omega}(\omega)$ that players 1 and 2 assign to the true state ω are depicted in the figure, in the given order. For example, at state ω_1, player 1 knows that the state is ω_1 and assigns probability 1 on ω_1, while player 2 assigns probability 1/2 on ω_1 and probability 1/2 on ω_0.*

Definition 2.12 *An information structure $(\Omega, (I_j)_{j\in N}, (p_{j,\omega})_{j\in N, \omega\in\Omega})$ is said to* admit a common prior *$P \in \Delta(\Omega)$ if*

$$p_{i,\omega} = P(\cdot|I_i(\omega)) \qquad \forall i, \omega. \tag{CPA}$$

In that case, the information structure is denoted as $(\Omega, (I_j)_{j\in N}, P)$*. A* common-prior model *is any model* $M = (\Omega, (I_j)_{j\in N}, P, \boldsymbol{s})$ *in which information structure admits a common prior.*

Here, $P(\cdot|I_i(\omega))$ is a conditional probability distribution on Ω given information set $I_i(\omega)$. Since it puts probability 1 on $I_i(\omega)$ it is also viewed as a probability distribution on $I_i(\omega)$. Recall that when Ω is finite and $P(I_i(\omega)) > 0$,

$$p_{i,\omega}(\omega') = P(\omega'|I_i(\omega)) = P(\{\omega'\})/P(I_i(\omega)) \qquad (\forall \omega' \in I_i(\omega)). \tag{2.7}$$

Example 2.2 *The information structure in Figure 2.2 admits a common prior* P *where*

$$P(\omega_0) = P(\omega_1) = P(\omega_2) = 1/3.$$

For example, $p_{1,\omega_0}(\omega_0) = P(\omega_0|\{\omega_0, \omega_1\}) = \frac{1}{3}/\left(\frac{1}{3} + \frac{1}{3}\right) = 1/2.$

The condition (CPA) in the above definition is called the Common-Prior Assumption. It states that all differences in beliefs can be attributed to differences in information. The common-prior assumption has been made throughout traditional economic models in regards to underlying economic environment. Here, the assumption is made about the players' beliefs about the strategies. Since the beliefs about strategies are highly subjective and there is no physical ex-ante stage, some may find the common-prior assumption especially unwarranted in the current setup (see Gul's (1998) critique and Aumann's (1998) response). This assumption is also made in Nash equilibrium. For us, the assumption (CPA) provides a good transition from rationalizability to Nash equilibrium, providing a middle solution concept (namely, correlated equilibrium) that puts more discipline in players' beliefs without requiring that the players know the other players' strategies.

Definition 2.13 (Correlated Equilibrium) *A* correlated equilibrium *is any common-prior model* $M = (\Omega, (I_j)_{j\in N}, P, \boldsymbol{s})$ *in which*

$$\boldsymbol{s}_i(\omega) \in B_i\left(P(\cdot|I_i(\omega)) \circ \boldsymbol{s}_{-i}^{-1}\right)$$

for all i *and* ω*.*

The displayed condition can be spelled as follows. The probability distribution $P(\cdot|I_i(\omega)) \circ \boldsymbol{s}_{-i}^{-1}$ is the probability distribution induced on the other players' strategy set S_{-i} by the belief $p_{i,\omega}$ of player i at state ω and the mapping $\boldsymbol{s}_i$. Hence, this is the belief of player i at state ω about the other players' strategies. The displayed condition requires that the strategy $\boldsymbol{s}_i(\omega)$ played by player i at state ω must be a best response to her belief about the other players' strategies at that state. That is, player i is rational at state ω. Since this is required at all states, we call it common knowledge of rationality (of player i). Of course, what distinguishes correlated equilibrium is that the model has common prior. In plain English, a correlated equilibrium is a theory about behavior that assumes common knowledge of rationality and a common prior. As such, the correlated equilibria characterize the implications of common knowledge of rationality and common-prior assumptions.

It is useful to spell out what a correlated equilibrium involves:

1. Information structure $(\Omega, (I_j)_{j\in N}, P)$ admits a common prior.
2. Each player i knows his own strategy. That is, $\boldsymbol{s}_i$ is adapted (i.e. it is constant over each information set $I_i(\omega)$).
3. Each player i plays a best response to other players' strategies at each ω.

Example 2.3 *Consider the game*

	a	b
a	$5,1$	$0,0$
b	$4,4$	$1,5$

(2.8)

and the epistemic model $M = (\Omega, I, P, \boldsymbol{s})$ *in Figure 2.2, where* $\Omega = \{\omega_0, \omega_1, \omega_2\}$, $I_1(\omega_0) = \{\omega_0, \omega_1\}$, $I_2(\omega_0) = \{\omega_0, \omega_2\}$, $P(\omega) = 1/3$ *for each* ω, *and* $\boldsymbol{s}(\omega_0) = (b,a)$, $\boldsymbol{s}(\omega_1) = (b,b)$, $\boldsymbol{s}(\omega_2) = (a,a)$. *Observe that each player plays a best response at each state. For example, at* ω_2, *Player 1 assigns probability 1 on* $s_2 = a$, *and plays* $\boldsymbol{s}_2(\omega_2) = a$. *Since* a *is a best response to* a, *her action is a best response. Likewise, at states* ω_0 *and* ω_1, *she assigns probability 1/2 on* $s_2 = a$ *and probability 1/2 on* $s_2 = b$. *Under this belief, her payoff from either strategy is 5/2 and she is indifferent between her strategies. She plays* b *as a best response. Therefore,* M *is a correlated equilibrium of game (2.8).*

Definition 2.14 *A* correlated equilibrium distribution *is a probability distribution*

$$q = P \circ \boldsymbol{s}^{-1}$$

on the set S of strategy profiles induced by a correlated equilibrium $(\Omega, (I_j)_{j\in N}, P, \boldsymbol{s})$.

Note that when S is finite, the probability of a strategy profile s is simply the probability of states at which s is played:

$$q(s) = P(\{\omega | \boldsymbol{s}(\omega) = s\}).$$

For example, the correlated equilibrium in Example 2.3 induces the following correlated equilibrium distribution for game (2.8):

$$\begin{aligned} q(a,a) &= P(\omega_2) = 1/3 \\ q(a,b) &= 0 \\ q(b,a) &= P(\omega_0) = 1/3 \\ q(b,b) &= P(\omega_1) = 1/3. \end{aligned} \tag{2.9}$$

The set of correlated equilibrium distributions has a very convenient structure. It is a convex and compact subset of $\Delta(S)$, characterized by a set of linear inequalities. Towards establishing this structure, for any $\alpha \in [0,1]$ and any two models $(\Omega, (I_j)_{j\in N}, P, \boldsymbol{s})$ and $\left(\Omega', (I'_j)_{j\in N}, P', \boldsymbol{s}'\right)$ with disjoint state spaces, define a mixture model $\left(\tilde{\Omega}, (\tilde{I}_j)_{j\in N}, \tilde{P}, \tilde{\boldsymbol{s}}\right)$ by

$$\begin{aligned} \tilde{\Omega} &= \Omega \cup \Omega' \\ \tilde{I}_j(\omega) &= \begin{cases} I_j(\omega) & \text{if } \omega \in \Omega \\ I'_j(\omega) & \text{if } \omega \in \Omega' \end{cases} \\ \tilde{\boldsymbol{s}}(\omega) &= \begin{cases} \boldsymbol{s}(\omega) & \text{if } \omega \in \Omega \\ \boldsymbol{s}'(\omega) & \text{if } \omega \in \Omega' \end{cases} \\ \tilde{P}(F) &= \alpha P(F \cap \Omega) + (1-\alpha) P(F \cap \Omega'). \end{aligned}$$

Observe that $\left(\tilde{\Omega}, (\tilde{I}_j)_{j\in N}, \tilde{P}, \tilde{\boldsymbol{s}}\right)$ has a further ex-ante stage at which one of the information structures is randomly selected.

Exercise 2.3 *Show that a mixture $\left(\tilde{\Omega}, (\tilde{I}_j)_{j\in N}, \tilde{P}, \tilde{\boldsymbol{s}}\right)$ of any two correlated equilibria is also a correlated equilibrium. Conclude that the set of correlated equilibrium distributions is convex.*

The following characterization of correlated equilibrium distributions is generally true; it is stated for finite games for expositional clarity.

Theorem 2.7 *For any finite (N, S, u), a probability distribution p on S is a correlated equilibrium distribution if and only if for each s_i with $p(s_i, s_{-i}) > 0$ for some s_{-i} and for each s'_i,*

$$\sum_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i})\, p(s_{-i}|s_i) \geq \sum_{s_{-i} \in S_{-i}} u_i(s'_i, s_{-i})\, p(s_{-i}|s_i). \tag{2.10}$$

The condition (2.10) is called the *obedience* condition. To see its logic, suppose that a disinterested moderator randomly selects a strategy profile s from the distribution p and recommends each player i to play s_i without giving any other information. Hearing the recommendation, player comes to believe that the other players' strategies are distributed by $p(\cdot|s_i)$. The obedience condition states that he follows the recommendation.

Formally, this corresponds to the simple model

$$\begin{aligned} \Omega &= S \\ I_i(s) &= \{s_i\} \times S_{-i} = \left\{ (s_i, s'_{-i}) \,|\, s'_{-i} \in S_{-i} \right\} \\ \boldsymbol{s}(s) &= s \end{aligned} \tag{2.11}$$

with common prior p. The condition (2.10) states that each player i is rational at each state s. Hence, (2.10) is a sufficient condition for a correlated equilibrium distribution. Conversely, in order to capture probability distributions induced by correlated equilibria with respect to arbitrary information structures, it suffices to consider this limited set of information structures. To see this, take any correlated equilibrium $((\Omega, I_1, \ldots, I_n, p), \boldsymbol{s})$. The distribution $\tilde{p}$ induced by $((\Omega, I_1, \ldots, I_n, p), \boldsymbol{s})$ on S is given by $\tilde{p}(s) = \sum_{\omega \in \Omega, \mathbf{s}(\omega) = s} p(\omega)$. Now suppose that instead of letting i know that the true state is in $I_i(\omega)$, we only inform him that he needs to play $\boldsymbol{s}_i(\omega)$ according to $\boldsymbol{s}_i$. Since he did not have an incentive to deviate under any information (by definition of correlated equilibrium), by the sure-thing principle,[7] he does not have an incentive to deviate. Hence, the new model with limited information is also a correlated equilibrium. Since u_i does not depend on ω, the latter information structure can be represented by (2.11).

[7] The Sure-Thing Principle is the main axiom for expected utility theory. It states that if a decision maker prefers a to b when she learns that an event occurs and she also prefers a to b when she learns that the event does not occur, she must then prefer a to b without any information.

Thanks to Theorem 2.7, the set of correlated equilibrium distributions is characterized by a finite set of linear inequalities.

Corollary 2.2 *For any finite (N, S, u), the set $CE \subset \Delta(S)$ of correlated equilibrium distributions is the set of solutions p to*

$$\sum_{s_{-i} \in S_{-i}} \left(u_i\left(s_i, s_{-i}\right) - u_i\left(s_i', s_{-i}\right)\right) p\left(s\right) \geq 0 \qquad \left(\forall i, s_i, s_i'\right). \tag{2.12}$$

It is compact and convex.

In practice, one uses (2.12) to study correlated equilibria, and it is often taken as the definition of correlated equilibrium, keeping in mind that it is induced by model as in (2.11). The next example illustrates how to compute the correlated equilibrium distributions.

Example 2.4 *Consider the following Stag Hunt game*

	Stag	*Hare*
Stag	*3,3*	*0,2*
Hare	*2,0*	*2,2*

in which two hunters face a choice between hunting a stag or a hare. Here, hare corresponds to a safe option that does not require the cooperation of the other party, while stag corresponds to a better option that requires all parties' cooperation. Such coordination and cooperation motives present in many situations as we will see in many later applications. Note that there are three Nash equilibria: (Stag, Stag), (Hare, Hare) and a mixed strategy equilibrium in which each player plays Stag with probability 2/3. To compute correlated equilibrium distributions, denote a distribution on S by the following table:

	Stag	*Hare*
Stag	p	q
Hare	r	$1-p-q-r$

where p is the probability of (Stag, Stag) for example. By (2.12), this is a correlated *equilibrium distribution if and only if*

$$\begin{aligned} p &\geq 2q && \text{(Player 1, Stag)} \\ p &\geq 2r && \text{(Player 2, Stag)} \\ 2 &\geq 2p + 2q + 3r && \text{(Player 1, Hare)} \\ 2 &\geq 2p + 2q + 3r. && \text{(Player 2, Here)} \end{aligned}$$

Each inequality is corresponds to the obedience condition for the indicated player and strategy. For example, $p \geq 2q$ *is the condition that player 1 does not deviate when he is asked to play Stag. By deviating, he loses 1 if the other player plays Stag and gains 2 if the other player plays Hare. We multiply these amounts by the ex-ante probabilities of (Stag,Stag) and (Stag, Hare), respectively. Next consider the symmetric correlated equilibria, where* $q = r$. *Such symmetric distributions can be represented by pairs* (p, q). *Since this is a probability distribution, the pair* (p, q) *must satisfy*

$$p + 2q \leq 1.$$

The above conditions reduces to

$$\begin{aligned} q &\leq p/2 \\ 2p + 5q &\leq 2. \end{aligned}$$

The set of correlated equilibrium distributions is the shaded area in Figure 2.3. Note that the Nash equilibria are also among the symmetric correlated equilibrium distributions where $(1, 0)$ *is (Stag, Stag),* $(0, 0)$ *is (Hare, Hare) and* $(4/9, 2/9)$ *is the mixed strategy equilibrium.*

Note that the set of symmetric correlated equilibrium distributions is simply the convex hull of the Nash equilibrium distributions. Of course, there are also asymmetric correlated equilibria, and they are outside of the convex hull of Nash equilibria. In general, under broad conditions, the Nash equilibria are located on the boundary of the set of correlated equilibrium distributions as in this example. Notice also that if $q > 0$, then $p > 0$ and $p + 2q < 1$. That is, both (Stag, Stag) and (Hare, Hare) are played with positive probability. It is illustrative to see that this is true in all correlated equilibria (including asymmetric ones). To see this consider a distribution

	Stag	Hare
Stag	p	q
Hare	r	0

where $q > 0$. Now if player 2 is recommended to play Hare, then she knows that player 1 plays Stag, in which case her best response must be Stag, violating the obedience condition. For a correlated equilibrium, q must be zero. Of course, if $q = 0$, then player 1 must also play Stag with probability 1, leading to the (Stag, Stag) Nash equilibrium. Note that, for (Stag, Stag) equilibrium, ex-ante obedience condition for Hare in Corollary 2.2 is satisfied (as $2 = 2$) because it puts zero probability on the above violation. That is why the obedience condition applies to all strategies in Corollary 2.2.

Exercise 2.4 *Show that if p and q are Nash equilibrium distributions, then any mixture $\alpha p + (1 - \alpha) q$ is a correlated equilibrium distribution, where $\alpha \in [0, 1]$.*

Exercise 2.5 *Compute the set of correlated equilibrium distributions in game (2.8). Compute also the set of expected utility pairs for these distributions and plot them.*

Existence and Continuity Properties

Existence of a correlated equilibrium follows from existence of a Nash equilibrium—as it will be clearer in the next section. The set of correlated equilibrium distributions is also upperhemicontinuous—thanks to Corollary 2.2. Formally, consider space X of some payoff-relevant parameters. Fix a set N of players and a finite set S of strategy profiles. The utility function of each player i depends on x as well as s. That is, $u_i : S \times X \to \mathbb{R}$. The utility functions are assumed to be continuous in the parameters. Write $CE(x)$ for the set of all correlated equilibrium distributions of game $(N, S, u(\cdot; x))$ in which it is common knowledge that the parameter value is x.

Corollary 2.3 *If S is a compact metric space and each u_i is continuous (in (s, x)), then CE is compact-valued and upperhemicontinuous.*

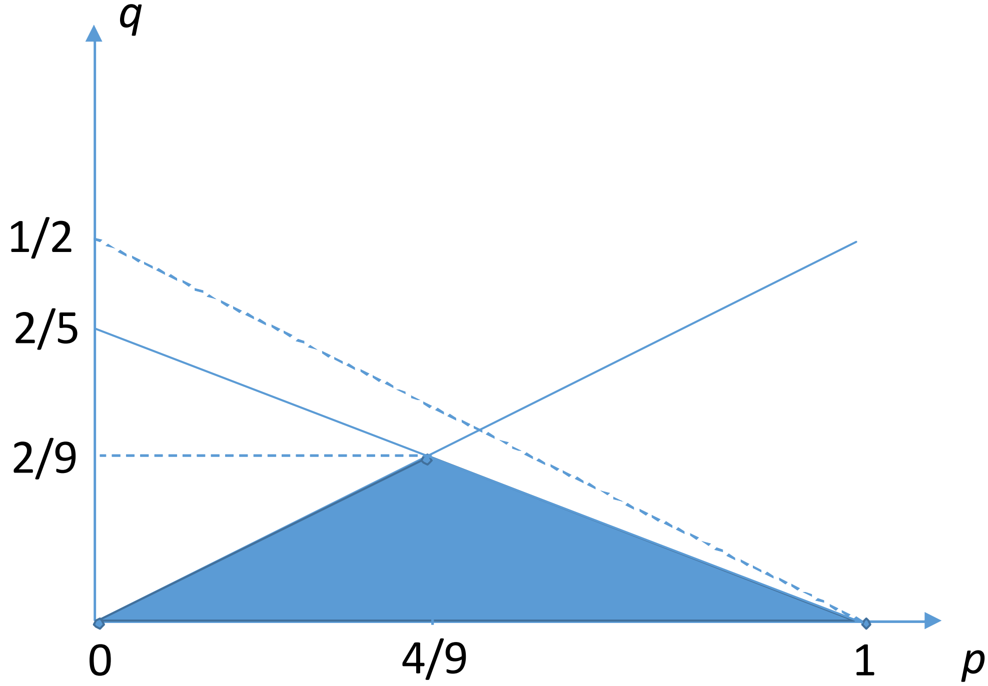

Figure 2.3: Symmetric correlated equilibria and Nash equilibria in Stag Hunt game.

2.4 Relation Between Nash Equilibrium, Correlated Equilibrium, and Rationalizability

Rationalizability corresponds to common-knowledge of rationality, the assumptions that are made in the definition of the game. Correlated equilibrium further assumes that the beliefs can be viewed as coming from a common prior. This is clearly a further restriction, and hence correlated equilibrium is a refinement of rationalizability leading to stronger predictions. Finally, (mixed strategy) Nash equilibrium puts even a further restriction by assuming that the strategies of players are stochastically independent. As such, it is a further refinement of correlated equilibrium distributions. This is formally stated next.

Theorem 2.8 *(1) A distribution $\sigma \in \Delta(S)$ is a Nash equilibrium distribution if and only if σ is a correlated equilibrium distribution and $\sigma = \sigma_1 \times \cdots \times \sigma_n$ for some $\sigma_1 \in \Delta(S_1)$, ..., $\sigma_n \in \Delta(S_n)$. (2) For any correlated equilibrium $(\Omega, (I_j)_{j \in N}, P, \boldsymbol{s})$,*

$$\boldsymbol{s}(\Omega) \subseteq S^{\infty}.$$

(3) Under any correlated equilibrium distribution σ,

$$\sigma(S^{\infty}) = 1.$$

Exercise 2.6 *Prove Theorem 2.8 for finite games.*

The first part states that Nash equilibria are correlated equilibria in which the players strategies are stochastically independent. By relaxing such independence assumption, correlated equilibrium results in a larger set of solutions—and a weaker solution concept. Since the convex combination of correlated equilibrium distributions are also correlated equilibrium distribution, the convex hull of Nash equilibrium distributions remains within the set of correlated equilibrium distributions as in Figure 2.3. The latter set can be strictly larger. For example, the correlated equilibrium in (2.9) is outside of the Nash equilibrium distributions for game (2.8). Indeed, the expected payoff vector for this distribution is $(10/3, 10/3)$, where players get $20/3$ in total, while the sum of payoffs can be at most 6 in any Nash equilibrium.

The second part states that only rationalizable strategies can be played in a correlated equilibrium. This part immediately follows from Theorem 10.2 because it is common knowledge that everybody is rational in model $(\Omega, (I_j)_{j \in N}, P, \boldsymbol{s})$. In turn, the last part follows from the second part immediately: every player plays a rationalizable strategy with probability 1 under any correlated equilibrium. In particular, every player must play a rationalizable strategy in every Nash equilibrium.

Exercise 2.7 *Find a finite game (N, S, u) and a strategy profile $s \in S^\infty$ such that no correlated equilibrium assigns positive probability on s.*

2.5 Classical Applications

Cournot Competition

Consider n firms. Each firm i produces $q_i \geq 0$ units of a good at marginal cost $c \geq 0$ and sell it at price

$$P = \max\{1 - Q, 0\} \tag{2.13}$$

where

$$Q = q_1 + \cdots + q_n \tag{2.14}$$

is the total supply. Each firm maximizes the expected profit. Hence, the payoff of firm i is

$$\pi_i = q_i (P - c). \tag{2.15}$$

Assuming all of the above is commonly known, we can write this as a game in normal form, by setting

- $N = \{1, 2, \ldots, n\}$ as the set of players
- $S_i = [0, \infty)$ as the strategy space of player i, where a typical strategy is the quantity q_i produced by firm i, and
- $\pi_i : S_1 \times \cdots \times S_n \to \mathbb{R}$ as the payoff function.

Best Response It is useful to know the best response of a firm i to the production levels of the other firms. Write

$$Q_{-i} = \sum_{j \neq i} q_j \tag{2.16}$$

for the total supply of the firms other than firm i. If $Q_{-i} > 1$, then the price $P = 0$ and the best firm i can do is to produce zero and obtain zero profit. Now assume $Q_{-i} \leq 1$. For any $q_i \in (0, 1 - Q_{-i})$, the profit of the firm i is

$$\pi_i (q_i, Q_{-i}) = q_i (1 - q_i - Q_{-i} - c). \tag{2.17}$$

(The profit is negative if $q_i > 1 - Q_{-i}$.) By setting the derivative of π_i with respect to q_i to zero, we obtain the best production level

$$q_i^B (Q_{-i}) = \max \left\{ \frac{1 - Q_{-i} - c}{2}, 0 \right\}. \tag{2.18}$$

The best response functions are plotted in Figure 2.4.

Cournot Duopoly

Now, consider the case of two firms. In that case, for $i \neq j$, we have $Q_{-i} = q_j$.

Nash Equilibrium In order to have a Nash equilibrium, we must have

$$q_1 = q_1^B (q_2) \equiv \frac{1 - q_2 - c}{2}$$

and

$$q_2 = q_2^B (q_1) \equiv \frac{1 - q_1 - c}{2}.$$

Solving these two equations simultaneously, one obtains

$$q_1^* = q_2^* = \frac{1 - c}{3}$$

as the only Nash equilibrium. Graphically, as in Figure 2.4, one can plot the best response functions of each firm and identify the intersections of the graphs of these functions as Nash equilibria. In this case, there is a unique intersection, and therefore there is a unique Nash equilibrium.

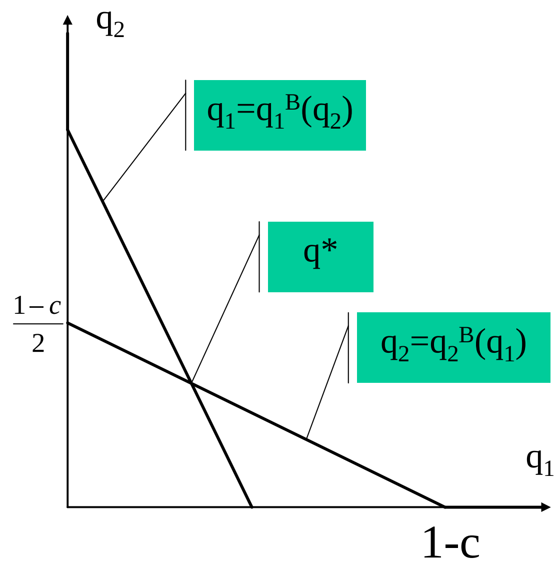

Figure 2.4:

Rationalizability The linear Cournot duopoly game is "dominance solvable" i.e. there is a unique rationalizable strategy. Let us first consider the first couple rounds of elimination to see this intuitively; I will then show mathematically that this is indeed the case.

Round 1 Any strategy $\hat{q}_i > (1-c)/2$ is strictly dominated by $(1-c)/2$. To see this, consider any q_j. Profit $\pi_i(q_i, q_j)$ is strictly increasing until $q_i = (1-c-q_j)/2$ and strictly decreasing thereafter. In particular, since $\hat{q}_i > (1-c)/2 \geq (1-c-q_j)/2$,

$$\pi_i((1-c)/2, q_j) > \pi_i(\hat{q}_i, q_j),$$

showing that $\hat{q}_i$ is strictly dominated by $(1-c)/2$. We therefore eliminate all $\hat{q}_i > (1-c)/2$ for each player i. The resulting strategies are as in Figure 2.5, where the shaded area is eliminated.

Round 2 In the remaining game $q_j \leq (1-c)/2$. Consequently, any strategy $\hat{q}_i < (1-c)/4$ is strictly dominated by $(1-c)/4$. To see this, take any $q_j \leq (1-c)/2$ and recall that π_i is strictly increasing until $q_i = (1-c-q_j)/2$, which is greater than

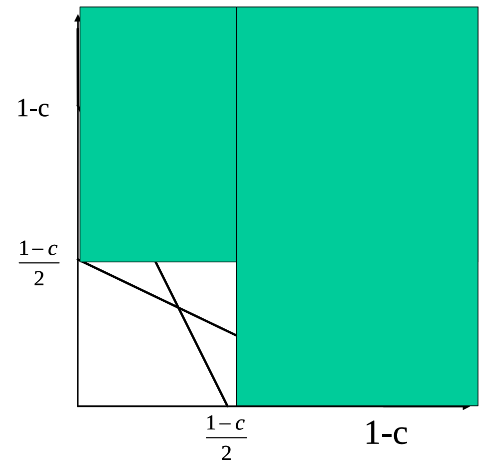

Figure 2.5: Remaining strategies after 1 round of iteration in Cournot Duopoly

or equal to $(1-c)/4$. Hence,

$$\pi_i(\hat{q}_i, q_j) < \pi_i((1-c)/4, q_j),$$

showing that $\hat{q}_i$ is strictly dominated by $(1-c)/4$. We will therefore eliminate all $\hat{q}_i$ with $\hat{q}_i < (1-c)/4$. The remaining strategies are as in Figure 2.6. The remaining game is a smaller replica of the original game. Applying the same procedure repeatedly we eliminate all strategies except for the Nash equilibrium. (After every two rounds, we obtain a smaller replica.) Therefore, the only rationalizable strategy is the unique Nash equilibrium strategy:

$$q_i^* = (1-c)/3.$$

A more formal treatment We can prove this more formally by invoking the following lemma repeatedly:

Lemma 2.1 *Given that $q_j \le \bar{q}$, every strategy $\hat{q}_i$ with $\hat{q}_i < q_i^B(\bar{q})$ is strictly dominated by $q_i^B(\bar{q}) \equiv (1-\bar{q}-c)/2$. Given that $q_j \ge \bar{q}$, every strategy $\hat{q}_i$ with $\hat{q}_i > q_i^B(\bar{q})$ is strictly dominated by $q_i^B(\bar{q}) \equiv (1-\bar{q}-c)/2$.*

Proof. Let's first prove the first statement. Take any $q_j \le \bar{q}$. Note that $\pi_i(q_i; q_j)$ is strictly increasing in q_i at any $q_i < q_i^B(q_j)$. Since $\hat{q}_i < q_i^B(\bar{q}) \le q_i^B(q_j)$,[8] this implies

[8] This is because q_i^B is decreasing.

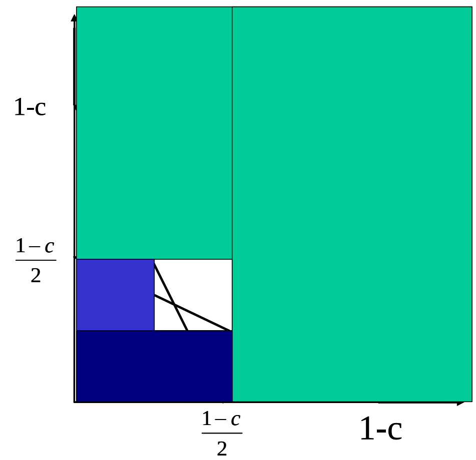

Figure 2.6: Remaining strategies after 2 rounds of elimination in Cournot Duopoly

that

$$\pi_i(\hat{q}_i, q_j) < \pi_i\left(q_i^B(\bar{q}), q_j\right).$$

That is, $\hat{q}_i$ is strictly dominated by $q_i^B(\bar{q})$.

To prove the second statement, take any $q_j \leq \bar{q}$. Note that $\pi_i(q_i; q_j)$ is strictly decreasing in q_i at any $q_i > q_i^B(q_j)$. Since $q_i^B(q_j) \leq q_i^B(\bar{q}) < \hat{q}_i$, this implies that

$$\pi_i(\hat{q}_i, q_j) < \pi_i\left(q_i^B(\bar{q}), q_j\right).$$

That is, $\hat{q}_i$ is strictly dominated by $q_i^B(\bar{q})$. ■

Now, define a sequence $q^0, q^1, q^2, \ldots$ by $q^0 = 0$ and

$$q^m = q_i^B\left(q^{m-1}\right) \equiv \left(1 - q^{m-1} - c\right)/2 = (1-c)/2 - q^{m-1}/2$$

for all $m > 0$. That is,

$$\begin{aligned}
q^0 &= 0 \\
q^1 &= \frac{1-c}{2} \\
q^2 &= \frac{1-c}{2} - \frac{1-c}{4} \\
q^3 &= \frac{1-c}{2} - \frac{1-c}{4} + \frac{1-c}{8} \\
&\ldots \\
q^m &= \frac{1-c}{2} - \frac{1-c}{4} + \frac{1-c}{8} - \cdots - (-1)^m \frac{1-c}{2^m} \qquad (2.19) \\
&\ldots
\end{aligned}$$

The sequence above is a geometric series in which the contribution of each term is as much as the contribution of the previous term. It converges to $(1-c)/3$.

Theorem 2.9 *The set of remaining strategies after any odd round m ($m = 1, 3, \ldots$) is $[q^{m-1}, q^m]$. The set of remaining strategies after any even round m ($m = 2, 4, \ldots$) is $[q^m, q^{m-1}]$. The set of rationalizable strategies is $\{(1-c)/3\}$.*

Proof. We use mathematical induction on m. For $m = 1$, we have already proven the statement. Assume that the statement is true for some odd m. Then, for any q_j available at even round $m+1$, we have $q^{m-1} \leq q_j \leq q^m$. Hence, by Lemma 2.1, any $\hat{q}_i < q_i^B(q^m) = q^{m+1}$ is strictly dominated by q^{m+1} and eliminated. That is, if q_i survives round $m+1$, then $q^{m+1} \leq q_i \leq q^m$. On the other hand, every $q_i \in [q^{m+1}, q^m] = \left[q_i^B(q^m), q_i^B(q^{m-1})\right]$ is a best response to some q_j with $q^{m-1} \leq q_j \leq q^m$, and it is not eliminated. Therefore, the set of strategies that survive the even round $m+1$ is $[q^{m+1}, q^m]$.

Now, assume that the statement is true for some even m. Then, for any q_j available at odd round $m+1$, we have $q^m \leq q_j \leq q^{m-1}$. Hence, by Lemma 2.1, any $\hat{q}_i > q_i^B(q^m) = q^{m+1}$ is strictly dominated by q^{m+1} and eliminated. Moreover, every $q_i \in [q^m, q^{m+1}] = \left[q_i^B(q^{m-1}), q_i^B(q^m)\right]$ is a best response to some q_j with $q^m \leq q_j \leq q^{m-1}$, and it is not eliminated. Therefore, the set of strategies that survive the odd round $m+1$ is $[q^m, q^{m+1}]$.

Finally, notice that

$$\lim_{m\to\infty} q^m = (1-c)/3.$$

Therefore the intersections of the above intervals is $\{(1-c)/3\}$, which is the set of rationalizable strategies. ■

Cournot Oligopoly

We will now consider the case of three or more firms. When there are three or more firms, rationalizability does not help, i.e., we cannot eliminate any strategy less than the monopoly production $q^1 = (1-c)/2$.

In the first round we eliminate any strategy $q_i > (1-c)/2$, using the same argument in the case of duopoly. But in the second round, the maximum possible total supply by the other firms is

$$(n-1)(1-c)/2 \geq 1-c,$$

where n is the number of firms. The best response to this aggregate supply level is 0. Hence, we cannot eliminate any strategy in round 2. The elimination process stops, yielding $[0, (1-c)/2]$ as the set of rationalizable strategies.

Of course, Cournot oligopoly has a unique Nash equilibrium as in the Cournot duopoly. While the Nash equilibrium remains to make strong predictions as we introduce new firms, the predictions of rationalizability become rather weak. (In equilibrium analysis the weak predictions of rationalizability reappears as instability of equilibrium, making equilibrium behavior highly sensitive to the specification of beliefs.)

Bertrand Duopoly

Consider two firms. Simultaneously, each firm i sets a price p_i. The firm i with the lower price $p_i < p_j$ sells $1 - p_i$ units and the other firm cannot sell any. If the firms set the same price, the demand is divided between them equally. That is, the amount of sales for firm i is

$$Q_i(p_1, p_2) = \begin{cases} 1 - p_i & \text{if } p_i < p_j \\ \frac{1-p_i}{2} & \text{if } p_i = p_j \\ 0 & \text{otherwise.} \end{cases}$$

We assume that it costs noting to produce the good (i.e. $c = 0$). Therefore, the profit of a firm i is

$$\pi_i(p_1, p_2) = p_i Q_i(p_1, p_2) = \begin{cases} (1 - p_i) p_i & \text{if } p_i < p_j \\ \frac{(1-p_i)p_i}{2} & \text{if } p_i = p_j \\ 0 & \text{otherwise.} \end{cases}$$

Assuming all of the above is commonly known, we can write this formally as a game in normal form by setting

- $N = \{1, 2\}$ as the set of players
- $S_i = [0, \infty)$ as the set of strategies for each i, with price p_i a typical strategy,
- π_i as the utility function.

Observe that when $p_j = 0$, $\pi_i(p_1, p_2) = 0$ for every p_i, and hence every p_i is a best response to $p_j = 0$. This has two important implications:

1. Every strategy is rationalizable (we cannot eliminate any strategy because each of them is a best reply to zero).

2. $p_1^* = p_2^* = 0$ is a Nash equilibrium.

This is indeed the only Nash equilibrium. Hence, even with two firms, when the firms compete by setting prices, the competitive equilibrium emerges. If the game is modified slightly by discretizing the set of allowable prices and putting a minimum price, then the game becomes dominance-solvable, i.e., only one strategy remains rationalizable. In the modified game, the minimum price is the only rationalizable strategy, as in competitive equilibrium. The model is highly sensitive to small search costs on the part of consumers: the equilibrium behavior is dramatically different from the equilibrium behavior in the original game and competitive equilibrium.

Exercise 2.8 *Show that Bertrand duopoly is better-reply secure.*

Differentiated Bertrand Competition

The Bertrand duopoly above assumes that the goods produced by the two firms are perfect substitutes. We next consider differentiated price competition with n firms, $1, \ldots, n$. Each firm i simultaneously sets a price p_i, and the demand for the product of firm i is $Q_i(p_1, \ldots, p_n)$ where Q_i is decreasing in p_i and increasing in p_j for each $j \neq i$. The profit (the payoff) of firm i is $p_i Q_i$. In this example, further assume that

$$Q_i = 1 - p_i + \frac{b}{n-1} \sum_{j \neq i} p_j$$

where $0 < b < 1$. There is a unique Nash equilibrium, in which each firm sets price

$$p^* = \frac{1}{2+b}.$$

When the firms can set any price, any price $p \geq p^*$ is rationalizable.

Exercise 2.9 *Suppose that the set of allowable prices is* $[0, \bar{p}]$ *for some* $\bar{p} > p^*$*. Compute the set of rationalizable strategies.*

2.6 Notes on Literature

In his dissertation, Nash (1950) introduced the non-cooperative game theory, defined the concept of Nash equilibrium, and proved that it exists in finite games using the proof technique in Theorem 2.4. Nash also considered iterated elimination of strictly dominated strategies—as a means to compute the set Nash equilibria—and analyzed the incomplete information game of three-person poker. Harsanyi (1967) showed that Nash's formulation can be extended to arbitrary games of incomplete information using type spaces, as we will see in the next chapter.

Glicksberg (1952) extended Nash's existence theorem to compact and continuous games (cf. Assumption 1.1). Dasgupta and Maskin (1986) studied existence of equilibrium in discontinuous game. The existence theorem and the concepts in Section 2.2.3 are due to Reny (1999).

Rationalizability has been introduced by Bernheim (1984) and Pearce (1984) in their dissertations. They have in addition assumed that the beliefs do not put correlation between different players' strategies. They relate rationalizability to the iterative elimination of strictly dominated strategies. Theorem 2.1, which characterizes rationality with undominated strategies, is sometimes referred to as Pearce's Lemma.

Aumann (1987) introduced the solution concept of correlated equilibrium and showed that it captures the idea of common knowledge of rationality under common-prior assumption using this formulation.

Cournot (1838) and Bertrand (1883) applied Nash's equilibrium analysis to their oligopoly models a century before Nash. Moulin (1984) shows that dominance solvability and Cournot stability are logically equivalent in two-player nice games, which include Cournot duopoly and differentiated Bertrand duopoly. Dominance solvability of these games can be obtained from his results. The analyses of oligopoly models in Section 2.5 are confined to Nash equilibrium and rationalizability. Correlated equilibria of these games coincide with Nash equilibria (Liu (1996), Jann and Schottmüller (2015)). In Bertrand competition, the solution is unique if and only if the monopolist's profit is finite; otherwise there can be mixed strategy equilibria with arbitrarily high prices.

2.7 Exercises

Exercise 2.10 *Consider the following complete unformation game* $G = (N, A, u)$*:*

	w	x	y	z
a	$0,3$	$1,2$	$2,1$	$3,0$
b	$4,0$	$2,3$	$1,2$	$0,1$
c	$1,0$	$2,0$	$2,3$	$1,2$
d	$2,0$	$1,0$	$1,0$	$1,3$

Compute the sets of rationalizable strategies, Nash equilibria, and the correlated equilibrium distributions.

Exercise 2.11 *Compute the set* CE *of correlated equilibrium distributions for game (2.8). Find the set* C *of corners of* CE*, where* CE *is the convex hull of* C *and if* $q = \alpha p + (1-\alpha) r$ *for some* $\alpha \in (0,1)$*,* $q \in C$ *and* $p, r \in CE$*, then* $p = q = r$*.*

Exercise 2.12 *Consider an information structure* (Ω, I, p) *with finite set of states. Define a* possibility chain *as a sequence* $(\omega_0, i_0, \omega_1, i_1, \ldots, i_{m-1}, \omega_m)$ *of states* ω_k *and players* i_k *such that* $\omega_{k+1} \in I_{i_k}(\omega_k)$ *for each* k*. An event* $F \subseteq \Omega$ *is said to be* possibility-connected *if for any* $\omega, \omega' \in F$ *there exists a possibility chain* $(\omega, i_0, \omega_1, i_1, \ldots, i_{m-1}, \omega')$*. An event* $F \subseteq \Omega$ *is said to be* possibility-closed *if for any* $\omega \in F$ *and any possibility chain* $(\omega, i_0, \omega_1, i_1, \ldots, i_{m-1}, \omega')$*,* $\omega' \in F$*.*

1. *State a possibility chain* $(\omega_0, i_0, \omega_1, i_1, \ldots, i_{m-1}, \omega_m)$ *in plain English.*

2. *For any distinct possibility-connected and possibility-closed events* F *and* F'*, show that* $F \cap F' = \varnothing$*, i.e., such events form a partition of* Ω*.*

3. *Show that, for any possibility-connected and possibility-closed event* F*, any* $\omega \in F$*, and any event* F'*,* F' *is common knowledge at* ω *if and only if* $F \subseteq F'$*. (That is, possibility-closed and possibility connected events form an information partition for common knowledge.)*

4. *Assume* $p_{i,\omega}(\omega') > 0$ *for every* i*,* ω*, and* ω' *with* $\omega' \in I_i(\omega)$*. Show that* (Ω, I, p) *admits a common prior if and only if for any two possibility chains* $(\omega_0, i_0, \omega_1, i_1, \ldots, i_{m-1}, \omega_m)$

and $\left(\omega'_0, i'_0, \omega'_1, i'_1, \ldots, i'_{m'-1}, \omega'_{m'}\right)$ *with* $\omega_0 = \omega'_0$ *and* $\omega_m = \omega'_m$, *we have*

$$\prod_{k=0}^{m-1} \frac{p_{i_k,\omega_k}\left(\omega_{k+1}\right)}{p_{i_k,\omega_k}\left(\omega_k\right)} = \prod_{k=0}^{m'-1} \frac{p_{i'_k,\omega'_k}\left(\omega'_{k+1}\right)}{p_{i'_k,\omega'_k}\left(\omega'_k\right)}.$$

Bibliography

[1] Aumann, Robert J. (1976):. "Agreeing to disagree." *The annals of statistics* 1236-1239.

[2] Aumann, Robert J. (1987): "Correlated equilibrium as an expression of Bayesian rationality." *Econometrica* 1-18.

[3] Aumann, Robert J. (1998): "Common priors: A reply to Gul." Econometrica 66.4 929-938.

[4] Aumann, Robert, and Adam Brandenburger. (1995): "Epistemic conditions for Nash equilibrium." *Econometrica* 1161-1180.

[5] Bernheim, B. Douglas (1984): "Rationalizable strategic behavior." *Econometrica* 1007-1028.

[6] Bertrand, J. (1883). "Review of Theorie mathematique de la richesse sociale and of Recherches sur les principles mathematiques de la theorie des richesses". *Journal des Savants*. 67: 499–508.

[7] Brandenburger, A. and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1391-1402.

[8] Cournot, A. (1838): *Researches into the mathematical principles of the theory of wealth*.

[9] Dasgupta, Partha, and Eric Maskin. (1986): "The existence of equilibrium in discontinuous economic games, I: Theory." *The Review of economic studies* 53.1 1-26.

[10] Glicksberg, Irving L. (1952): "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points." *Proceedings of the American Mathematical Society* 3.1 170-174.

[11] Gul, Faruk (1998): "A comment on Aumann's Bayesian view." *Econometrica* 66.4 923-927.

[12] Harsanyi, J. (1967): "Games with Incomplete Information played by Bayesian Players. Part I: the Basic Model," *Management Science* 14, 159-182.

[13] Jann, Ole, and Christoph Schottmüller. (2015): "Correlated equilibria in homogeneous good Bertrand competition." *Journal of Mathematical Economics* 57 31-37.

[14] Liu, L. (1996). "Correlated equilibrium of Cournot oligopoly competition." *Journal of Economic Theory* 68 (2), 544–548.

[15] Moulin, Herve (1984): "Dominance solvability and Cournot stability." *Mathematical Social Sciences* 7.1 83-102.

[16] Nash, J.F. (1950): *Noncooperative Games*, Ph.D. Dissertation, Department of Mathematics, Princeton University.

[17] Pearce, David G. (1984): "Rationalizable strategic behavior and the problem of perfection." *Econometrica* 1029-1050.

[18] Reny, Philip J. "On the existence of pure and mixed strategy Nash equilibria in discontinuous games." Econometrica 67.5 (1999): 1029-1056.

[19] Tan, Tommy Chin-Chiu, and Sérgio Ribeiro da Costa Werlang (1988): "The Bayesian foundations of solution concepts of games." *Journal of Economic Theory* 45.2 370-391.

Chapter 3

Bayesian Games

In any complete information game, it is assumed that the payoff functions and the strategies available to any player are all common knowledge. Of course, most economic applications involve some form of incomplete information, as one may not know another player's payoffs or beliefs. Since Harsanyi (1967), such situations are formalized as Bayesian games, in which the private information is modeled through types.

Formally, Bayesian games can also be viewed as complete-information games, with some special structure. Hence, they inherit the properties of complete-information games. In this chapter, I exploit this fact to extend the existence and continuity properties of equilibrium in the previous section to Bayesian games. All in all, under equilibrium analysis, Bayesian games will not be any different from complete information games.

In contrast, rationalizability analysis in Bayesian games uncovers many subtle fundamental issues. Indeed, there are at least three well-known notions of rationalizability in Bayesian games, each reflecting a different view of Bayesian games. I will present these three notions. I will show that the weakest of these three concepts, namely Interim Correlated Rationalizability (ICR), characterizes the common knowledge of rationality. I will also present a very general upper-hemicontinuity result for ICR.

3.1 Basic Definitions

Now, I will introduce a formulation of Bayesian games that will be used throughout the course.

Definition 3.1 *A Bayesian game is a list* (N, A, Θ, T, u, p) *where*

- N *is the set of players (with generic members* i, j*)*
- $A = \prod_{i \in N} A_i$ *is the set of action profiles (with generic member* $a = (a_i)_{i \in N}$*)*
- Θ *is a set of payoff parameters* θ
- $T = \prod_{i \in N} T_i$ *is the set of type profiles (with generic member* $t = (t_i)_{i \in N}$*)*
- $u_i : \Theta \times A \to \mathbb{R}$ *is the payoff function of player* i*, and*
- $p_i(\cdot|t_i) \in \Delta(\Theta \times T_{-i})$ *is the belief of type* t_i *about* (θ, t_{-i})*.*

Here, each player i knows his own type t_i and does not necessarily know θ or the other players' types, about which he has a belief $p_i(\cdot|t_i)$. The game is defined in terms of players' interim beliefs $p_i(\cdot|t_i)$, which they obtain after they observe their own type but before taking their action. The game can also be defined by ex-ante beliefs $p_i \in \Delta(\Theta \times T)$ for some belief p_i. The game has a common prior if there exists $\pi \in \Delta(\Theta \times T)$ such that

$$p_i(\cdot|t_i) = \pi(\cdot|t_i) \qquad \forall t_i \in T_i, \forall i \in N \tag{CPA}$$

In that case, the game is simply denoted by $(N, A, \Theta, T, u, \pi)$. In line with Assumption 1.1, I will maintain the following technical assumption in formal results throughout.

Assumption 3.1 *The set* $N = \{1, \ldots, n\}$ *is finite; the sets* A*,* Θ*, and* T *are separable metric spaces, and* A *is also compact. Each utility function* u_i *is continuous.*

Once again, an important special case is that of finite Bayesian games, in which all of the above sets are finite. Another important case is the one in which the sets A, Θ, and T are subsets of $\mathbb{R}^n$ and A is closed and bounded.

Sometimes a Bayesian game is defined by a list (N, A, T, u, π) where utility function

$$u_i : T \times A \to \mathbb{R} \tag{3.1}$$

depended on type profile t and the action profile. Here, the utility function depends explicitly on payoff parameters but not on type profiles. The formulation here is slightly more general. Given a game (N, A, T, u, π) with (3.1), one can simply introduce the set $\Theta = T$ of parameters and a new prior $\hat{\pi}$ on $\Theta \times \ T$ with support on the diagonal $\{(t,t)|t \in T\}$. Conversely, given a game $(N, A, \Theta, T, u, \pi)$ in our formulation with u_i : $\Theta \times A \to \mathbb{R}$, one can define a new utility function $v_i : T \times A \to \mathbb{R}$ by

$$v_i(t,a) = E[u_i(\theta,a)|t] = \int_{\theta\in\Theta} u_i(\theta,a)d\pi(\theta|t).$$

Note, however, that by suppressing the dependence on the payoff parameter θ, the formulation with (3.1) loses some information. Such information is not needed for Bayesian Nash equilibrium, but that information is used by interim correlated rationalizability, the main concept we will introduce in this chapter. Also, the interim formulation here reflects the idea behind incomplete information better.

Sometimes, yet another formulation is used for Bayesian games, which seems to be more general than either of the formulations above. In that formulation one takes the utility functions

$$u_i : \Theta \times T \times A \to \mathbb{R} \tag{3.2}$$

to depend on *both* the payoff parameters and the type profiles. Note that this is equivalent to our formulation. To incorporate the dependence on the type profiles, one can simply take $\tilde{\Theta} = \Theta \times T$ to be the set of payoff functions and the belief function to be $\tilde{p}_i$ defined by

$$\tilde{p}_i\left(\Theta' \times \{t_i\} \times T'_{-i} \times T'_{-i}|t_i\right) = p_i\left(\Theta' \times T'_{-i}|t_i\right)$$

for all $\Theta' \subseteq \Theta$, $T'_{-i} \subseteq T_{-i}$ and $t_i \in T_i$. This ensures that it is common knowledge that players know their own types, which are als incorporated in the underlying payoff parameters.

Meaning of a Type Space When a researcher models an incomplete information, there is often no ex-ante stage or an explicit information structure in which players observe values of some signals. In the modeling stage, each player i has

- some belief $\tau_i^1 \in \Delta(\Theta)$ about the payoffs (and the other aspects of the physical world), a belief that is referred to as *the first-order belief* of i,
- some belief $\tau_i^2 \in \Delta\left(\Theta \times \ \Delta(\Theta)^{N\setminus\{i\}}\right)$ about the payoffs and the other players' first-order beliefs (i.e. $\theta, \tau_{-i}^1)$), a belief that is referred to as *the second-order belief* of i,

- some belief τ_i^3 about the payoffs and the other players' first and second order beliefs (i.e. $\theta, \tau_{-i}^1, \tau_{-i}^2$),
- up to infinity.

(It is an understatement that some of these beliefs may not be fully articulated even in players' own minds.)

Modeling incomplete information directly in this form is considered to be quite difficult. Harsanyi (1967) has proposed a tractable way to model incomplete information through a type space. In this formalization, one models the infinite hierarchy of beliefs above through a type space (Θ, T, p) and a type $t_i \in T_i$ as follows. Given a type t_i and a type space (Θ, T, p), one can compute an infinite hierarchy of beliefs for type t_i. For example for finite $\Theta \times T$, one can compute the first-order belief of type t_i by

$$h_i^1(\cdot|t_i) = \text{marg}_\Theta p(\cdot|t_i),$$

so that

$$h_i^1(\theta|t_i) = \sum_{t-i} p(\theta, t_{-i}|t_i),$$

the second-order belief $h_i^2(\cdot|t_i)$ of type t_i by

$$h_i^2\left(\theta, \hat{h}_{-1}^1|t_i\right) = \sum\{t_{-i}|h_{-i}^1(\cdot|t_{-i}) = \hat{h}_{-i}^1\}p(\theta, t_{-i}|t_i),$$

and so on.[1] A type space (Θ, T, p) and a type $t_i \in T_i$ model a belief hierarchy $(\tau_i^1, \tau_i^2, ...)$ if and only if $h_i^k(\cdot|t_i) = \tau_i^k$ for each k.

It is important to keep in mind that in a type space only one type profile corresponds to the actual incomplete-information situation that is meant to be modeled. All the remaining type profiles are hypothetical situations that are introduced in order to model the players' beliefs.

---

[1]For any probability distribution P on product space $X \times Y$, $\text{marg}_X P$ denotes the marginal distribution of P on X, yielding a probability distribution on X. In general, for any measurable event $X' \subset X$, $\text{marg}_X P(X') = P(X' \times Y)$. When $X \times Y$ is finite, the marginal distribution is computed by

$$\text{marg}_X P(x) = P(\{x\} \times Y) = \sum_{y \in Y} P(x, y).$$

Bayesian Game as a Complete-Information Game A Bayesian game can be viewed as a complete information game. Indeed, there are two complete-information representation of any Bayesian game, each representing a different view of a Bayesian game. First, given any Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, \pi)$ with common prior π and finite type space T, one can define *ex-ante game*

$$G(\mathcal{B}) = (N, S, U)$$

where

$$S_i = A_i^{T_i}$$

is the set of strategies $s_i : T_i \to A_i$ and

$$U_i(s) = E_\pi[u_i(\theta, s(t))]$$

is the ex-ante expected utility from strategy profile s for each $i \in N$ and $s \in S$. (Without the common-prior assumption, one can simply take the expectation with respect to the belief of player i in the definition of U_i.) In this view, a Bayesian game is simply a complete-information game in which the players observe different dimensions of the Nature's moves. Here, types arise from individuals' observations of the Nature's move and the ex-ante stage prior to the observation seems real. In some situations, this may indeed be the case. For example, in a poker game, each player is randomly dealt a set of cards from a deck and each player observes his own cards. Here, a type is any possible combination of cards a player can be dealt. For another example, firms may literally sample the amount of oil under a mountain before bidding for the right to extract the oil in an auction. In genuine cases of incomplete information, however, the ex-ante stage is just a technical device to model the interim beliefs of a type, beliefs that is formed after observing the type.

A second complete-information game takes the interim view, in which the ex-ante stage has no relevance. Consider any Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ with finite T; without loss of generality assume that the types of different players are denoted differently so that $T_i \cap T_j = \emptyset$ for any distinct i and j. One can define *interim game*

$$AG(\mathcal{B}) = \left(\hat{N}, \hat{S}, U\right)$$

where

$$\begin{aligned}\hat{N} &= \bigcup_{i\in N} T_i \\ \hat{S}_{t_i} &= A_i, \\ \hat{U}_{t_i}(\hat{s}) &= E\left[u_i\left(\theta,\hat{s}\right)|p_i(\cdot|t_i)\right] \equiv \sum_{(\theta,t_{-i})} u_i\left(\theta,\hat{s}_{t_i},\hat{s}_{t_{-i}}\right)p_i(\theta,t_{-i}|t_i)\end{aligned}$$

for each $t_i \in \hat{N}$. Here, each type is viewed as a different player. Since each type t_i chooses just an action $a_i \in A_i$, his strategy is just an action. Note that a strategy profile $s : T \to A$ in the Bayesian game is represented as a list $s = (s_i(t_i))_{t_i\in\hat{N}}$ of actions. Finally, his payoff from a strategy profile $s = (s_{t_i})_{t_i\in\hat{N}}$ is the expected utility of i from s conditional on type t_i. I only consider finite type spaces in the definition in order to avoid measurability issues.

3.2 Bayesian Nash Equilibrium

In this section, I present the main equilibrium concept for Bayesian games: *Bayesian Nash equilibrium*. Defining a Bayesian Nash equilibrium as a Nash equilibrium of the interim game, I establish its existence and continuity properties using the corresponding results for Nash equilibrium.

Recall that a strategy of a player i in a Bayesian game is a mapping $s_i : T_i \to A_i$, as in the ex-ante game. I will focus throughout the section on measurable strategies (when T_i is uncountable).

Definition 3.2 *Given any Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$, a strategy profile $s^* : T \to A$ is said to be a* Bayesian Nash equilibrium of $\mathcal{B}$ *if*

$$E_i\left[u_i\left(\theta, s_i^*\left(t_i\right), s_{-i}^*\left(t_{-i}\right)\right)|t_i\right] \geq E_i\left[u_i\left(\theta, a_i, s_{-i}^*\left(t_{-i}\right)\right)|t_i\right] \qquad (\forall a_i \in A_i, \forall t_i \in T_i, \forall i \in N)$$

where $E_i[\cdot|t_i]$ is the expectation operator under the probability distribution $p_i(\cdot|t_i)$. A mixed strategy Bayesian Nash equilibrium is defined similarly. The set of all Bayesian Nash equilibria of $\mathcal{B}$ is denoted by $BNE(\mathcal{B})$.

That is, s^* is a Bayesian Nash equilibrium if and only if each type plays a best response to other players' strategies. In other words, s^* is a Nash equilibrium of the interim game $AG(\mathcal{B})$, when it is viewed as a collection of actions. Clearly, when player i plays a best-response at each given type, then he also optimizes his ex-ante payoffs, i.e.,

$$U_i(s^*) = E\left[u_i\left(\theta, s_i^*(t_i), s_{-i}^*(t_{-i})\right)\right] \geq E\left[u_i\left(\theta, s_i(t_i), s_{-i}^*(t_{-i})\right)\right] = U_i\left(s_i, s_{-i}^*\right)$$

for every $s_i : T_i \to A_i$. Therefore, a Bayesian Nash equilibrium is also a Nash equilibrium of the ex-ante game $G(\mathcal{B})$. Conversely, ex-ante optimality of a strategy requires that the player maximizes at each type that has positive ex-ante probability. In that case, every Nash equilibrium of the ex-ante game is also a Bayesian Nash equilibrium. These facts are formally stated as follows.

Fact 3.1 *Let $\mathcal{B} = (N, A, \Theta, T, u, p)$ be any Bayesian game. Whenever T is finite,*

$$BNE(\mathcal{B}) = NE(AG(\mathcal{B})).$$

Moreover,

$$BNE(\mathcal{B}) \subseteq NE(G(\mathcal{B})).$$

Conversely, if $p_i(t_i) > 0$ for all $t_i \in T_i$ and $i \in N$, then

$$BNE(\mathcal{B}) = NE(G(\mathcal{B})).$$

Exercise 3.1 *Formally prove these facts.*

I will use these facts to extend the existence and continuity properties of Nash equilibrium to Bayesian Nash equilibrium. Beforehand, I must note that some authors define Bayesian Nash equilibrium as a Nash equilibrium of the ex-ante game. Here, I use the interim definition for two reasons. First, conceptually, I take a Bayesian game to be a model of interim beliefs, and hence rationality of players corresponds to interim optimization; an ex-ante distribution need not be specified. Second, pragmatically, in many economic applications, there are a continuum of types. For example, the value of a object for a buyer in bargaining or auction would typically be modeled as a real number coming from an interval. Hence, in such models, even if one specifies an ex-ante distribution, ex-ante distribution assigns zero probability on some (and often all) types. But any action for such zero-probability types is a best-response, leading to a multitude of spurious equilibria. Such spurious equilibria are ruled out here.

3.2.1 Existence of Bayesian Nash Equilibrium

Building on Fact 3.1 and Theorem 2.4, the following result establishes a general existence result for Bayesian Nash equilibria.

Theorem 3.1 *Let $\mathcal{B} = (N, A, \Theta, T, u, p)$ be a Bayesian game where*

- *each A_i is a convex, compact subset of a Euclidean space,*
- *each $u_i : \Theta \times A \to \mathbb{R}$ is continuous and concave in a_i,*
- *Θ is a separable metric space, and*
- *T is finite.*

Then, there exists a Bayesian Nash equilibrium $s^* : T \to A$ of game $\mathcal{B}$ in pure strategies.

Proof. Since $BNE(\mathcal{B}) = NE(AG(\mathcal{B}))$ by Fact 3.1, it suffices to show that $AG(\mathcal{B})$ has a pure strategy Nash equilibrium, by checking that $AG(\mathcal{B})$ satisfies the conditions of Theorem 2.4. Since T is finite, $AG(\mathcal{B})$ is a finite-player game in which the strategy set A_i of each player t_i is a convex, compact subset of a Euclidean space. Moreover, since u_i is continuous and $\Theta \times A \times T_{-i}$ is a compact metric space,

$$\hat{U}_{t_i}(s) = E\left[u_i(\theta, s(t)) \,|p_i(\cdot|t_i)\right]$$

is continuous by Lemma A.1. Since u_i is concave in a_i, $\hat{U}_{t_i}(s)$ is concave (and hence quasi-concave) in $s_i(t_i)$. Therefore, there exists a Nash equilibrium $s^* : T \to A$ of game $AG(\mathcal{B})$. ■

For games with finite type spaces, compact and convex strategy sets and concave utility functions, Theorem 2.4 proves existence of a *pure* strategy Bayesian Nash equilibrium. The finiteness of type space is made in order to use the equivalence of Nash equilibrium in interim game to Bayesian Nash equilibria. Note that Theorem 3.1 requires that the utility functions are concave in own action, while Theorem 2.4 requires only quasi-concavity. This is because quasi-concavity is not preserved under expectation. One can use this result to establish existence of Bayesian Nash equilibrium in classical economic models with finite type spaces. As in the case of Nash equilibrium, Theorem 2.4 also implies that finite games have Bayesian Nash equilibria:

Corollary 3.1 *Every finite Bayesian game* $\mathcal{B} = (N, A, \Theta, T, u, p)$ *has a (possibly mixed) Bayesian Nash equilibrium* σ^*.

3.2.2 Upperhemicontinuity of Bayesian Nash Equilibrium

I will now use Fact 3.1 and the upperhemicontinuity of Nash equilibrium to obtain two upper-hemicontinuity results for Bayesian Nash equilibrium, one with respect to payoff parameters, and one with respect to the beliefs (about the payoff parameters and types).

Fix a set N of players, a set Θ of payoff parameters, another set X of some known, payoff-relevant parameters, and a set A of action profiles. Assume that the sets Θ, X, and A are all separable metric spaces, and A is compact. The utility functions

$$u_i : \Theta \times A \times X \to \mathbb{R}$$

are continuous and depend on x as well as θ and a. Fix also a finite set T of type profiles. Write $BNE(x)$ for the set of all Bayesian Nash equilibria of game $\mathcal{B}^x = (N, A, \Theta, T, u(\cdot; x), p)$, in which it is common knowledge that the parameter value is x. Towards analyzing the continuity with respect to the beliefs, let also P be a compact set of probability distribution tuples $(p_1, \ldots, p_n)$ on $\Theta \times T$ such that

$$p_i(t_i) \equiv p_i(\Theta \times \{t_i\} \times T_{-i}) > 0$$

for each $t_i \in T_i$ and $i \in N$. When p varies in P, write $BNE(p)$ for the set of all Bayesian Nash equilibria of game $\mathcal{B}^{p,x} = (N, A, \Theta, T, u(\cdot; x), p)$.

Theorem 3.2 *Under the assumptions on* (N, A, Θ, T, u, X) *in the previous paragraph,* $BNE(x)$ *is upperhemicontinuous in* x. *Moreover,* $BNE(x, p)$ *is upperhemicontinuous in* (x, p) *on* P.

Proof. To prove the first statement, observe that, by Fact 3.1,

$$BNE(x) = BNE(\mathcal{B}^x) = NE(AG(\mathcal{B}^x)).$$

Observe also that, in game $AG(\mathcal{B}^x)$, every strategy space A_i is a compact metric space, and every utility function

$$\hat{U}_{t_i}(s, x) = E[u_i(\theta, s(t), x) | p_i(\cdot | t_i)]$$

is continuous by Lemma A.1. Hence, by Theorem 2.5, $NE\left(AG\left(\mathcal{B}^{x}\right)\right)$ is upperhemicontinuous in x.

To prove the second statement, observe that, by Fact 3.1, $BNE\left(p\right) = NE\left(G\left(\mathcal{B}^{P}\right)\right)$ on P. Moreover, in the ex-ante game $G\left(\mathcal{B}^{P}\right)$, each strategy space $S_i = A_i^{T_i}$ is a compact metric space, and each utility function

$$U_i\left(s, x, p\right) = E_p[u_i(\theta, s(t), x)]$$

is continuous by Lemma A.1. Once again, by Theorem 2.5, this implies that $BNE\left(x, p\right) = NE\left(G\left(\mathcal{B}^{x,p}\right)\right)$ is upperhemicontinuous in (x, p). ■

For Bayesian games with finite type spaces, under the usual compactness and continuity assumptions, the first part states that Bayesian Nash equilibrium correspondence is upperhemicontinuous. Here, finiteness of the type space allowed us to simply use the equivalence to the Nash equilibrium of the interim game and the upper-hemicontinuity of Nash equilibrium.

The second part further states that the Bayesian Nash equilibrium correspondence is upperhemicontinuous in ex-ante beliefs as well, provided that the probabilities of all own-types are away from zero. Positiveness of those probabilities allowed the use of equivalence to the Nash equilibrium of the ex-ante game in the proof. This assumption, however, plays a more substantial role than the proof may suggest. Note that the Bayesian Nash equilibrium actions depend on the interim beliefs $p_i\left(\cdot|t_i\right)$. When $p_i\left(t_i\right) > 0$,

$$p_i\left(\cdot|t_i\right) = \frac{p_i\left(\cdot, t_i\right)}{p_i\left(t_i\right)}$$

is a continuous function of p_i, while it is discontinuous at $p_i\left(t_i\right) = 0$. The assumption that $p \in P$ ensures that the beliefs are away from the discontinuities of the Bayes' rule.

This discussion also suggests that the main continuity property here is with respect to the interim beliefs, regardless of the conditions on the ex-ante beliefs, which need not even be specified in the definition of a Bayesian game. The next corollary shows that this is indeed the case:

Corollary 3.2 *Under the assumptions in Theorem 3.2, $BNE\left(x, p\right)$ is upperhemicontinuous in (x, p) where $p = \left(p_i\left(\cdot|t_i\right)\right)_{i\in N, t_i\in T_i}$ is the vector of interim beliefs.*

Proof. Given any list $(p_i(\cdot|t_i))_{i\in N, t_i\in T_i}$ of interim beliefs, one can define a profile $(p_i)_{i\in N}$ of ex-ante beliefs so that $p_i(t_i) = 1/|T_i| \gg 0$ for each t_i by setting

$$p_i(\theta, t) = p_i(\theta, t_{-i}|t_i) / |T_i|$$

at each (θ, t). (In the formula, I assumed that Θ is finite for clarity.) Hence, one can pick $P = \left\{(p_i)_{i\in N} | p_i(t_i) = 1/|T_i| \quad \forall t_i\right\}$. Since BNE does not depend on the specification of the interim beliefs and $p_i(\theta, t)$ is continuous in the interim beliefs above, Theorem 3.2 shows then that BNE is upperhemicontinuous in those beliefs. ■

Note however that economic models often assume a common prior, by assuming $p_1 = p_2 = \cdots = p_n$. In that case, one cannot pick p_i arbitrarily, and the positiveness assumption for types is a substantial assumption. Under the common-prior assumption, Theorem 3.2 establishes that the Bayesian Nash equilibrium is upperhemicontinuous with respect to the prior, so long as the probability of all types are away from zero. That is, the equilibrium is continuous with respect to the beliefs as long as they are foreseen by the ex-ante prior (i.e. the modeler). Nevertheless, it is usually discontinuous with respect to the variations that are not foreseen (i.e. with zero ex-ante probability), and that is the robustness check one must do in order to check robustness with respect to the misspecification error in the modeling stage. This will be clearer in Chapter 11, where robustness is studied formally.

3.3 Rationalizability

In Bayesian games there are many rationalizability notions, each reflecting a different perspective for Bayesian games. Here, I will discuss three main concepts. The first concept is simply the rationalizability with respect to the ex-ante game, and will be called *ex-ante rationalizability*. This corresponds to the view that takes the ex-ante stage as real. The second concept is simply the rationalizability with respect to the interim game, and will be called *interim-independent rationalizability* (henceforth, IIR). Here, independence refers to the assumption one makes when taking marginals with respect to θ in the formulation of the interim game–before specifying the players' beliefs about the other players' actions in the rationalizability analysis. Finally, the third solution concept, called *interim-correlated rationalizability* (henceforth, IIR), drops this assumption.

3.3.1 Ex-ante Rationalizability

Here, I will present the definition of ex-ante rationalizability and illustrate the assumptions made in such an analysis on an example.

Definition 3.3 *Given any Bayesian game* $\mathcal{B} = (N, A, \Theta, T, u, p)$ *and any player* $i \in N$, *a strategy* $s_i : T_i \to A_i$ *is said to be* ex-ante rationalizable *if* s_i *is rationalizable in ex-ante game* $G(\mathcal{B})$.

Ex-ante rationalizability makes sense if there is an ex-ante stage. In that case, ex-ante rationalizability captures precisely the implications of common knowledge of rationality as perceived in the ex-ante planning stage of the game. It does impose unnecessary restrictions on players' beliefs from an interim perspective, however. In order to illustrate the idea of ex-ante rationalizability and its limitations consider the following example.

Example 3.1 *Take*

$$\begin{aligned} \Theta &= \{\theta, \theta'\}; \\ T_1 &= \{t_1, t_1'\}; T_2 = \{t_2\}; \\ p(\theta, t_1, t_2) &= p(\theta', t_1', t_2) = 1/2. \end{aligned}$$

In this type space, there are two parameter values, and Player 1 knows the true value while Player 2 does not, assigning probability 1/2 to each value. And this is common knowledge. Take the payoff functions and the action spaces as in the following table

θ	L	R
U	$1, \varepsilon$	$-2, 0$
D	$0, 0$	$0, 1$

θ'	L	R
U	$-2, \varepsilon$	$1, 0$
D	$0, 0$	$0, 1$

where $u_1(\theta, U, L) = 1$ *and* $u_1(\theta', U, L) = -2$ *for example. To compute the ex-ante rationalizability, one first needs to find the ex-ante game* $G(\mathcal{B})$*:*

	L	R
UU	$-1/2, \varepsilon$	$-1/2, 0$
UD	$1/2, \varepsilon/2$	$-1, 1/2$
DU	$-1, \varepsilon/2$	$1/2, 1/2$
DD	$0, 0$	$0, 1$

Here, for the strategies of player 1, the first entry is the action for t_1 *and the second entry is the action for* t_1'*, e.g.,* $UD(t_1) = U$ *and* $UD(t_1') = D$*. The payoffs are computed by taking ex-ante expectation. For example, payoff of player 1 from* (UD, L) *is computed by* $\frac{1}{2}u_1(\theta, U, L) + \frac{1}{2}u_1(\theta', D, L)$*. This game has a unique rationalizable strategy profile*

$$S^{\infty}(G(\mathcal{B})) = \{(DU, R)\}.$$

In computing the rationalizable strategies, one first eliminates UU, noting that it is dominated by DD, and then eliminates L and finally eliminates UD and DD. Note, however, that elimination of UU crucially relies on the assumption that player 1's belief about the other player's action is independent of player 1's type. Otherwise, we could not eliminate DD. For example, if type t_1 believed player 2 plays L, he could play U as a best response, and if type t_1' believed player 2 plays R, he could play U as a best response. The assumption that the beliefs of types t_1 and t_1' are identical is embedded in the definition of ex-ante game. Moreover, the conclusion that (DU, R) is the only rationalizable strategy profile crucially relies on the assumption that player 1 knows that player 2 knows that player 1's belief about player 2's action is independent of player 1's belief about the state.

From an interim perspective, such invariance assumption for beliefs (and common knowledge of it) is unwarranted. This is because distinct types of a player correspond to distinct hypothetical situations that are used in order to encode players' beliefs. There is no reason to assume that in those hypothetical situations a player's belief about the other player's action is independent of his beliefs about the payoffs. (Of course, if it were actually the case that player 1 observes a signal about the state without observing a signal about the other player's action and player 2 does not observe anything, then it would have been plausible to assume that player 1's belief about player 2's action does depend on his signal. This is what ex-ante rationalizability captures. This is not the story however in a genuine incomplete information.)

Exercise 3.2 *Consider any Bayesian game* $\mathcal{B} = (N, A, \Theta, T, u, p)$ *in which each type has positive probability. Show that if a strategy* $s_i : T_i \to A_i$ *is played with positive probability by a player* i *in a Bayesian Nash equilibrium, then* s_i *is ex-ante rationalizable.*

3.3.2 Interim Independent Rationalizability

In order to capture the implication of common knowledge of rationality from an interim perspective without imposing any restriction on the beliefs of distinct types, one then needs to eliminate actions for type in the interim stage. While most contemporary game theorists would agree on the relevant notion of ex-ante rationalizability and the relevant notion of rationalizability in complete-information games, there is a disagreement about the relevant notion of interim rationalizability in incomplete information games.

One straightforward notion of interim rationalizability is to apply rationalizability to interim game $AG(\mathcal{B})$. An embedded assumption on the interim game is however that it is common knowledge that the belief of a player i about (θ, t_{-i}), which is given by $p_i(\cdot|t_i)$, is independent of his belief about the other players actions. That is, his belief about (θ, t_{-i}, a_i) is derived from some belief $p_i(\cdot|t_i) \times \mu_{t_i}$ for some $\mu_{t_i} \in \Delta\left(A_{-i}^{T_{-i}}\right)$. This is because we have taken the expectations with respect to $p_i(\cdot|t_i)$ in defining $AG(\mathcal{B})$, before considering his beliefs about the other players' actions. Because of this independence assumption, such rationalizability notion is called *interim independent rationalizability* (IIR).

Definition 3.4 *Given any Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ and any type t_i of player $i \in N$, an action $a_i \in A_i$ is said to be* interim independent rationalizable *(IIR) for t_i if a_i is rationalizable for t_i in interim game $AG(\mathcal{B})$.*

As an illustration, I will next apply interim independent rationalizability to the previous Bayesian game.

Example 3.2 *Consider the Bayesian game in the previous example. The interim game $AG(\mathcal{B})$ is a 3-player game with player set $\hat{N} = \{t_1, t_1', t_2\}$ with the following payoff table, where t_1 chooses rows, t_2 chooses columns, and t_1' chooses the matrices:*

U :

	L	R
U	$1, \varepsilon, -2$	$-2, 0, 1$
D	$0, \varepsilon/2, -2$	$0, 1/2, 1$

D :

	L	R
U	$1, \varepsilon/2, 0$	$-2, 1/2, 0$
D	$0, 0, 0$	$0, 1, 0$

(The first entry is the payoff of t_1, the second entry is the payoff of t_2, and the last entry is the payoff of t_1'.) In $AG(\mathcal{B})$, no strategy eliminated, and all actions are rationalizable for all types, i.e., $S^{\infty}_{t_1}(AG(\mathcal{B})) = S^{\infty}_{t_1'}(AG(\mathcal{B})) = \{U, D\}$ and $S^{\infty}_{t_2}(AG(\mathcal{B})) = \{L, R\}$. For example, for type t_1, who is a player in $AG(\mathcal{B}), U$ is a best response to t_2 playing L (regardless of what t_1' would have played). For t_i', U is a best response to t_2 playing R. For t_2, L is a best response to (U, U) and R is a best response to (D, D).

Exercise 3.3 *For any Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ with finite type space, show that if an action a_i is played with positive probability by a type t_i in a Bayesian Nash equilibrium, then a_i is interim independent rationalizable for t_i.*

3.3.3 Interim Correlated Rationalizability

As discussed in Section 2.1 (see Remark 1.1), the fact that two players choose their actions independently does not mean that a third player's belief about their actions will have a product form. In particular, just because all of player j's information about θ, which is the action of the nature, is summarized by t_j does not mean the belief of i about the state θ and the action of j does not have any correlation once one conditions on t_j. Once again i might find it possible that the factors that affect the payoffs may also affect how other players will behave given their beliefs (regarding the payoffs). This leads to the following notion of rationalizability, called *interim correlated rationalizability.*

Iterated Elimination of Strictly Dominated Actions Consider a Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$. For each $i \in N$ and $t_i \in T_i$, set $S_i^0[t_i] = A_i$, and define sets $S_i^k[t_i]$ for $k > 0$ iteratively, by letting $a_i \in S_i^k[t_i]$ if and only if

$$a_i \in B_i(\pi) \equiv \arg\max_{a_i'} \int u_i(\theta, a_i', a_{-i}) d\pi(\theta, t_{-i}, a_{-i})$$

for some $\pi \in \Delta(\Theta \times T_{-i} \times A_{-i})$ such that

$$\text{marg}_{\Theta \times T_{-i}} \pi = p_i(\cdot|t_i) \text{ and } \pi\left(a_{-i} \in S_{-i}^{k-1}[t_{-i}]\right) = 1.$$

That is, a_i is a best response to a belief of t_i that puts positive probability only on the actions that survive the elimination in round $k-1$. Write $S_{-i}^{k-1}[t_{-i}] = \prod_{j \neq i} S_j^{k-1}[t_j]$ and $S^k[t] = \prod_{i \in N} S_i^k[t_i]$.

Definition 3.5 *The set of all* interim correlated rationalizable *(ICR) actions for player i with type t_i is*

$$S_i^\infty[t_i] = \bigcap_{k=0}^{\infty} S_i^k[t_i].$$

Since interim correlated rationalizability allows more beliefs, interim correlated rationalizability is a weaker concept than interim independent rationalizability, i.e., if an action is interim independent rationalizable for a type, then it is also interim correlated rationalizable for that type. When all types have positive probability, ex-ante rationalizability is stronger than both of these concepts because it imposes not only independence but also the assumption that a player's conjecture about the other actions is independent of his type. Since all of the equilibrium concepts are refinements of ex-ante rationalizability, interim correlated rationalizability emerges as the weakest solution concept we have seen so far, i.e., all of them are refinements of interim correlated rationalizability.

Exercise 3.4 *Consider a Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ in which each type t_i has positive ex-ante probability of $p_i(t_i)$.*

1. *Show that if a player i plays a strategy s_i with positive probability in a Bayesian Nash equilibrium, then s_i is ex-ante rationalizable.*

2. *For any ex-ante rationalizable strategy $s_i : T_i \to A_i$ and for any t_i show that $s_i(t_i)$ is interim independent rationalizable for t_i.*

Exercise 3.5 *Show that if a_i is interim independent rationalizable for some type t_i, then a_i is interim correlated rationalizable for t_i.*

Fixed-Point Definition of ICR I will next present a fixed point definition of ICR. A solution concept $\Sigma : T \rightrightarrows A$ is said to have the *best-response property* (or closed under rational behavior) if for every $t_i \in T_i$ and $a_i \in \Sigma(t_i)$, there exists $\mu^{a_i,t_i} \in \Delta(\Theta \times T_{-i} \times A_{-i})$ such that

$$a_i \in BR_i\left(\mu^{a_i,t_i}\right), \tag{3.3}$$

$$p_i(\cdot|t_i) = \operatorname{marg}_{\Theta \times T_{-i}} \mu^{a_i,t_i}, \tag{3.4}$$

$$\mu^{a_i,t_i}(a_{-i} \in \Sigma_{-i}(t_{-i})) = 1. \tag{3.5}$$

As in the case of complete information games, the next result establishes that ICR is the largest solution concept with best-response property.

Theorem 3.3 *If $\Sigma : T \rightrightarrows A$ has best-response property, then $\Sigma \subseteq S^\infty$. Moreover, under Assumption 3.1, S^∞ has the best response property.*

Exercise 3.6 *Prove this result. (See the proofs of Theorem 2.3 and Theorem 3.4 below.)*

Upperhemicontinuity of ICR I will now present a general upperhemicontinuity theorem for interim correlated rationalizability and illustrate a couple of its applications.

Theorem 3.4 *Let $\mathcal{B} = (N, A, \Theta, T, u, p)$ be a Bayesian game where*

- *A, Θ, and T are compact metric spaces, and*
- *$u_i : \Theta \times A \to R$ and $t_i \mapsto p_i(\cdot|t_i)$ are continuous for every $i \in N$.*

Then, $S^\infty : T \rightrightarrows A$ is upperhemicontinuous and compact valued.

Proof. For each finite $m \in \mathbb{N}$, I will show that S^m has the closed-graph property, i.e., the graph

$$G(S^m) = \{(\theta, t, a) \,|a \in S^m[t]\}$$

of S^m is closed. This further implies that

$$G(S^\infty) = \bigcap_{m \geq 0} G(S^m)$$

is also closed. Since A is compact, by Proposition A.1, this is indeed the desired result: S^∞ is upperhemicontinuous and compact-valued.

Clearly, $G(S^0) = T \times A$ is closed. Towards an induction, assume that $G\left(S^{m-1}_{-i}\right)$ is closed for some $i \in N$ and $m \in \mathbb{N}$. Take a sequence $(t_{i,k}, a_{i,k}) \in G(S_i^m)$ with limit (t_i, a_i). For each k, since $(t_{i,k}, a_{i,k}) \in G(S_i^m)$, there exists $\pi_k \in \Delta(\Theta \times T_{-i} \times A_{-i})$ such that

$$a_{i,k} \in B_i(\pi_k), \tag{3.6}$$

$$\kappa_{t_{i,k}} = \text{marg}_{\Theta \times T_{-i}} \pi_k, \tag{3.7}$$

$$\mu_k\left(G\left(S^{m-1}_{-i}\right)\right) = 1. \tag{3.8}$$

Note that, since $\Theta\times T_{-i}\times A_{-i}$ is a compact metric space, so is $\Delta(\Theta\times T_{-i}\times A_{-i})$. Hence, μ_k has a convergent subsequence with a limit $\pi\in\Delta(\Theta\times T_{-i}\times A_{-i})$. Switching to the convergent subsequence, I will show that μ satisfies the conditions in the definition of S_i^m, showing that $a_i\in S_i^m(t_i)$, as desired.

Firstly,

$$p(\cdot|t_{i,k})=\text{marg}_{\Theta\times T_{-i}}\pi_k\to\ \text{marg}_{\Theta\times T_{-i}}\pi,$$

where the equality is by (3.7), and the convergence is by continuity of the projection mapping. On the other hand, since $t_{i,k}\to t_i$ and the mapping is continuous, $p(\cdot|t_{i,k})\to p(\cdot|t_i)$. Hence, the above convergence implies that

$$\text{marg}_{\Theta\times T_{-i}}\pi=p(\cdot|t_i).$$

Secondly, since B_i is upperhemicontinuous by Lemma A.1, (3.6) implies that $a_i\in B_i(\pi)$. Finally, since $G\left(S_{-i}^{m-1}\right)$ is closed (by the inductive hypothesis) and $\pi_k\to\pi$, by Portmanteau Theorem,

$$\pi\left(a_{-i}\in S_{-i}^{m-1}[t_{-i}]\right)\equiv\pi\left(G\left(S_{-i}^{m-1}\right)\right)\geq\limsup\pi_k\left(G\left(S_{-i}^{m-1}\right)\right)=1,$$

where the last equality is by (3.8). ■

Under the standard continuity and compactness assumptions, Theorem 3.4 establishes upperhemicontinuity of ICR with respect to the types. The compactness of A and the continuity of payoffs and beliefs are not superfluous because the Maximum Theorem, the special case of Theorem 3.4 for the single-player case, could fail when either of these assumptions fail.

Theorem 3.4 considers upperhemicontinuity with respect to the types only, while earlier upperhemicontinuity results considered upperhemicontinuity with respect to the known payoff relevant parameters. Indeed, Theorem 3.4 implies upperhemicontinuity with respect to such parameters as well. To see this, recall that the above formulation in which $u_i:\Theta\times A\to\mathbb{R}$ is as general as the formulation in which $u_i:\Theta\times T\times A\to\mathbb{R}$ (by considering a new space $\Theta\times T$ of payoff parameters). Now suppose that there is an additional known payoff-parameter x from a compact metric space X so that

$$u_i:\Theta\times X\times T\times A\to\mathbb{R}$$

is continuous. One can now define a new type space $\tilde{T}$ with types

$$\tilde{t}_i=(t_i,x)$$

and beliefs

$$p\left(\left(\theta,(t_j,x)_{j\neq i}\right)|(t_i,x)\right)=p\left(\left(\theta,(t_j)_{j\neq i}\right)|t_i\right)\qquad(\forall(\theta,t,x)).$$

One can then apply Theorem 3.4 to type space $\tilde{T}$ to obtain upperhemicontinuity with respect to payoff parameters, obtaining the following corollary.

Corollary 3.3 *Let X be compact metric space and consider the family of Bayesian games $\mathcal{B}^x=(N,A,\Theta,T,u(\cdot,x),p)$ where*

- *A, Θ, and T are compact metric spaces,*
- *$u_i:\Theta\times T\times A\times X\to\mathbb{R}$ and $t_i\mapsto p_i(\cdot|t_i)$ are continuous for every $i\in N$, and*
- *x is commonly known.*

Then, S^∞ is upperhemicontinuous in (x,t).

One can also consider upperhemicontinuity with respect to belief parameters. Here, I will confine myself to continuity with respect to prior beliefs as in Theorem 3.2.

Corollary 3.4 *Consider the family of Bayesian games $\mathcal{B}^p=(N,A,\Theta,T,u(\cdot,x),p)$, $p\in P$, where*

- *A, Θ, and T are finite, and*
- *P is a compact set of probability distribution tuples $p=(p_1,\ldots,p_n)$ on $\Theta\times T$ such that*

$$p_i(t_i)\equiv p_i(\Theta\times\{t_i\}\times T_{-i})>0$$

Then, S^∞ is upperhemicontinuous in p.

Exercise 3.7 *Formally prove the above corollaries from Theorem 3.4.*

One can present three justifications for using interim correlated rationalizability in genuine cases of incomplete information, which I described in the previous section. First, interim correlated rationalizability captures the implications of common knowledge of rationality precisely, as Theorem xx will establish formally. Second, interim independent rationalizability depends on the way the hierarchies are modeled, in that there can be multiple representations of the same hierarchy with distinct set of interim independent rationalizable actions. Finally, and most importantly, at it will be clear in Chapter 11, one cannot have any extra robust prediction from refining interim correlated rationalizability. Any prediction that does not follow from interim correlated rationalizability alone crucially relies on the assumptions about the infinite hierarchy of beliefs. A researcher cannot verify such a prediction in the modeling stage without the knowledge of infinite hierarchy of beliefs.

3.3.4 Invariance to Representation of Belief-Hierarchies

Now, by way of an example, I will show that Bayesian Nash equilibrium and IIR are not invariant to the way the hierarchies are modeled. That is, there can be two types in two different Bayesian games with the same hierarchy, and yet the sets of IIR actions are distinct for those types (and the sets of Bayesian Nash equilibrium actions are distinct). Among other things this shows that these concepts cannot be upperhemicontinuous with respect to the belief hierarchies.

Example 3.3 *Fix* $\Theta = \{\theta, \theta'\}$ *and* $N = \{1, 2\}$. *Fix also the action spaces and the payoff functions as in the following table:*

θ	L	R
U	$1,0$	$0,0$
D	$0.6,0$	$0.6,0$

θ'	L	R
U	$0,0$	$1,0$
D	$0.6,0$	$0.6,0$

First consider the type space $T = \{t_1, t_1'\} \times \{t_2, t_2'\}$ *with the following common prior*

θ	t_2	t_2'
t_1	$1/6$	$1/12$
t_1'	$1/12$	$1/6$

θ'	t_2	t_2'
t_1	$1/12$	$1/6$
t_1'	$1/6$	$1/12$

Here, the probability of (θ, t_1, t_2) *is* $1/6$ *while the probability of* (θ, t_1, t_2') *is* $1/12$. *Write* $\mathcal{B}$ *for the resulting Bayesian game. In* $\mathcal{B}$, *every action can be played in a Bayesian Nash equilibrium and hence is interim independent rationalizable.*

To see this, consider the following Bayesian Nash equilibrium s^* first:

$$\begin{aligned} s_1^*(t_1) &= U; s_1^*(t_1') = D; \\ s_2^*(t_2) &= L; s_2^*(t_2') = R. \end{aligned}$$

In order to check that this is a BNE, it suffices to check that player 1 plays a best response, as player 2 is always indifferent. To this end, observe that type t_1 puts probability 1/3 on each of (θ, t_2) and (θ', t_2'). Hence, his expected payoff from U is 2/3, which is higher than the expected payoff from D, which is only 0.6. Likewise, the expected payoff of t_1' from U is 1/3, enticing that type to play D and get 0.6. One can, of course, switch the actions of types to get another equilibrium s^{**}:

$$\begin{aligned} s_1^{**}(t_1) &= D; s_1^{**}(t_1') = U; \\ s_2^{**}(t_2) &= R; s_2^{**}(t_2') = L. \end{aligned}$$

Therefore, each action is played by each type in a Bayesian Nash equilibrium, showing that they are all interim independent rationalizable. Note that there is yet another Bayesian Nash equilibrium in which Player 1 plays D and player 2 plays the same action at both types.

Example 3.4 *Now consider a new type space $\hat{T} = \{(\hat{t}_1, \hat{t}_2)\}$ with $p(\theta, \hat{t}) = p(\theta', \hat{t}) = 1/2$. The interim game with the new type space is given by*

	L	R
U	$1/2, 0$	$1/2, 0$
D	$.6, 0$	$.6, 0$

The only rationalizable action for player 1 in this game is D. Therefore, D is the only interim independent rationalizable action for type $\hat{t}_1$. Hence, in any Bayesian Nash equilibrium, $\hat{t}_1$ must play D with probability 1.

The solutions to the above games are quite different under both the Bayesian Nash equilibrium and the interim independent rationalizability. Everything is a solution in the first game, while player 1 must play D in every solution of the second game.

Nevertheless, the two games represent the same hierarchy of beliefs! Indeed, each type $t_i \in T_i$ in the first game assigns probability 1/2 on θ. Therefore, it is common knowledge that each player assigns probability 1/2 to each value, precisely as it is modeled in the complete-information game with trivial type space $\hat{T}$.

This example demonstrates that Bayesian Nash equilibrium and interim independent rationalizability use some information embedded in the type space that is irrelevant to the hierarchies of beliefs. If one views the types spaces in Bayesian games merely as devices to represent the belief hierarchies, this can be viewed as an important critique of these solution concepts, as they refer to irrelevant aspects of the type space. Alternatively, one may take one of the above solution concepts as the relevant model of strategic behavior and view this result as a critique of the idea of belief hierarchies: there may be relevant information in the strategic environment that is not captured by the belief hierarchies.

No matter how one views this result, it shows an important fact regarding continuity of the above solution concepts with respect to the belief hierarchies: Bayesian Nash equilibrium and IIR are not upperhemicontinuous with respect to belief hierarchies, regardless of the topology on these hierarchies. I will next show that the latter fact is an instance of a general theorem: ICR does not have an non-empty refinement that is upperhemicontinuous with respect to belief hierarchies.

3.4 Correlated Equilibrium

Once again there are multiple formulations of correlated equilibrium for Bayesian games. I will present two definitions. First definition is straightforward: one simply extends epistemic model in the definition of correlated equilibrium by identifying players' types at each state and requiring that the the types' beliefs are preserved. Such a model does not allow players to have further information about the underlying fundamentals and the other players' beliefs. Another definition allows players to have more information in the epistemic model, leading to a weaker solution concept. For simplicity, I will fix a finite Bayesian game $\mathcal{B} = (N, A, \Theta, u, T, p)$ with a common prior $p \in \Delta(\Theta \times T)$ throughout this section. I start with the first definition.

Definition 3.6 *A* correlated equilibrium *is an epistemic model* $((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ *for* $\mathcal{B}$ *such that*

1. (Ω, I, π) *admits a common prior* P *on* Ω *(i.e.,* $\pi_{i,\omega} = P(\cdot|I_i(\omega))$ *for all* i *and* ω*) and*

2. *it is common knowledge that every player* i *is rational:*

$$\boldsymbol{a}_i(\omega) \in B_i\left(\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{a}_{-i})^{-1}\right) \qquad \forall i \in N, \omega \in \Omega.$$

This definition simply imposes the common prior assumption on top of common knowledge of rationality. Hence, it is a refinement of interim correlated rationalizability in that if a type t_i plays an action a_i at some ω in a correlated equilibrium, then a_i is interim correlated rationalizable for t_i, i.e., $a_i \in S_i^\infty[t_i]$. As in the complete information case, the likelihood of actions profiles for types can be computed using the prior P.

Definition 3.7 *A* correlated equilibrium distribution *is a probability distribution*

$$q = P \circ (\boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})^{-1}$$

on the set $\Theta \times T \times A$ *induced by a correlated equilibrium* $((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$*.*

Note that we are considering a joint distribution on $\Theta \times T \times A$, relating the action to types and the underlying fundamentals. One can ignore the dependence on θ by considering the distribution $P \circ (\boldsymbol{t}, \boldsymbol{a})^{-1}$, which is marginal distribution of q on $T \times A$, but it is somewhat necessary to consider the joint distribution of types and action profiles to determines which types play which actions and how they are correlated to each other.

Exercise 3.8 *Show that for any correlated equilibrium distribution* q*,* $marg_{\Theta\times T} q = p$*. That is, the type distribution is preserved.*

Exercise 3.9 *Extend the definition of mixture of correlated equilibria to Bayesian games and show that a mixture of any two correlated equilibria is also a correlated equilibrium. Conclude that the set of correlated equilibrium distributions is convex.*

The next result extends the characterization of correlated equilibrium distributions in Theorem 2.7 to Bayesian games.

Proposition 3.1 *For any correlated equilibrium distribution* q *of a finite game* $\mathcal{B} = (N, A, \Theta, u, T, p)$*, we have (i)* $marg_{\Theta\times T} q = p$*, and (ii)*

$$\sum_{(\theta, t_{-i}, a_{-i}) \in \Theta \times T_{-i} \times A_{-i}} (u_i(\theta, a_i, a_{-i}) - u_i(\theta, a_i', a_{-i}))\, q(\theta, a, t) \geq 0 \qquad (\forall i, t_i, a_i, a_i'). \quad (3.9)$$

The first part is a coherence property stating that the distribution on type space does not change just because we added a new dimension about how those types play a game. The second part is the usual obedience condition. It states that asked to take an action a_i, type t_i does not have an incentive to deviate—under the possibly additional information provided by the recommendation (i.e. conditioning on (t_i, a_i)). In a complete information game, the coherence condition is vacuous, and the obedience condition is also sufficient for a correlated equilibrium as we have seen in Corollary 2.2. It is tempting to conjecture that these conditions are also sufficient in incomplete information games. This is far from true as the next example illustrates.

Example 3.5 *Consider a one person game in which the player has two actions $\{L, R\}$ where the payoff from L is $\theta \in \{-1, 2\}$ and the payoff from right is 0. The player assigns equal probabilities to $\theta = -1$ and $\theta = 2$; his type is denoted by t. In this decision problem, clearly, there is a unique rationalizable solution: L. Then, the unique correlated equilibrium distribution q is given by $q(-1, t, L) = q(2, t, L) = 1/2$. However, there is another distribution that satisfies the coherence and obedience conditions: $q^*(-1, t, R) = q^*(2, t, L) = 1/2$. According to this solution, there is an omniscient moderator who recommends player to play L when $\theta = 2$ and play R when $\theta = -1$. He is clearly happy to follow the recommendation.*

That is, under uncertainty, correlated equilibrium imposes also that the players do not have any further information about the fundamentals and other players' beliefs about fundamentals—in addition to obedience and coherence conditions. Formally, this is the coherence requirement for information structures as stated in condition (10.9).

A second notion, called Bayes correlated equilibrium, simply ignores these informational constraints:

Definition 3.8 *For any finite game $\mathcal{B} = (N, A, \Theta, u, T, p)$, a* Bayes correlated equilibrium distribution *is a distribution q on $\Theta \times T \times A$ that satisfies the coherence condition $marg_{\Theta \times T} q = p$ and the obedience condition (3.9).*

Note that any distribution q on $\Theta \times T \times A$ yields a conditional probability system $q(\cdot|\cdot)$ conditional on (θ, t). Such a system can be represented as a function

$$q(\cdot|\cdot) : \Theta \times T \to \Delta(A),$$

which is called a *decision rule*. A decision rule $q(\cdot|\cdot)$ and a prior p on $\Theta \times T$ induces a probability distribution q on $\Theta \times T \times A$ by

$$q(\theta, t, a) = q(a|t, a)\, p(t, a)\,. \tag{3.10}$$

Thus defined, q satisfies the coherence condition $\text{marg}_{\Theta\times T} q = p$. Hence, obedience condition is sufficient for Bayes correlated equilibrium. Indeed, this is how Bayes correlated equilibrium is defined.

Definition 3.9 *For any finite game $\mathcal{B} = (N, A, \Theta, u, T, p)$, a decision rule $q(\cdot|\cdot)$ is* Bayes correlated equilibrium *if the distribution q induced by $q(\cdot|\cdot)$ and p satisfies the obedience condition (3.9).*

Bayes correlated equilibrium is defined by dropping the condition that the players' beliefs about the fundamentals and the other players' beliefs remain intact. Consequently, the set of Bayes correlated equilibrium distributions is the union of all correlated equilibrium distributions for games in which players can obtain information (in an arbitrary fashion) in addition to what they know in $\mathcal{B}$. Thus, it characterizes the strategic implications of common prior assumption and common knowledge assumption when players can obtain more information.

Note that Bayesian Nash equilibrium is stronger than correlated equilibrium, and the correlated equilibrium is stronger than interim correlated rationalizability and Bayes correlated equilibrium. But as Example 3.5 demonstrates, Bayes correlated equilibrium can be even weaker than interim correlated rationalizability. This is formally stated in the following exercise.

Exercise 3.10 *Consider a finite game $\mathcal{B} = (N, A, \Theta, u, T, p)$.*

1. *For any Bayesian Nash equilibrium $\sigma^* : T \to \Delta(A)$, show that the associated epistemic model $((\Omega^*, I^*, \pi^*), \boldsymbol{\theta}^*, \boldsymbol{t}^*, \boldsymbol{a}^*)$ is a correlated equilibrium where*

$$\begin{aligned} \Omega^* &= \{(\theta, t, a) \,|\sigma^*(a|t) > 0\} \\ I_i^*(\theta, t, a) &= \{(\theta', t', a') \,|t_i' = t_i, a_i' = a_i\} \\ \pi^*(\theta, t, a) &= p(\theta, t)\, \sigma^*(a|t) \end{aligned}$$

and $\boldsymbol{\theta}^*$, $\boldsymbol{t}^*$, and $\boldsymbol{a}^*$ are projection mappings.

2. *Show that for any correlated equilibrium* $((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$,

$$\boldsymbol{a}(\omega) \in S^{\infty}[\boldsymbol{t}(\omega)] \qquad \forall \omega \in \Omega,$$

and conclude that for any correlated equilibrium distribution q,

$$q(S^{\infty}[t] | \theta, t) = 1 \qquad \forall (\theta, t).$$

3. *Find an example in which a correlated equilibrium puts positive probability on some* (θ, t, a) *where* a *is not interim independent rationalizable at* t. *(Hint: consider the games studied in Section 3.3.4.)*

3.5 Bayesian Games in Extensive Form

Bayesian games are included in extensive form games and have some extra structure. Indeed, any Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ can be considered as an extensive form game in which the nature moves at the initial node (i.e. $\iota(\phi) = 0$) selecting each (θ, t) with probability $p(\theta, t)$, and then each player i observes only his own type. That is, there are $|T|$ information sets following ϕ, each player i moving at $|T_i|$ information sets. At each information set $h \in H_i$, the set of available actions is $A(h) = A_i$. A terminal history is of the form $z = (\theta, t, a)$, and the payoff of each player i at any terminal node is $u_i(\theta, a)$.

Extensive-form games subsume many more interesting incomplete-information game with richer information structure, as the nature can move at many different points and $A(h)$ can vary arbitrarily. For example one can allow different types to have different set of actions in a Bayesian game by allowing $A(t_i)$ to vary with t_i in the extensive form game above.

The following class of dynamic Bayesian games consists of important incomplete information games. In these games, the players' types are persistent, in that their information does not change in the course of the game, i.e., they do not get any new information about the underlying payoff parameters or other players' types as they play the game, although they could infer valuable information from other players' actions and update their beliefs as they play the game, of course. Moreover, the players' set of allowable action plans does not depend on their information.

Definition 3.10 *A game form is any list* $\Gamma = (N, (X, <), H, \iota)$ *with no chance moves. A* dynamic Bayesian game *is a any list* $\mathcal{B} = (\Gamma, \Theta, \mathcal{T}, \pi, u)$ *where* Γ *is a game form,* $(\Theta, \mathcal{T}, \pi)$ *is a type space, and* $u_i : \Theta \times Z \to \mathbb{R}$ *is the von-Neumann utility function of player* i*.*

Here, a game form is an extensive-form game without payoffs specified. I also rule out nature's moves for clarity. One could have an extensive form game if one also include a payoff function. There is payoff uncertainty, as u_i depends on parameter θ as well as the terminal history. The incomplete information about the payoff uncertainty is modeled by type space $(\Theta, \mathcal{T}, \pi)$. Recall that $\mathcal{T} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_n$ is the set of type profiles, $\pi_i(\cdot|\tau_i)$ is the belief of type τ_i on $\Theta \times \mathcal{T}_{-i}$, and Z is the terminal histories in game form Γ. I write τ for types $\mathcal{T}$ for the set of type profiles. (The slight change of the notation is due to the fact that, in dynamic games, $t \in T$ often denotes the time.) As in the Bayesian games in normal form, each player observes his own type and plays according to the game form Γ and gets his payoff according to u_i at the end. Although he uses his information in deciding which action to take at each history he moves, the set of allowable actions $A(h)$ at any given history does not depend on his type.

A dynamic Bayesian game $\mathcal{DB} = (\Gamma, \Theta, \mathcal{T}, \pi, u)$ with persistent types can be represented as a normal-form Bayesian game $\mathcal{B} = (N, \mathcal{A}, \Theta, \mathcal{T}, \pi, u)$ by taking the set of actions $\mathcal{A}$ as

$$\mathcal{A}_i = \prod_{h \in H_i} A(h).$$

To avoid confusion, I will refer $\boldsymbol{a}_i \in \mathcal{A}_i$ as an action plan (or a plan of action). Note that a strategy of a player i in a dynamic Bayesian game $\mathcal{DB} = (\Gamma, \Theta, \mathcal{T}, \pi, u)$ is any mapping

$$s_i : \mathcal{T}_i \to \mathcal{A}_i;$$

mixed strategies are defined similarly.

3.6 Notes on Literature

Harsanyi (1967) has formulated Bayesian games as a way to incorporate incomplete information as complete information games with chance moves. The latter games have been studies since Nash (1950), who studied the equilibria of a three-person poker game.

Harsanyi informally discussed the hierarchies of beliefs and showed how to formulate them using type spaces. Mertens and Zamir (1985) and later Brandenburger and Dekel (1993) studied the hierarchies of model and show that Harsanyi's approach is without loss of generality. These formalizations will be discussed in greater detail in Chapter 11.

The formulation of interim-correlated rationalizability is due to Dekel, Fudenberg, and Morris (2007), who showed that ICR is invariant with respect to the representation of belief hierarchies. This paper also contains a characterization of common knowledge of rationalizability in terms of ICR, extending the characterization in the complete information games to Bayesian games. Battigalli (1999) has an extensive discussion of rationalizability concepts in incomplete information games. Battigalli and Siniscalchi (2003) introduces a concept called Δ-rationalizabilty, which captures the strategic implications of rationalizability and the restrictions Δ, where Δ can vary. Different set Δ of assumptions corresponds to a different solution concept. Non-invariance of Bayesian Nash equilibrium and IIR has been discussed in Ely and Peski (2006). A general analysis of invariance of solution concepts to representation of information can be found in Yildiz (2015).

Bayes correlated equilibrium is introduced by Bergeman and Morris (2016); see also Forges (1993) and Liu (2015) for other notions of correlated equilibrium for Bayesian games.

3.7 Exercises

Exercise 3.11 *Consider the folowing Bayesian game in which each payoff matrix is equally likely, and privately known by player 2*

	a	b
a	$1,1$	$0,0$
b	$0,0$	$1,2$

	a	b
a	$1,0$	$0,2$
b	$0,1$	$1,0$

	a	b
a	$1,1$	$1,0$
b	$0,1$	$0,0$

1. *Compute a Bayesian Nash equilibrium.*

2. *Compute a correlated equilibrium that is not a mixture of Bayesian Nash equilibria.*

3. *Compute the sets of ex-ante, interim independent, and interim correlated rationalizable solutions.*

Exercise 3.12 *Consider the two-player Bayesian game B with $\Theta = T_1 = \{-2, 2\}$, $T_2 = \{t_2\}$, $A_1 = A_2 = \{a, b\}$, $\Pr(\theta = t_1 = -2) = \Pr(\theta = t_1 = 2) = 1/2$, and payoffs*

	a	b
a	θ, θ	$0, 0$
b	$1, 0$	$\theta, 1$

1. *Write the (ex-ante) normal-form game G for B.*

2. *Compute the set of correlated equilibrium distributions for G.*

3. *For each correlated equilibrium distribution q for G, compute the distribution $\hat{q}$ induced by q on $\Theta \times A_1 \times A_2$. Construct a correlated equilibrium for B that induces distribution $\hat{q}$ on $\Theta \times A_1 \times A_2$.*

4. *Construct a correlated equilibrium for B such that the distribution $\tilde{q}$ induced by the correlated equilibrium on $\Theta \times A_1 \times A_2$ cannot be induced by any correlated equilibrium for G.*[2]

Exercise 3.13 *Consider a Cournot oligopoly with n firms in which the inverse-demand function is given by*

$$P = \theta - Q$$

where

$$\theta = 1 + t_1 + \cdots + t_n;$$

each t_i is privately known by t_i and $(t_1, \ldots, t_n)$ are i.i.d. with zero mean. (Each firm i is to produce $q_i \geq 0$; make any reasonable assumption about the distribution that you find helpful.)

1. *Compute the Bayesian Nash equilibria.*

2. *Compute the ex-ante, interim independent, and interim correlated rationalizable solutions for $n = 2$ and $n = 4$.*

Exercise 3.14 *Consider the complete unformation game $G = (N, A, u)$ in Exercise 2.10.*

[2] You can work with the distributions on $\Theta \times A_1 \times A_2$. Find a distribution $\tilde{q}$ that satisfies the conditions emposed by correlated equilibria of B but does not satisfy the conditions for $\hat{q}$.

1. *Construct a finite Bayesian game $B = (N, A, T, p, \tilde{u})$ with a Bayesian Nash equilibrium $s^* : T \to A$ such that*

   (a) $\tilde{u}_i(a,t) = u_i(a)$ for all $i \in N$, $a \in A$, and $t \in T$

   (b) $s^*(T) = S^\infty$.

2. *Consider the family of finite Bayesian games $B = (N, A, T, P, \tilde{u})$ with property (a) above and with a common prior P that puts positive probability on each type.. What is the largest set $\{s(t) \,|P(t) > 0\}$ that one can obtain using Bayesian Nash equilibria s for this family of Bayesian games?*

Exercise 3.15 *A game $G = (N, S, u)$ is said to be symmetric if $S_1 = S_2 = \cdots = S_n$ and there is some function $f : S_1 \times S_1^{n-1} \to \mathbb{R}$ such that $f(s_i, s_{-i})$ is symmetric with respect to the entries in s_{-i}, and $u_i(s) = f(s_i, s_{-i})$ for every player i.*

1. *Consider a symmetric game $G = (N, S, u)$ in which S_1 is a compact and convex subset of a Euclidean space and u_i is continuous and quasiconcave in s_i. Show that there exists a symmetric pure-strategy Nash equilibrium (i.e. a pure-strategy Nash equilibrium where every player uses the same strategy).*

2. *Suggest a definition for symmetric Bayesian games, $G = (N, A, \Theta, u, T, p)$, and find broad conditions on such a game G that ensure that G has a symmetric Bayesian Nash equilibrium.*

3. *Consider a Cournot oligopoly with inverse-demand function P and a cost function γ that is common to all firms. Each firm's cost depends on its production level and its idiosyncratic cost parameter, which is drawn from a finite set C. Assume the vector of cost parameters $(c_1, \ldots, c_n)$ is symmetrically distributed. Each firm i privately knows its own cost c_i, but not the others' costs, and independently chooses a quantity q_i to produce. Find conditions on P and γ that guarantee existence of a symmetric Bayesian Nash equilibrium in this game. (Note that the profit of each firm i is $q_i P(q_1 + \cdots + q_n) - \gamma(q_i, c_i)$.)*

Exercise 3.16 *(Due to Jonathan Weinstein) Consider two Bayesian games $G = (N, A, \Theta, T, u, p)$ and $G' = (N, A, \Theta, T, u', p)$, where*

$$u'_i = f_i \circ u_i$$

for each $i \in N$ for some strictly increasing concave function f_i.

1. *Show that*

$$S_i^{\infty}[t_i|G] \subseteq S_i^{\infty}[t_i|G']$$

for all $t_i \in T_i$ and $i \in N$, where $S_i^{\infty}[t_i|G]$ and $S_i^{\infty}[t_i|G']$ are the sets of interim correlated rationalizable action profiles for type t_i in games G and G', respectively.

2. *Provide an example in which the above inclusion is strict.*

Exercise 3.17 *This question asks you to analyze a well-known game, namely* the email game *or* the coordinated attack problem*; we will refer to this game later when we study the global games. Two players play the game*

	a	b
a	θ, θ	$\theta - 1, 0$
b	$0, \theta - 1$	$0, 0$

where $\theta \in \{-1, 2/5, 2\}$. Ex-ante, θ can take either of the values $\theta_0 = -1$ or $\theta_1 = 2/5$, each with probability $1/2$. Player 1 knows the value of θ. If $\theta = \theta_0$, she does not communicate with Player 2. If $\theta = \theta_1$, her computer automatically sends a message to Player 2, indicating that $\theta = \theta_1$. The message is lost with probability $\epsilon \in (0, 1)$ on the way. If the message arrives, Player 2's computer sends back an automatic confirmation of receipt message, which can also be lost with probability $\epsilon \in (0, 1)$. If the message arrives, Player 1's computer sends back an automatic confirmation of receipt message, which can also be lost with probability $\epsilon \in (0, 1)$. This goes on back and forth until a message is lost, when each player chooses an action.

1. *Write the formally as a Bayesian game; define the type space carefully. For each type, briefly describe what that type knows.*

2. *Compute the set of interim correlated rationalizable actions for each type.*

3. *Compute the set of interim correlated rationalizable actions for each type, assuming instead that $\theta_0 = 2$.*

4. *Briefly discuss your findings by comparing the answers in Parts 2 and 3 and the Nash equilibria of the game in which it is common knowledge that $\theta = \theta_1$.*

Bibliography

[1] Battigalli, P. (1999): “Rationalizability in Incomplete Information Games,” EUI working paper 99/15.

[2] Battigalli, Pierpaolo, and Marciano Siniscalchi (2003): "Rationalization and incomplete information," Advances in Theoretical Economics, 3 (1), Article 3.

[3] Brandenburger, A. and E. Dekel (1993): “Hierarchies of Beliefs and Common Knowledge,” *Journal of Economic Theory*, 59, 189-198.

[4] Dekel, E. D. Fudenberg, S. Morris (2006): “Topologies on Types,” *Theoretical Economics*, 1, 275-309.

[5] Dekel, E. D. Fudenberg, S. Morris (2007): “Interim Correlated Rationalizability,” *Theoretical Economics*, 2, 15-40.

[6] Ely, J. and M. Peski (2006): “Hierarchies of belief and interim rationalizability”, *Theoretical Economics* 1, 19–65.

[7] Forges, Françoise (1993): "Five legitimate definitions of correlated equilibrium in games with incomplete information," *Theory and Decision* 35, 277–310.

[8] Harsanyi, J. (1967): “Games with Incomplete Information played by Bayesian Players. Part I: the Basic Model,” *Management Science* 14, 159-182.

[9] Liu, Qingmin (2015): "Correlation and Common Priors in Games with Incomplete Information," *Journal of Economic Theory*, 157, 49–75.

[10] Mertens, J. and S. Zamir (1985): “Formulation of Bayesian Analysis for Games with Incomplete Information,” *International Journal of Game Theory*, 10, 619-632.

[11] Yildiz, Muhamet (2015): "Invariance to representation of information." *Games and Economic Behavior* 94 142-156.

Chapter 4

Extensive-Form Games

This chapter is devoted to extensive-form solution concepts, such as sequential equilibrium, extensive-form rationalizability, and forward induction. The reader is assumed to know more basic concepts, such as multi-stage games, games of perfect information, backward induction, subgame-perfect Nash equilibrium and perfect Bayesian equilibrium.

The key issues are what to believe and how to behave in contingencies that are not supposed to arise. First, if a player assigns zero probability to some contingency, she may plan to take a suboptimal action at that contingency, as her action at the contingency does not affect her expected payoff. For example, in Figure 4.1, at the beginning, Player 2 may plan to take action R believing that Player 1 will play Out and she will not get to move. Indeed, in normal-form representation,

	L	R
In	10,1	0,0
Out	2,2	2,2

strategy R is a best response to Out, and the strategy profile (Out, R) is a Nash equilibrium. However, in the extensive-form game, if Player 2 gets to move, she knows that Player 1 has played In. Hence, she cannot maintain the assumption that Player 1 plays Out. Knowing that Player 1 played In, she must play L. Anticipating all these, Player 1 plays In.

Hence, in extensive-form games, one assumes that a player's strategy is a best response not only at the beginning of the game, but also at every information set. This

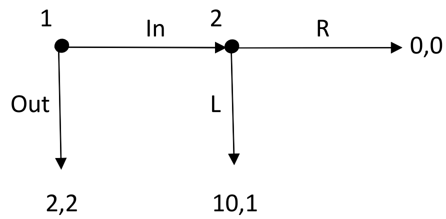

Figure 4.1:

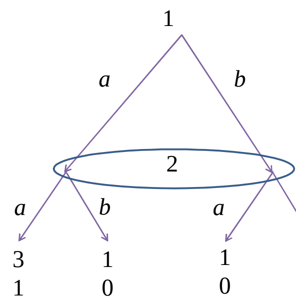

Figure 4.2: A dominance-solvable game

stronger requirement is called *sequential rationality*.

Sequential rationality alone does not put much restriction on what players play in general. For example, consider the game in Figure 4.2. In this game, strategy profile (a, b) is consistent with players' sequential rationality, as b is sequentially rational if Player 2 assigns probability one on the node at the right, believing that Player 1 plays b. But the game is dominance solvable with unique rationalizable strategy profile (a, a). To explore the strategic implications of higher-order knowledge of sequential rationality, one needs to impose further restriction on players' beliefs.

In equilibrium, players must use the Bayes rule to update their beliefs using the other players' strategies. For example, if Player 2 knows that Player 1 plays a, then she must assign probability one on the left node in her information set. Then, sequential rationality leads to the equilibrium (a, a). However, updating beliefs becomes an issue when a player faces uncertainty at an information set that were not supposed to arise.

For example, consider the game in Figure 4.3, where Player 1 has an additional strategy, namely Out. Suppose that, at the beginning, Player 2 believes that Player 1

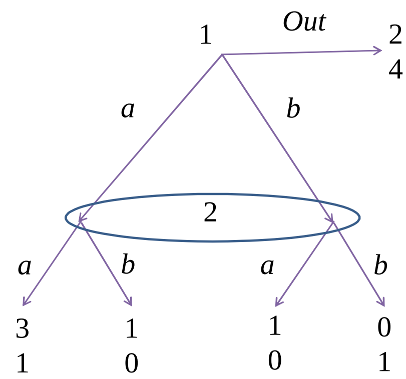

Figure 4.3: Dominance solvable game with outside option

plays *Out* for sure, believing that she will not get to move. When she gets to move, she would learn that a zero probability event occurred. In that case, she cannot use the Bayes rule to update her belief, and any belief is consistent with Bayesian updating at her information set. In particular, she may put probability 1 on the right node, believing that Player 1 played his dominated strategy b. In contrast, if she believed at the beginning that Player 1 plays a with probability 1/2 and *Out* with probability 1/2, she would learn at her information set that Player 1 played either a or b, concluding that Player 1 played a. She would then assign probability one on the left node in her information set.

The game in Figure 4.3 will play a prominent role in this chapter for illustrations. Hence, it is useful to compute its Nash equilibria and rationalizable strategies at the outset. Since the game does not have a proper subgame, all Nash equilibria are subgame-perfect. The game can be written in normal form as follows:

	a	b
a	$3,1$	$1,0$
b	$1,0$	$0,1$
Out	$2,4$	$2,4$

(4.1)

Strategy b is strictly dominated for Player 1; there are no other dominated strategies, yielding the set of rationalizable strategies $\{a, Out\}$ for Player 1 and $\{a, b\}$ for Player 2. There is a strict Nash equilibrium (a, a). Another pure-strategy Nash equilibrium is (Out, b), where Player 2 is indifferent between her strategies. There is also a continuum of mixed strategy Nash equilibria. To characterize those equilibria, observe that b is strictly dominated by a for Player 1. Moreover, strategy a weakly dominates strategy b for Player 2 once we eliminate b for Player 1. Hence, in any mixed Nash equilibrium, Player 1 plays Out for sure. Player 2 mixes between a and b assigning probability

$$\sigma_2(a) \leq 1/2$$

on a, so that Out is a best response for Player 1. Any such strategy profile (Out, σ_2) is a Nash equilibrium.

4.1 Sequential Equilibrium

In general extensive-form game one faces a fundamental problem: how to update one's beliefs at information sets that are not supposed to be reached. At such information sets, the Bayes rule is silent and any belief updating is consistent with Bayesian updating. Unfortunately, if one does not put any restriction on such off-path beliefs, sequential rationality has little bite beyond Nash equilibrium. Sequential equilibrium, due to Kreps and Wilson (1982), restricts off-path beliefs by requiring that those beliefs can be derived by the Bayes rule from some perturbation of equilibrium strategies.

Formally, consider an extensive-form game with finite number of nodes and perfect recall; denote each node by the sequence of moves that lead to the node. For each player i, let H_i be the set of information sets at which player i moves, and write H for the set of all information sets (aka histories) with generic element $h \in H$. Denote the set of available histories at $h \in H_i$ by $A_i(h)$. We will use behavioral mixed strategies:

Definition 4.1 *A* strategy *for a player i is a mapping $\sigma(\cdot|\cdot)$ that yields a probability distribution*

$$\sigma_i(\cdot|h) \in \Delta(A_i(h))$$

on the set of available moves at every history $h \in H_i$. A strategy profile is denoted by σ as usual.

We have so far defined solution concepts as strategy profiles by keeping players' beliefs that lead to those strategies implicit. In dynamic games, the beliefs become an explicit part of the solution, as reasonableness of a solution is closely linked to the plausibility of the beliefs that leads to the solution.

Definition 4.2 *A* belief system *is a list μ of probability distributions on information sets:*

$$\mu(\cdot|h) \in \Delta(h) \qquad (h \in H);$$

for each information set h, μ gives a probability distribution $\mu(\cdot|h)$ on h.

A sequential equilibrium will be a pair of a strategy profile and a belief system:

Definition 4.3 *An* assessment *is a pair (σ, μ) of a strategy profile σ and a belief system μ.*

At any information set h of a player i, a strategy profile yields a probability distribution $\sigma_i(\cdot|h)$ on the moves available at h, describing probabilistically what moves player i takes at that information set, and a belief system yields a probability distribution on the information set itself, describing what player i believes about the past history at that information set, including the Nature's moves so far. As such, an assessment is a complete description of the beliefs and the moves at every contingency. Sequential equilibrium imposes sequential rationality and a consistency requirement on beliefs.

Definition 4.4 *An assessment (σ, μ) is said to be* sequentially rational *if for each information set $h \in H_i$ of each player i, her strategy σ_i is a best response to σ_{-i} given that she is at information set h and given her beliefs $\mu(\cdot|h)$ at that information set.*

Example 4.1 *In Figure 4.3, consider the assessment $((Out, b), \hat{\mu})$, consisting of pure strategy profile (Out, b) and the probability distribution $\hat{\mu}(\cdot|\{a, b\})$ on the information set of Player 2*[1] *where*

$$\hat{\mu}(b|\{a, b\}) = 1.$$

Since Player 1 is supposed to play Out, Player 2's information set $\{a, b\}$ is not supposed to be reached. If Player 2 does get to move, however, she assigns probability 1 on b believing that Player 1 has played b. The assessment $((Out, b), \hat{\mu})$ is sequentially rational. To check it, first consider the information set of Player 1. Since Player 2 plays b with probability 1 according to the assessment, Player 1 gets 1 *from a,* 0 *from b, and* 2 *from Out; hence Out is a sequential best response. Now consider Player 2's information set.*

[1] The beliefs at the information set of Player 1 are trivial, as the information set contains a single node.

Since Player 2's information set is not reached, any strategy is a best response to Out. Sequential rationality is stronger than a simple best response however. Her strategy must be a best response to Out given her belief at her information set. At her information set $\{a, b\}$*, Player 2 assigns probability 1 on* b*. Hence, her expected payoff from* a *is* $u_2(b, a) = 0$*, and her expected payoff from* b *is* $u_2(b, b) = 1$*. Since* $u_2(b, b) \geq u_2(b, a)$*, her strategy is a best response at her information set and the assessment* $((Out, b), \hat{\mu})$ *is sequentially rational. This also shows that the assessment* $((Out, a), \hat{\mu})$ *would not be sequentially rational because* a *is not a best response to Out at Player 2's information set given her belief* $\hat{\mu}(\cdot|\{a, b\})$*.*

As we have seen in the introduction, sequential rationality without any restriction on beliefs has weak predictions. In Figure 4.2, the assessment $((a, R), \mu)$ with $\mu(b|\{a, b\}) = 1$ is sequentially rational although the game is dominance solvable with unique rationalizable strategy profile (a, L). As an equilibrium refinement, sequential equilibrium imposes the following consistency requirement on beliefs.

Definition 4.5 *An assessment* (σ, μ) *is said to be* consistent *if there exists a sequence* (σ^m, μ^m) *of assessments such that*

1. $\sigma_i^m(a|h) \to \sigma_i(a|h)$ *and* $\mu^m(n|h) \to \mu(n|h)$ *as* $m \to \infty$ *for every information set* $h \in H_i$ *for every available move* $a \in A_i(h)$ *at* h *and for every node* $n \in h_i$*;*

2. σ^m *assigns positive probability to each available move at every information set, and*

3. μ^m *is derived from* σ^m *using the Bayes' rule:*

$$\mu^m(n|h) = \frac{\Pr(n|\sigma^m)}{\Pr(h|\sigma^m)}.$$

Recall that σ^m and σ prescribe probability distributions $\sigma_i^m(\cdot|h)$ and $\sigma_i(\cdot|h)$ on the available moves at every information set h of every player i. Likewise, μ^m and μ prescribe probability distributions $\mu^m(\cdot|h)$ and $\mu(\cdot|h)$ on every information set h. The first condition states that $\lim_{m\to\infty} \sigma_i^m(a|h) \to \sigma_i(a|h)$ and $\lim_{m\to\infty} \mu^m(n|h) \to \mu(n|h)$ for all (i, h, n, a) where i moves at h, $n \in h$, and a is available at h. The second condition requires that $\sigma_i^m(a|h) > 0$ everywhere (i.e., every available move is played with positive probability). Under any such strategy profile, every information set is reached with positive probability, and hence one can apply Bayes rule to obtain the beliefs. Observe that if an information set h is reached with positive probability under (σ,μ) consistency reduces to having $b(\cdot|\mu)$ derived from σ using the Bayes' rule; perturbations are required only when there are information sets that are not reached with positive probability.

Example 4.2 *In Example 4.1, the assessment* $((Out,b),\hat{\mu})$ *is consistent. To check this, it suffices to construct a sequence of mixed assessments* (σ^m,μ^m) *that converge to the assessment* $((Out,b),\hat{\mu})$*:*

$$\begin{aligned}\sigma_1^m(Out) &\to 1\\ \sigma_2^m(b) &\to 1\\ \mu^m(b|\{a,b\}) &\to 1,\end{aligned}$$

and the beliefs are consistent with the Bayes rule:

$$\mu^m(b|\{a,b\}) = \frac{\sigma_1^m(b)}{\sigma_1^m(a)+\sigma_1^m(b)}$$

To this end, for each m*, let* $\epsilon = 1/m$*, and consider the mixed strategy profile:* $\sigma_1^m(a)=\epsilon^2$, $\sigma_1^m(b)=\epsilon$, $\sigma_1^m(Out)=1-\epsilon-\epsilon^2$, $\sigma_2^m(a)=\epsilon$ *and* $\sigma_2^m(b)=1-\epsilon$*. Clearly* σ^m *converges to* (Out,b)*. Moreover,*

$$\mu^m(b|\{a,b\}) = \frac{\sigma_1^m(b)}{\sigma_1^m(a)+\sigma_1^m(b)} = \frac{\epsilon}{\epsilon^2+\epsilon} = \frac{1}{1+\epsilon} \to 1.$$

Sequential equilibrium is defined as an assessment that is sequentially rational and consistent:

Definition 4.6 *An assessment* (σ,μ) *is said to be a* sequential equilibrium *if* (σ,μ) *is sequentially rational and consistent.*

A sequential equilibrium is a pair, not just a strategy profile. Hence, in order to identify a sequential equilibrium, one must identify a strategy profile σ, which describes what a player does at every information set, **and** a belief system μ, which describes what a player believes at every information set. In order to check that that (σ,μ) is a sequential equilibrium, one must check that

1. **(Sequential Rationality)** σ_i is a best response to belief $\mu(\cdot|h)$ and the belief that the other players will follow σ_{-i} in the continuation game for every information set $h \in H_i$, and

2. **(Consistency)** there exist trembling probabilities that go to zero such that the conditional probabilities derived from Bayes rule under the trembles approach $\mu(\cdot|h)$ at every information set h.

Example 4.3 *In Example 4.1, the assessment $((Out, b), \hat{\mu})$ is a sequential equilibrium of the game in Figure 4.3. Indeed, as shown above, it is sequentially rational and consistent. One can characterize the sequential equilibria of the game in Figure 4.3 as follows. Since strategy b is strictly dominated, Player 1 will assign positive probability only on a and Out. In a sequential equilibrium, if Player 1 plays a with positive probability, then Player 2's information set is reached with positive probability, and Player 2 must assign probability 1 on a by consistency, and thus she must play a with probability 1 by sequential rationality. Therefore, Player 1 must play a with probability 1 as a best response. This leads to the sequential equilibrium*

$$((a, b), \mu_a)$$

where $\mu_a(a|\{a,b\}) = 1$. In all other sequential equilibria, Player 1 plays Out with probability 1: $\sigma_1(Out) = 1$. For this to be sequentially rational, Player 2 must play a with probability less than 1/2: $\sigma_2(a) \leq 1/2$. There is a continuum of such sequential equilibria. First, there is a continuum of pure-strategy sequential equilibria

$$\left((Out, b), \mu_q\right) \text{ where } \mu_q(a|\{a,b\}) = q,\ q \leq 1/2.$$

Sequential rationality of Player 2 reduces to the condition that $q \leq 1/2$, and any such belief can be obtained consistently from a perturbation with $\sigma_1^\epsilon(a) = q\epsilon$ and $\sigma_1^\epsilon(a) = (1-q)\epsilon$; perturbation for the corner case $q = 0$ is as in the previous example. For $q = 1/2$, there is also a continuum of sequential equilibria in mixed strategies:

$$\left((Out, \sigma_2), \mu_{1/2}\right) \text{ where } \sigma_2(a) \leq 1/2,\ \mu_{1/2}(a|\{a,b\}) = 1/2,$$

where the condition $\sigma_2(a) \leq 1/2$ ensures sequential rationality for Player 1.

The sequential equilibria in this example are depicted in Figure 4.4. There are two components of equilibria; a component is a set of equilibria that lead to a common outcome. One of them is the isolated point at $(1, 1)$, which corresponds to the pure-strategy

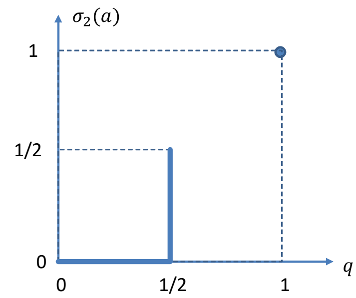

Figure 4.4: Sequential equilibria of the game in Figure 4.3 ($q = \mu(a| \{a, b\})$).

equilibrium $((a, a), \mu_a)$. The outcome of this equilibrium is (a, a). The second component is the L-shaped set at the bottom, corresponding to the continuum of equilibria derived above. All these equilibria lead to the same outcome: *Out*. This illustrates several properties of sequential equilibria:

1. All sequential equilibrium strategy profiles are Nash equilibria. Indeed, if (σ, μ) is a sequential equilibrium, then σ is a subgame-perfect Nash equilibrium. It so happens that in this example, each Nash equilibrium is supported by a sequential equilibrium.

2. There is typically a continuum of sequential equilibria, multiplicity often arising multiplicity of beliefs at off-path histories.

3. But for generic payoff functions (from terminal nodes to real numbers), there are finitely many components of sequential equilibria, and all equilibria in each component leads to the same outcome. There are finitely many sequential equilibrium outcome, generically.

As in the case of subgame-perfect Nash equilibrium, sequential equilibrium depends on the details of the game tree, as illustrated in the next example.

Example 4.4 *Consider the game in Figure 4.7. As discussed before, this game can be viewed as representing the same strategic situation as the one depicted in Figure 4.3. In contrast to the game in Figure 4.3, this game has a unique subgame perfect Nash equilibrium:* (In, a, a). *Indeed,* (a, a) *is the only Nash equilibrium in the proper subgame,* *which is dominance solvable; hence Player 1 plays In at the beginning. This leads to the unique sequential equilibrium: $((In, a, a), \mu)$ where $\mu(In, a|h_2) = 1$ at the information set h_2 of Player 2. To see this, observe that sequential rationality of Player 1 requires that Player 1 plays a after playing In. Of course, if Player 1 plays In with positive probability in a sequential equilibrium, then consistency requires that Player 2 assigns probability 1 on In, a at her information set, and then she must play a with probability 1 by sequential rationality. Player 1 must then play In with probability 1. This leaves out only one possibility: Player 1 plays Out for sure. Then, his strategy must be (Out, a) by sequential rationality—as established above. But then consistency requires that $\mu(In, a|h_2) = 1$ at the information set h_2 of Player 2, which then leads to Player 1 playing In, leading to a contradiction. To see that we must have $\mu(In, a|h_2) = 1$, consider any perturbation $\sigma_1^{\epsilon_1,\epsilon_2}$ of Player 1's strategy where $\sigma_1^{\epsilon_1,\epsilon_2}(In|\varnothing) = \epsilon_1$ and $\sigma_1^{\epsilon_1,\epsilon_2}(b|In) = \epsilon_2$ for some $\epsilon_1, \epsilon_2 \in (0, 1)$, where $\varnothing$ denotes the empty history at the beginning of the game. Clearly, the conditional belief under the perturbed strategy is*

$$\mu^{\epsilon_1,\epsilon_2}(In, a|h_2) = \frac{\epsilon_1(1-\epsilon_2)}{\epsilon_1(1-\epsilon_2) + \epsilon_1\epsilon_2} = 1 - \epsilon_2$$

at the information set of Player 2. As $(\epsilon_1, \epsilon_2) \to (0, 0)$, $\mu^{\epsilon_1,\epsilon_2}(In, a|h_2) \to 1$. Therefore, we must have $\mu(In, a|h_2) = 1$.

I will conclude this section by stating some further properties of sequential equilibria, and comparing it to Perfect Bayesian equilibrium.

1. Every finite game has at least one sequential equilibrium.

2. Sequential equilibrium is upper-hemicontinuous with respect to the payoff functions for a given game. However, as Example 4.4 illustrates, sequential equilibrium may be sensitive to the way game tree is drawn, and it is not a well-defined function of reduced normal form.

Sequential equilibrium can be compared to perfect Bayesian equilibrium as follows. First, both solution concepts pick assessments. Their domains are different. Sequential equilibrium is defined for arbitrary extensive-form games with a finite number of nodes. Perfect Bayesian equilibrium is defined on a specific class of dynamic Bayesian games, which are specific extensive-form games. Within the class of finite games, perfect Bayesian equilibrium applies to a narrower class of games. In this class, sequential equilibrium refines perfect Bayesian equilibrium, as consistency requirement of sequential equilibrium is stronger than the conditions perfect Bayesian equilibrium imposes on beliefs. The two solution concepts coincide in some specific classes of dynamic Bayesian games, e.g., when each player has at most two types or there are at most two periods (as in signaling games). In applications, perfect Bayesian equilibrium is used frequently as it is often easier to apply (for it circumvents consistency). The finiteness restriction in the domain of sequential equilibrium turns out to be an important one; whereas finiteness does not play a role in the domain of perfect Bayesian equilibrium.

4.2 Equilibrium Refinements in Normal Form

This section is devoted to some equilibrium refinements, such as (trembling-hand) perfect equilibrium and proper equilibrium. Although these concepts are defined for normal-form games, they are motivated by extensive-form games. Throughout this section, we will consider a finite game $G = (N, S, u)$ in normal form.

4.2.1 Perfect Equilibrium

Some Nash equilibria crucially rely on the players' assumption that the other players all play their Nash equilibrium strategies. For example, consider the game

	a	b
a	$1,1$	$0,0$
b	$0,0$	$0,0$

(4.2)

In this game, (b, b) is a Nash equilibrium even though (a, a) is a dominant-strategy equilibrium. A player will not play weakly dominated strategy b if she has any doubt that the other player will play b. In equilibrium (b, b), if a player's hand is anticipated to tremble with some small probability $\epsilon > 0$ and play an unintended strategy (in this case a), then the other player will deviate from her equilibrium strategy, regardless of how small ϵ is. In that sense, this equilibrium is highly fragile.

Selten (1975) introduced the concept of perfect equilibrium to rule out such fragile equilibria. Perfect equilibrium requires that a player's equilibrium strategy remains to be a best response to *some* perturbation of the other players' equilibrium strategy.[2] (A strategy σ_i is said to be totally mixed or completely mixed if it assigns positive probability to each strategy.)

Definition 4.7 *A (mixed) strategy profile σ is a* (trembling-hand) perfect equilibrium *if there exists a sequence σ^m of totally mixed strategy profiles such that*

1. *$\sigma^m \to \sigma$ and*
2. *σ_i is a best response to each σ_{-i}^m for each i and m.*

The key property is that the equilibrium strategy σ_i remains to be a best response to *some* perturbation σ_{-i}^m that puts positive probability on each strategy. For example, in game (4.2), the equilibrium (b, b) is not a perfect equilibrium, because b is not a best response to *any* perturbation σ_j^m of b so long as σ_j^m puts positive probability on a. The dominant-strategy equilibrium (a, a) is a perfect equilibrium because players do not have incentive to deviate even if the other players slightly deviate from their equilibrium strategy. Here, the dominant-strategy equilibrium remains a best response to all perturbations, but this stronger requirement is not needed.

Example 4.5 *Consider the game (4.1), which is normal-form representation of the game in Figure 4.3 and a reduced normal-form representation of the game in Figure 4.7. The Nash equilibrium (Out, b) is a perfect equilibrium. To establish this, one only needs to construct one perturbation against which Out and b are best responses. Write $\epsilon = 0.5/m$ and let $\sigma_1^m(a) = \epsilon^2$, $\sigma_1^m(b) = \epsilon$, $\sigma_1^m(Out) = 1 - \epsilon - \epsilon^2$, and $\sigma_2^m(a) = \epsilon$. As $m \to \infty$, $\sigma_1^m(Out) \to 1$ and $\sigma_2^m(b) \to 1$, satisfying the first condition in the definition. To check the second condition, first observe that strategy Out remains a best response to σ_2^m. To check that strategy b remains a best response to σ_1^ϵ, observe that Player 2 is indifferent between her strategies when the likely strategy Out is played. The best response depends on relative likelihood of unlikely strategies a and b. Since b is much more likely than a, the best response depends on which strategy fairs better against b. Since b is the best response to b, strategy b is a best response to σ_1^m.*

It is useful to observe some general properties of perfect equilibrium:

[2] Selten has also introduced subgame-perfect Nash equilibrium in an earlier paper.

1. By definition, a perfect equilibrium must be a Nash equilibrium. Indeed, since σ_i is a best response to σ_{-i}^m and $\sigma_{-i}^m \to \sigma_{-i}$, by the Maximum Theorem, σ_i is a best response to σ_{-i}.

2. Conversely, any strict Nash equilibrium is a perfect equilibrium because a strict best response remains a best response when the other players' strategies are perturbed.

3. Likewise, any Nash equilibrium σ in totally mixed strategies is a perfect equilibrium because one can set $\sigma^m = \sigma$ for each m.

Therefore, perfect equilibrium is a refinement of Nash equilibrium and has a bite only when a player plays a weak best response against an equilibrium strategy that assigns zero probability to some strategies. It does not have a bite in generic normal-form games. However, such weak best responses are common in Nash equilibria of dynamic games, and perfection is often used to refine Nash equilibrium in dynamic games. The next examples illustrates the subtleties in applying perfect equilibrium to dynamic games.

Example 4.6 *Consider, again, the normal-form game (4.1). As discussed in the introduction, this game has a continuum of Nash equilibria, and all of them are perfect. The strict Nash equilibrium (a, a) is a perfect equilibrium, and we have already seen that (Out, b) is a perfect equilibrium. For the remaining Nash equilibria (Out, σ_2) with $0 < \sigma_2(a) \leq 1/2$, one can construct perturbations σ_1^m, σ_2^m as $\sigma_1^m(a) = \sigma_1^m(b) = \epsilon$, $\sigma_1^m(Out) = 1 - 2\epsilon$ and $\sigma_2^m = \sigma_2$ where $\epsilon = 0.25/m$.*

The game in Figure 4.3 does not have a proper subgame. As we have seen above, all Nash equilibria are subgame perfect and sequential. Perfect equilibrium joins them as an equilibrium refinement that does not have a bite. However, in the game Figure 4.3, (a, a) is the unique subgame-perfect Nash equilibrium and thus the unique sequential equilibrium strategy profile. As a normal-form refinement, perfect equilibrium still does not refine any Nash equilibrium when applied to reduced normal-form. It remains to be toothless when it is applied to the normal-form representation of Figure 4.3, where strategy Out is split to two equivalent strategies Out, a and Out, b. In that sense, perfect equilibrium can be viewed as even weaker than subgame-perfect Nash equilibrium.

Perfect equilibrium is applied to "agent-normal-form" representation of the extensive form games, in which each information set of each player i is considered as a separate player with payoffs equal to the payoffs of player i. This is illustrated next.

Example 4.7 *Consider the extensive-form game in Figure 4.7. The agent-normal form game for this game is a three-player game: Player 1 at the beginning (denoted as Player 1.1), Player 1 after playing In (denoted as Player 1.2) and Player 2. The three player game is represented as follows*

	a	b
a	$3,3,1$	$1,1,0$
b	$1,1,0$	$0,0,1$

In

	a	b
a	$2,2,4$	$2,2,4$
b	$2,2,4$	$2,2,4$

Out

where Player 1.1 chooses the matrices, Player 1.2 chooses the rows, and Player 2 chooses the columns; the first and the second payoffs are the payoffs of Players 1.1 and 1.2, respectively, and the last payoff is the payoff of Player 2. The pure-strategy Nash equilibria are $(In, a, a), (Out, b, b), (Out, a, b), (Out, b, a)$*. Here,* (Out, b, a) *is a Nash equilibrium because the selves 1.1 and 1.2 cannot coordinate their actions; such miscoordination of selves can lead to new equilibria as in this example. The only perfect equilibrium is the strict Nash equilibrium* (In, a, a)*, corresponding to the unique subgame-perfect Nash equilibrium. For example, to see why* (Out, a, b) *is not a perfect equilibrium, consider any perturbation* $\left(\sigma_{1,1}^{\epsilon_{11}}, \sigma_{1,2}^{\epsilon_{12}}, \sigma_2^{\epsilon_2}\right)$ *where* $\sigma_{1,1}^{\epsilon_{11}}(In) = \epsilon_{11}$*,* $\sigma_{1,2}^{\epsilon_{12}}(b) = \epsilon_{12}$*, and* $\sigma_2^{\epsilon_2}(a) = \epsilon_2$ *for some* $(\epsilon_{11}, \epsilon_{12}, \epsilon_2)$*. For Player 2, the equilibrium strategy* b *is not a best response to* $\left(\sigma_{1,1}^{\epsilon_{11}}, \sigma_{1,2}^{\epsilon_{12}}\right)$*. Indeed,*

$$u_2\left(a, \sigma_{1,1}^{\epsilon_{11}}, \sigma_{1,2}^{\epsilon_{12}}\right) - u_2\left(b, \sigma_{1,1}^{\epsilon_{11}}, \sigma_{1,2}^{\epsilon_{12}}\right) = \epsilon_{11}\left(1 - 2\epsilon_{12}\right),$$

which is positive whenever $\epsilon_{12} < 1/2$*. Now, since Player 2 is indifferent when Player 1 plays Out, her move matters only when Player 1.1 trembles (hence the term* ϵ_{11}*). Her best response depends only on what Player 1.2 plays. However, since the trembles of Players 1.1 and 1.2 are independent, the probability of tremble for Player 1.2 is still* ϵ_{12}*. Therefore, Player 2 plays a best response in the proper subgame disturbing the putative perfect equilibrium* (Out, a, b)*.*

Examples 4.5 and 4.7 illustrate the power of applying perfect equilibrium to agent-normal representation. When applied to normal form representation, perfect equilibrium considers trembles to strategies, which maps information sets to moves, allowing correlation between the trembles at various nodes. For example, when applied to normal form representation, (Out, a, b) is a perfect Nash equilibrium because b remains a best response when tremble In, b is more likely than tremble In, a, assuming that it becomes likely that Player 1 trembles one more time if he trembles at the beginning. When applied to agent-normal form representation, trembles by different selves are assumed to be independent, and hence such correlated trembles are not allowed. In that case, tremble In, a must be "infinitely" more likely then the tremble In, b because the latter involves further unlikely trembles.

Throughout we will assume that the perfect equilibrium is applied to the agent-normal form. When applied to agent-normal form it is clear that a perfect equilibrium must be a subgame-perfect Nash equilibrium because the trembles within a given subgame are independent of the trembles that reach the subgame, implying that the players play best response to each other within each subgame. The perfect equilibrium is related to sequential equilibrium as follows.

1. If σ is a perfect equilibrium of the agent-normal form, then (σ, μ) is a sequential equilibrium for some belief system μ.

2. Conversely, given any game tree and information partition, for generic payoff functions, if (σ, μ) is a sequential equilibrium, then σ is a perfect equilibrium of the agent-normal form game.

When there are ties, perfect equilibrium can be stronger than sequential equilibrium. For example, in simultaneous action games, sequential equilibrium coincides with Nash equilibrium while perfect equilibrium refines it. For example, in game (4.2), (b, b) is a sequential equilibrium strategy profile but it is not perfect.

4.2.2 Proper Equilibrium

Perfect equilibrium does not put any restriction on trembles. In particular, as we have seen in Example 4.5, a perfect equilibrium (Out, b) may rely on the assumption that tremble to b is more likely than trembling to a although a dominates b. Myerson (1978) argues that the costlier trembles must be less likely. In static games, this amounts to arguing that players will be more careful when their mistakes are costlier. More importantly, in normal-form representation of dynamic games, this relates to the fact that trembles that involve deviations in multiple information sets are costlier as they involve deviating from best responses in multiple steps. This notion is formalized by proper equilibrium, which is a normal-form solution concept.

Definition 4.8 *For any game $G = (N, S, u)$, a strategy profile σ is a proper equilibrium if there exist a sequence of totally mixed strategy profiles σ^ϵ such that*

1. $\sigma^\epsilon \to \sigma$ *as* $\epsilon \to 0$, *and*

2. *for any $i \in N$ and $s_i, s_i' \in S_i$,*

$$u_i\left(s_i, \sigma_{-i}^\epsilon\right) < u_i\left(s_i', \sigma_{-i}^\epsilon\right) \Longrightarrow \sigma_i^\epsilon\left(s\right) \leq \epsilon \sigma_i^\epsilon\left(s_i'\right).$$

Observe that conditions 1 and 2 imply that σ_i is a best response to σ_{-i}^ϵ.[3] Hence, they together imply that if σ is a proper equilibrium, then it is also a perfect equilibrium. Proper equilibrium refines perfect equilibrium by restricting the allowable perturbations as in condition 2. Fewer perturbations justify (if anything) fewer equilibria, and hence proper equilibrium can be stronger. Since strict Nash equilibria remain best response to all perturbations, they are always proper.

Example 4.8 *Consider the game (4.1). As we have seen before, there is a continuum of Nash equilibria, which are all perfect. Only the strict Nash equilibrium (a, a) is proper. For example, consider the Nash equilibrium (Out, b). There cannot be a sequence σ^ϵ with $(\sigma_1^\epsilon(Out), \sigma_2^\epsilon(b)) \to (1, 1)$ as in the definition above. To see this, observe that*

$$u_1(a, b) > u_1(b, b);$$

hence it must be that $u_1(a, \sigma_2^\epsilon) > u_1(b, \sigma_2^\epsilon)$ when ϵ is small. Thus,

$$\sigma_1^\epsilon(b) \leq \epsilon \sigma_1^\epsilon(a) \text{ and } \sigma_1^\epsilon(a) \leq \epsilon \sigma_1^\epsilon(Out).$$

Then, b becomes "infinitely less likely" than a as ϵ vanishes, yielding $u_2(b, \sigma_1^\epsilon) < u_2(a, \sigma_1^\epsilon)$. Thus, by condition 2, $\sigma_2^\epsilon(b)$ must go to zero, a contradiction. Similarly, one can show that none of the Nash equilibria (Out, σ_2) is proper.

[3] If $\sigma_i(s_i) > 0$, then s_i must be a best response to σ_{-i}^ϵ for small ϵ; for otherwise $\sigma_i^\epsilon(s_i)$ would converge to zero.

Proper equilibrium is special among the equilibrium refinements that we considered in this chapter as follows. The extensive-form games in Figures 4.3 and 4.7 are strategically equivalent in the sense that they have the same reduced normal-form representation as in (4.1). However, all extensive-form solution concepts (namely subgame-perfect Nash equilibrium, sequential equilibrium and agent-normal form perfect equilibrium) treated these games differently. While they have not eliminated any of the Nash equilibria in Figure 4.3, they picked a unique equilibrium (a, a) in Figure 4.7. In contrast, proper equilibrium uniquely selected equilibrium (a, a) when applied to the reduced normal-form. The next result shows that this is not a coincidence:

Theorem 4.1 (Kohlberg and Mertens, 1986) *If σ is a proper equilibrium of a normal-form game G, then every extensive-form game with reduced normal-form representation G has a sequential equilibrium $(\hat{\sigma}, \hat{\mu})$ that induces strategy profile σ on G (i.e., $\sigma_i(s_i)$ is equal to the sum of the probabilities $\hat{\sigma}_i$ assigns to the strategies that are equivalent to s_i).*

Thus, proper equilibrium provides a refinement of sequential equilibrium that ignores the "strategically irrelevant" details in extensive-form representations. Van Damme (1984) strengthens this result further by considering a refinement of sequential equilibrium, called "quasi-perfect equilibrium". Proper equilibrium achieves this invariance property by implicitly making deviations from best response at more information sets infinitely less likely—as those deviations yield lower payoffs.

In summary, proper equilibrium is the strongest solution concept that we have learned: every proper equilibrium is a perfect equilibrium; every perfect equilibrium applied to agent-normal form game is a sequential equilibrium strategy profile, and every sequential equilibrium strategy profile is a subgame-perfect Nash equilibrium. In finite games, there exists a proper equilibrium, showing that all these solution concepts are non-empty in finite games.

4.2.3 Strategic Stability

Kohlberg and Mertens (1986) introduces a concept of strategic stability in order to obtain a solution concept that is invariant to extensive-representation of "equivalent" strategic situations and satisfy sequential rationality as well as iterative elimination of strictly dominated strategies. This section briefly introduces their concept of strategic stability. This concept is also related to forward induction that will be introduced later.

Proper equilibrium appears to fit the bill above: it is a refinement of sequential rationality that is invariant to extensive form representation of the same reduced form normal game (by Theorem 4.1), and it does not involve weakly dominated strategies by definition. Proper equilibrium does not satisfy a stronger invariance property—illustrated in the next example.

Example 4.9 *Consider the following games:*

	a	b
a	$3,3$	$0,0$
b	$0,0$	$1,1$
Out	$2,2$	$2,2$

	a	b
a	$3,3$	$0,0$
b	$0,0$	$1,1$
Out	$2,2$	$2,2$
mix	$\frac{7}{3},\frac{7}{3}$	$\frac{4}{3},\frac{4}{3}$

The game on the right is created from the one on the left by introducing a mixed strategy that puts probabilities 2/3 and 1/3 on Out and a, respectively. The game on the left has two proper equilibria (a,a) and (Out,b) in pure strategies while the game on the right has a unique proper equilibrium (a,a). In the game on the left, proper equilibrium (Out,b) is justified by trembles $(\varepsilon^2,\varepsilon,1-\varepsilon-\varepsilon^2)$ on (a,b,Out) as b is a better response than a to strategy b. On the right, mix strictly dominates b. Hence, for any proper equilibrium, trembling probability for b must be less than ε times the trembling probability on mix. Since a is a better response than b against mix, Player 2 must play a in any proper equilibrium, leading to (a,a) as the only proper equilibrium.

One may see this as a deficiency of proper equilibrium (and the other equilibrium refinements above). It appears that introduction of a mixed strategy in the form of a "redundant" strategy should not matter. Nonetheless, this leads to fewer proper equilibria, showing that proper equilibrium is sensitive to introduction of mixed strategies as pure strategies.[4] Kohlberg and Mertens (1986) requires invariance with respect to introduction of mixed strategies as pure strategies in addition to invariance to extensive-form representation with identical reduced normal form.

[4] The same criticsm applies to other refinement, as one can construct multistage representation of the above games in which proper equilibria coincide with subgame-perfect equilibria (and also with perfect and sequential equilibria.

Their key technical result that motivates their formal definitions is the following structure theorem for Nash equilibrium. (This result appears to be more useful than strategic stability notions they introduce.) To state the result, fix a set $N = \{1, \ldots, n\}$ of players, a finite set $S = S_1 \times \cdots \times S_n$ of strategy profiles, and set $\Gamma = \left(\mathbb{R}^S\right)^N$ of profiles $(u_1 \ldots, u_n)$ of payoff functions $u_i : S \to \mathbb{R}$. Let

$$G^{NE} = \{(u, \sigma) \in \Gamma \times \Sigma | \sigma \text{ is a Nash equilibrium of } (N, S, u)\}$$

be the graph of Nash equilibrium correspondence, where Σ is the set of all mixed strategy profiles. Let $p : G^{NE} \to \Gamma$ be the projection mapping, where $p(u, \sigma) = u$.

Theorem 4.2 (Structure Theorem for Nash Equilibrium) *The projection mapping $p : G^{NE} \to \Gamma$ from Nash equilibrium graph to the space of games is homotopic to a homeomorphism.*

That is, one can deform the projection mapping continuously to obtain a one-to-one continuous mapping from the graph to the space of games where the inverse is also continuous. In particular, they construct a homeomorphism[5] $\phi : \Gamma \to G^{NE}$ and a continuous mapping q on $[0, 1] \times G^{NE}$ such that $q(0, \cdot) = p \circ \phi$ and $q(1, \cdot)$ is the identity mapping. That is, $p \circ \phi$ can be continuously deformed to obtain the identity mapping. All these mappings continuously extend to one-point compactification of the spaces.

In other words, the Nash equilibrium graph can be deformed continuously to obtain the space of payoff functions. To aid visualization schematically, Figure 4.5 plots the Nash equilibrium graph as if both the space of games and the set of Nash equilibria are one-dimensional. As one can see, the one dimensional graph can be transformed to the line, the space of payoff function, deforming it continuously. The number of Nash equilibria is finite and odd generically, but the "knife-edge" cases are important, as they usually correspond to the normal-form representation of extensive form games. Consider one of them, u^*, as plotted in the figure. In this game there are two isolated Nash equilibria and two connected components with a continuum of equilibria. For each isolated equilibrium, there remains a nearby equilibrium when we perturb the payoffs regardless of how we perturb (i.e. whether to the left or to the right). These equilibria are stable with respect to all perturbations. The component at the bottom does not

[5] A continuous one-to-one and onto function with continuous inverse.

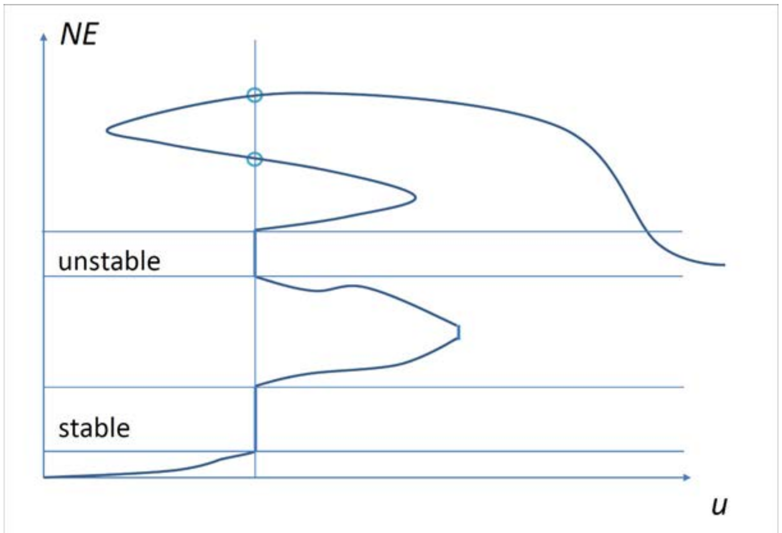

Figure 4.5: Structure of Nash equilibrium graph

contain any stable equilibrium in that sense; the top equilibrium disappears when we move to the left and the bottom equilibrium disappears when we move to the right, and all the remaining equilibria disappear in either direction. However, the component itself—as a set of equilibria— is stable in that there is a nearby equilibrium to the set regardless of how we perturb the game. The other component is not stable, as there is no nearby equilibrium to this set when we move to the left.

Kohlberg and Mertens define strategic stability as a property of sets of equilibria, as follows.

Definition 4.9 *A set $\hat{\Sigma}(u)$ of Nash equilibria of a game (N, S, u) is said to be* stable with respect to all perturbations *if for every $\epsilon > 0$, there exists $\delta > 0$ such that for every payoff function u' with $\|u - u'\| < \delta$ there exist $\sigma \in \hat{\Sigma}(u)$ and a Nash equilibrium σ' of game (N, S, u') such that $\|\sigma - \sigma'\| < \epsilon$.*

That is, there is always a nearby equilibrium to the set $\hat{\Sigma}(u)$ regardless of how we perturb the game. In the figure, each isolated equilibria and the component at the bottom are examples of such equilibria. The unstable component is an example of a set that does not satisfy this property. Observe that the set of Nash all Nash equilibria satisfies this property because the Nash equilibrium correspondence is upper-hemicontinuous.

Strategic stability requires that the considered set is also "minimal". There are various strategic stability notions, depending on what perturbations are allowed:

Definition 4.10 *A set $\hat{\Sigma}(u)$ of Nash equilibria of a game (N, S, u) is said to be* hyper-stable *if it is stable with respect to all perturbations and has no proper subset that is also stable with respect to all perturbations.*

In the figure each of isolated equilibria forms a hyper-stable set. For the bottom component, the two-element set consisting of the end-points of the component is a hyper-stable set.

Requiring stability with respect to all perturbation may require that we include some equilibria in the set that may be deemed as implausible. Kohlberg and Mertens (1986) introduce stability with respect to more restrictive sets of perturbations. One set of perturbations allow perturbations that put minimum trembling probabilities where the probabilities may depend on the strategy. Stability with respect to such perturbation leads to full-stability. A set $\hat{\Sigma}(u)$ of Nash equilibria of a game (N, S, u) is said to be *fully stable* if it is stable with respect to all perturbations by trembles as above and has no proper subset that is also stable with respect to such perturbations. Any fully stable set of a normal-form game contains a proper equilibrium of that game. In particular, for every-extensive-form game, any stable set contains a proper equilibrium of the extensive-form game, and hence it contains a perfect and a sequential equilibrium of the game. Unfortunately for Kohlberg and Mertens, one may need to include Nash equilibria in weakly dominated strategies to obtain a fully stable set. Hence, they consider a weaker concept as their concept of strategic stability.

Definition 4.11 *A $\hat{\Sigma}(u)$ of Nash equilibria of a game (N, S, u) is said to be* stable *if for every $\epsilon > 0$ and every completely mixed strategy profile σ^* there exists $\delta_0 > 0$ such that for all $\delta = (\delta_1, \ldots, \delta_n) \in (0, \delta_0)^n$ there exist $\sigma \in \hat{\Sigma}(u)$ and a Nash equilibrium σ' of game (N, S, u') such that $\|\sigma - \sigma'\| < \epsilon$ where $u'_i(s) = u_i(s(1-\delta) + \delta\sigma^*)$ for each s, and $\hat{\Sigma}(u)$ has no proper subset that satisfies this property.*

Strategic stability may seem similar to perfect equilibrium. It differs in two ways. First it is a property of a set of equilibria, while perfectness is a property of an equilibrium. Second, stability requires that equilibrium strategies remain best response to

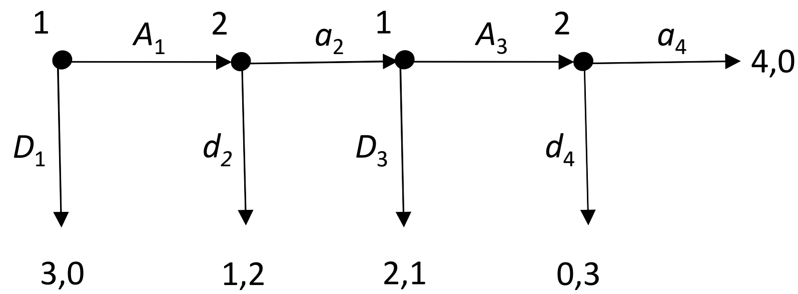

Figure 4.6:

trembling to *all* mixed strategy profiles, while perfect equilibrium requires that they remain best response to trembling to *some* mixed strategy profile. Stable sets do not contain weakly dominated strategies by definition. However, they may not contain a sequential equilibrium strategy profile. They do have the following appealing properties:

- Every stable set contains a stable set of a game obtained by deletion of a weakly dominated strategy.
- Every stable set contains a stable set of any game obtained by deletion of a strategy that is an inferior response to all the equilibria of the set.

The latter property is used formalizations of forward induction, such as the intuitive criterion of Cho and Kreps.

4.3 Backward Induction & Iterative Conditional Dominance

In perfect-information games, iterative application of sequential rationality can lead to sharp predictions. In particular, when the game has finite horizon and no relevant ties, backward induction amounts to applying sequential rationality iteratively, starting from the end decision nodes. In those games, backward induction leads to a unique solution.

For example, consider the game in Figure 4.6, which is taken from Battigalli (1997). One can apply backward induction as follows. At the last round, d_4 is picked because

Player 2 prefers d_4 to a_4 at that node. Then, Player 1 expects to get 0 from playing A_3 and 2 from playing D_3; hence D_3 is picked at that node. Similarly, d_2 and D_1 are picked in the subsequent rounds. Therefore, the backward induction solution is (D_1D_3, d_2d_4), yielding outcome D_1.

In this example, backward induction applies sequential rationality iteratively to eliminate actions that are never a best response at the time the action is to be taken. One can apply this idea beyond the perfect-information games with finite horizon. The resulting solution concept is called *iterative elimination of conditionally dominated actions*. Within the context of multi-stage games[6], it is formalized by Fudenberg and Tirole (1991) as follows.

Definition 4.12 *In a multi-stage game, an action a at a history h is conditionally dominated if in the subgame that start at h every strategy that picks a at h is strictly dominated.*

For example, in Figure 4.6, at the last stage, the subgame is a one-person game in which Player 2 has two strategies a_4 and d_4. Since a_4 is dominated in this one-person game, it is conditionally dominated in the larger game. In contrast, the move a_2 at history A_1 is not conditionally dominated. The subgame that starts at A_1 is the three-period game that starts with Player 2 choosing between a_2 and d_2. In this game, the strategy a_2d_4 is a best response to the strategy A_3, and thus it is not strictly dominated, while strategy a_2a_4 is strictly dominated by D_2d_4. Since there is an undominated strategy that picks move a_2, the move a_2 is not conditionally dominated. Indeed, in the larger game, the only conditionally dominated move is a_4.

Definition 4.13 *Iterative elimination of conditionally dominated actions is the iterative procedure in which all conditionally dominated moves are eliminated until no conditionally dominated move remains.*

For example, in Figure 4.6, only action a_4 is conditionally dominated, and it is eliminated in the first round. After the elimination, at the last stage Player 2 has only one move: d_4. Then, in the second round, the move A_3 is conditionally dominated.

[6] These are the games in which at each stage a subset of players move simultaneousl and all previous moves are observable.

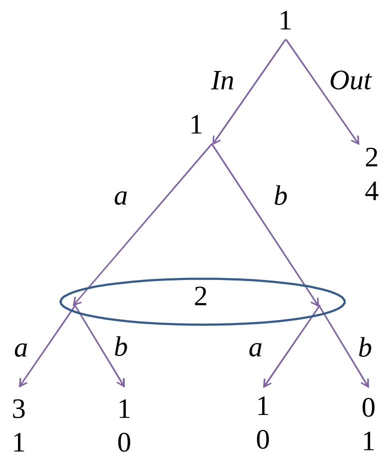

Figure 4.7: A variation of Game in Figure 4.3

(Now, at the subgame after A_1a_2, Player 1 has strategies A_3 and D_3 while Player 2 has only one strategy d_4, and D_3 is strictly better than A_3 for Player 1.) There is no other conditionally dominated strategies at this round. However, in the next rounds, actions a_2 and A_1 become conditionally dominated and eliminated (in that order).

This elimination procedure mimics backward induction leading to the backward-induction solution. This is not a coincidence. In perfect-information games with finite horizon and with no relevant ties, iterative elimination of conditionally dominated actions coincides with backward induction (by definition). When there are ties so that a player has multiple best responses at some node, the two solution concepts differ. As an equilibrium solution concept, backward induction picks a move (or a mixed action) that is a best response at that node, and in the following rounds it is assumed (by all players) that that action will be played at that node. In contrast, as a disequilibrium solution concept, iterative elimination procedure keeps *all* best responses at that node. In the subsequent round, at various nodes, players could maintain distinct beliefs about what will be played in this particular nodes, and this may lead to elimination of fewer actions. In general, iterative elimination of conditionally dominated actions is a weaker solution concept: all backward induction solutions survive and some surviving may not be played under any backward induction solution.

For another illustration, consider the game in Figure 4.3. This game is not a multi-

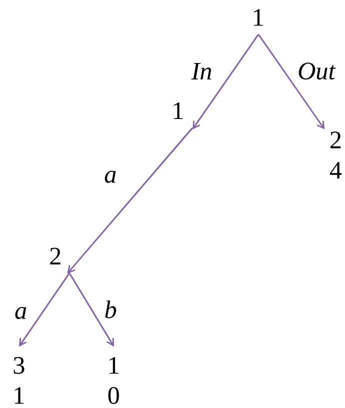

Figure 4.8: Conditional dominance in Figure 4.7

stage game, and it is not clear how apply to procedure in such games. However, one can represent the same strategic situation as a multi-stage game as in Figure 4.7; they are strategically equivalent in the sense that they have the same reduced normal-form representation.

Example 4.10 *Consider the game in Figure 4.7. The game in the proper subgame is dominance solvable, and the unique solution is* (a, L)*. Given* (a, L)*, Player 1 prefers* In*. The unique subgame-perfect Nash equilibrium is* (In, a, L)*. In this game, the same unique solution is obtained by the iterative elimination of conditionally dominated action. At history* In*, the move* b *is conditionally dominated because substrategy* a *strictly dominates the substrategy* b *at the subgame that starts at* In*. Once* b *is eliminated, the remaining game is another multistage game as in Figure 4.8. In this game, action* R *is conditionally dominated and eliminated. Finally, action Out is eliminated, and the only remaining strategy profile is the unique subgame-perfect equilibrium.*

Once again, in multi-stage games, subgame-perfect Nash equilibrium actions survive iterative elimination of conditionally dominated actions, as stated in the following proposition.

4.4 Iterated Conditional Dominance in Bargaining

see the lecture notes on bargaining

4.5 Extensive-Form Rationalizability

Extensive-form rationalizability, due to Pearce (1984), strengthens rationalizability by considering a stronger form of rationality: sequential rationality. The original definition involves hierarchies of conjectures. Here, I will present a simpler characterization of it due to Shimoji and Watson (1998) and Battigalli (1997).

Given any extensive-form game with a normal-form representation $G = (N, S, u)$, one can represent information sets of a player as subsets of S as follows; there is no chance moves. For any information set h, define $S(h) \subseteq S$ as the set of strategy profiles that reach information set h. By perfect-recall, if h is an information set of player i, then $S(h)$ is of the form $S_i(h) \times S_{-i}(h)$. A strategy $s_i \in S_i(h)$ is said to be *conditionally dominated at* $S(h)$ if there exists a mixed strategy σ_i such that

$$u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}) \qquad (\forall s_{-i} \in S_{-i}(h)).$$

Extensive-form rationalizability can be defined as the strategies that survive iterated elimination of conditionally dominated strategies. This procedure can be applied to reduced normal-form as well as normal-form representation.

Example 4.11 *Consider the game in Figure 4.6. The game has the following normal-form representation:*

	a_2a_4	a_2d_4	d_2a_4	d_2d_4
A_1A_3	$4,0$	$0,3$	$2,1$	$2,1$
A_1D_3	$2,1$	$2,1$	$1,2$	$1,2$
D_1A_3	$3,0$	$3,0$	$3,0$	$3,0$
D_1D_3	$3,0$	$3,0$	$3,0$	$3,0$

The first information set (of Player 1) is S; the second information set (of Player 2) is $\{A_1A_3, A_1D_3\} \times S_2$ (as Player 2 knows that Player 1 played A_1); the third information set (of Player 1) is $\{A_1A_3, A_1D_3\} \times \{a_2a_4, a_2d_4\}$ (as Player 1 knows that A_1 and a_2 are played), and the last information set (of Player 2) is $\{A_1A_3\} \times \{a_2a_4, a_2d_4\}$ as Player 2 knows that Player 1 played A_1A_3 and she played a_2. Now, strategy A_1D_3 is conditionally dominated at the first information set (entire game); hence it is eliminated in the first round. Moreover, in the last information set $\{A_1A_3\} \times \{a_2a_4, a_2d_4\}$, strategy a_2a_4 is conditionally dominated by a_2d_4 and eliminated. These are the only eliminations in the *first round. The remaining game is as follows:*

	a_2d_4	d_2a_4	d_2d_4
A_1A_3	$0,3$	$2,1$	$2,1$
D_1A_3	$3,0$	$3,0$	$3,0$
D_1D_3	$3,0$	$3,0$	$3,0$

In the second round, the second information set, which belongs to Player 2, reduces to $\{A_1A_3\} \times S_2 \backslash \{a_2a_4\}$, as the strategies A_1D_3 and a_2a_4 have been eliminated. Now, strategies d_2a_4 and d_2d_4 become conditionally dominated by a_2d_4, and hence they are eliminated. The only remaining strategy is a_2d_4. Similarly, strategy A_1A_3 has become conditionally dominated by D_1A_3 at the initial history and eliminated; both strategies D_1A_3 and D_1D_3 of Player 1 survive this round. There are no further eliminations; the surviving strategies are $\{D_1A_3, D_1D_3\}$ for Player 1 and $\{a_2d_4\}$ for Player 2.[7]

Observe that the outcome of extensive-form rationalizability is D_1, the same as the unique backward induction solution. This is generally true: the outcome of extensive-form rationalizability coincides with the unique backward induction outcome in generic perfect-information games of finite horizon. However, the solutions are different: Player 2 plays strategy a_2d_4 according to extensive-form rationalizability, while she plays d_2d_4 according to backward induction solution. Since iterative elimination of conditionally dominated actions coincides with backward induction, this also shows that the iterative elimination of conditionally dominated strategies is different from iterative elimination of conditionally dominated actions.

4.6 Forward Induction

Solutions based on backward induction analyze subgames in isolation. For example, subgame-perfect Bash equilibrium requires that the solution induced in a subgame should be a Nash equilibrium of the subgame when it is considered in isolation, but it does not put any restriction based on rationality of the induced behavior within the context of the larger game. Markov-perfect equilibrium, a widely used refinement in applications,

[7] Observe that only the initial information set is reached by the surviving strategies. One cannot maintain higher-order assumptions about sequential rationality at the remaining information sets.

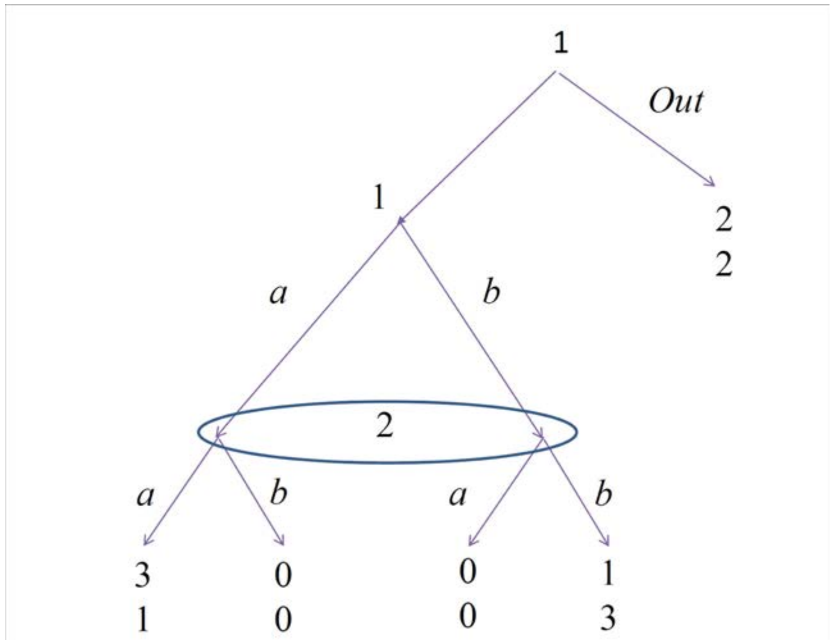

Figure 4.9: Battle of the Sexex game with outside option

goes even further and asks us to let bygones be bygones, and do not even make inference about future behavior from the past ones. In the same vein, when players reach an information set that are not supposed to be reached, perfect and sequential equilibria all attribute this to mistakes (or trembles). However, it may be more prudent to interpret any such deviation as a deliberate rational choice when it is possible. Forward induction, an informal idea proposed by Kohlberg and Mertens (1986), stipulates that one should maintain rationality of players when possible. In particular, when there is a rational interpretation of a move, one should not resort to an irrational interpretation. This imposes restriction on the possible strategies within a subgame that would not be imposed if the game were considered in isolation, leading to sharper predictions.

For a concrete example, consider the game in Figure 4.9. The proper subgame is the classical Battle of the Sexes game, in which the players want to coordinate their actions but they have opposing interest about which strategy they should coordinate on. The subgame has two pure strategy Nash equilibria, (a, a) and (b, b), as well as a mixed strategy equilibrium in which each player plays her favorite strategy with probability $3/4$, yielding expected payoff of $3/4$ to each player. Each of these equilibria lead to a subgame-perfect Nash equilibrium in the larger game, in which player 1 chooses between an outside option and playing this game. In one equilibrium, $(In, (a, a))$, she goes in expecting that they will coordinate on a in the subgame. In the other two equilibria, $(Out, (b, b))$ and (Out, mix), she stays out anticipating that they will end up coordination on b or worse playing mixed strategies. But are those expectations all reasonable?

Consider the equilibrium $(Out, (b, b))$. In this equilibrium, observing that player 1 has chosen to play the game, Player 2 thinks that she will play b in the subgame. However, this would give only 1 maximum to Player 1, and Player 1 has given up a payoff of 2 to play the game. Thus, Player 2 no longer maintain rationality assumption for Player 1, attributing her choice of In to a mistake (or tremble). He does so despite the fact that there is a perfectly rational interpretation for Player 1's move: she intends to play a, anticipating that Player 2 also plays a. This would have given her a payoff of 3, better than the outside option. Forward induction requires that one should not attribute a move to irrationality when a rational interpretation is available. In this example, it requires that Player 2 infer that Player 1 intends to play a. More generally, forward induction would require that one maintains higher order beliefs in rationality when possible. In particular, anticipating that Player 2 will interpret her going in rationally, Player 1 foresees that Player 2 will play a and plays In and then a, leading to the equilibrium $(In, (a, a))$.

Note that subgame-perfect Nash equilibrium is not alone in allowing Player 2 attribute In to a mistake in this game. In fact, $(In, (b, b))$ is a proper equilibrium, as shown in Example 4.9; the game on the left in that example is a reduced form representation of the Battle of Sexes game with outside option. Therefore, it is perfect and sequential equilibrium as well. Similarly, all equilibria survive iterated elimination of conditionally dominated strategies, as that procedure would not eliminate any subgame-perfect Nash equilibrium.

Both strategic stability of Kohlberg and Mertens (1986) and extensive-form rationalizability would choose $(In, (a, a))$ as the unique solution in our game, which is represented in reduced form as

	a	b
a	$3, 3$	$0, 0$
b	$0, 0$	$1, 1$
Out	$2, 2$	$2, 2$

In this game, $\{(a,a)\}$ is the only stable component. To see that there is no other stable component, observe that b is strictly dominated by Out, leading to the reduced game

	a	b
a	$3,3$	$0,0$
Out	$2,2$	$2,2$

In this game, a weakly dominates b for player 2, and iterative elimination of weakly dominated strategies leads to the unique solution $\{(a,a)\}$. Since some equilibria in a stable component must survive iterative elimination of weakly dominated strategies (within a stable component of the game obtained this way), this shows that there is no other stable component, which should be disjoint from $\{(a,a)\}$.

One can apply extensive-form rationalizability as follows. There are two information sets: the whole game, corresponding to the first information set of Player 1, and $\{a,b\} \times \{a,b\}$, which is common to both player (the simultaneous move). Strategy b is eliminated for Player 1 in her first information set. No other strategy is eliminated in the first round. In the second round, the game looks like the reduced game 2×2 game above, where the proper information set reduces to $\{a\} \times \{a,b\}$. In this information set, for Player 2, a strictly dominates b, and it is eliminated for player 2. Then, in the first information set of Player 1, which is now $\{a,b\} \times \{a\}$, a strictly dominates b, and thus b is eliminated. The unique solution is (a,a).

Observe that the eliminations in extensive form rationalizability mirrored our informal arguments. First, b is eliminated on the basis of Player 1's rationality. Then, b is eliminated for Player 2 because in her information set b were no longer available for Player 1, and thus he assumed that Player 1 plays a. Then, in her first information set, b were eliminated for Player 1 because a were the only remaining strategy for player 2. This is not a coincidence.

Extensive-form rationalizability is one of the prominent formalizations of the informal forward-induction argument. Battigalli and Siniscalchi (2002) characterize extensive-form rationalizability in terms of "common strong belief in sequential rationality", which requires that players are sequentially rational, they maintain believing that the players are sequentially rational until proven otherwise, they maintain believing "that the players are sequentially rational and that they maintain believing that the players are sequentially rational until proven otherwise," and so on. Another prominent formalization of

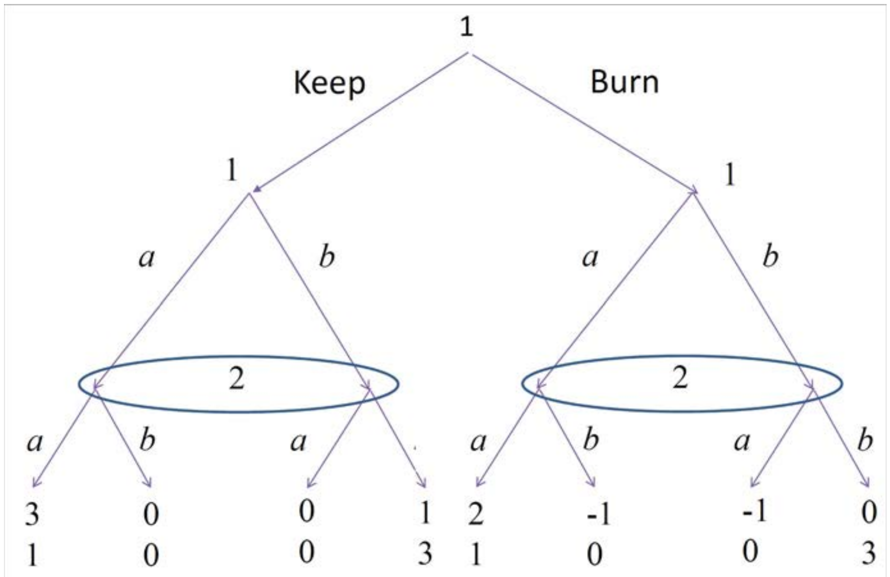

Figure 4.10: Money-Burning game

forward induction is the intuitive criteria of Cho and Kreps in signaling games.

Since extensive-form rationalizability picks the backward induction outcome (but not the solution) in generic finite games of perfect information, under this formalization, forward induction also implies backward induction outcome in those games.

Iterative application of forward induction leads to interesting (and perhaps perplexing) conclusions. The following example illustrates this.

Consider the Money-Burning game in Figure 4.10. In this game, before starting the game, Player 1 has the option of burning a dollar, reducing her payoff by one at all contingencies. In this game there are many subgame perfect Nash equilibria, obtained by selecting one of the three SPNE in each subgame. In all of them, Player 1 ends up not burning the dollar, but any of the equilibrium behavior is possible after that, in particular playing (b, b).

Now, by forward induction, if she burns a dollar, Player 2 should expect that Player 1 is going to play a. This is because burning a dollar and playing b is strictly dominated by not burning the dollar and playing b. Thus, Player 2 would play a after Player 1 burns a dollar. Therefore, by burning a dollar, Player 1 can guarantee a payoff of 2 for herself. Then, the reduced game is as in the Battle of Sexes with outside option, which gives payoff vector $(2,1)$. Then, applying backward induction further (as we did above), one concludes that they will play (a,a) if she does not burn the dollar. This leads to the unique solution: Player 1 does not burn the dollar, and they play (a,a) regardless of what she does.

Now, what if Player 1 burns the dollar anyway? The fourth-order mutual "strong belief in sequential rationality" would imply that Player 1 does not burn the dollar. Once he sees that Player 1 burns the dollar, Player 2 must conclude that Player 1 does not maintain the fourth order "strong belief in sequential rationality", and we cannot maintain the fifth order strong belief for Player 2 at that information set. What should then Player 2 believe at that information set? The extensive form-rationalizability maintains the fourth order—although players know that the higher order strong belief would contradict the information set, and any of the assumptions could have failed.

4.7 Exercises

Exercise 4.1 *Consider the extensive-form game in Figure 4.11 where $\theta \in \{0,2\}$. Assume that it is common knowledge that $\theta = 2$.*

1. *Compute the set of rationalizable strategies.*
2. *Compute the set of Nash equilibria.*
3. *Compute the sets of perfect, proper and sequential equilibria.*
4. *Compute the set of extensive-form rationalizable strategies; take this as the implications of forward induction for this game.*
5. *Briefly discuss your results.*

Exercise 4.2 *In the previous question, assume that $\theta = 2$ with probability $p \in (3/4,1)$ and Player* 1 privately observes *the realization of θ.*

1. *Compute the set of ex-ante and interim-independent rationalizable strategies.*
2. *Compute the set of (Bayesian) Nash equilibria.*
3. *Compute the sets of perfect, proper and sequential equilibria.*

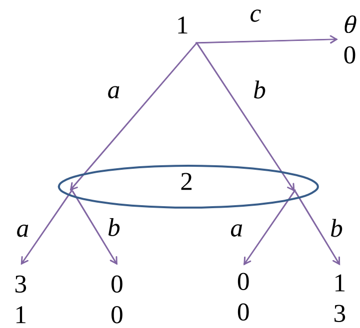

Figure 4.11:

4. *For each sequential equilibrium, check whether it satisfies forward induction by checking whether Player 2 assigns positive probability only on the sequentially rational strategies of Player 1.*

5. *Briefly discuss your results by comparing them to the results in the previous exercise.*

Exercise 4.3 *Consider a twice repeated "prisoners' dilemma" game in which the previous actions are observable and the stage-game payoff function is*

	C	D
C	$1, 1$	$1-2\theta, 1+\theta$
D	$1+\theta, 1-2\theta$	$0, 0$

where $\theta \in \{-1, 1\}$ *is privately known by Player 1 and* $\theta = -1$ *with some small probability* $\epsilon > 0$. *Compute a sequential equilibrium.*

Exercise 4.4 *Consider the Money Burning game in Figure 4.10.*

1. *Compute the subgame-perfect Nash equilibria in pure strategies.*

2. *Write the game in reduced normal form.*

3. *Compute the set of proper Nash equilibria in pure strategies.*

4. *Iteratively eliminate all weakly dominated strategies (eliminating all weakly dominated strategies at each round).*[8]

5. *Compute the set of all stable equilibrium sets.*

6. *Apply extensive-form rationalizability.*

7. *Briefly discuss your results.*

Exercise 4.5 *Consider the following variation of the Money Burning game above. We have a multistage game. First, Player 1 decides whether to burn a dollar. Then, observing Player 1, Player 2 decides whether to burn a dollar. Finally, observing above behavior, they play the Battle of the sexes game with payoffs as in the previous example, where we subtract one from all payoffs of a player if the player burned a dollar. Apply extensive-form rationalizability, and briefly discuss your finding.*

Exercise 4.6 *In the previous exercise, assume that first both players simultaneously decide whether to burn a dollar, and then they play the battle of the sexes game after observing the burning decisions. Apply extensive-form rationalizability again, and briefly discuss your results.*

Exercise 4.7 *Consider a finitely repeated Prisoners' Dilemma game, in which the following stage game is played at every period and all the previous actions are publicly observable:*

	C	D
C	$5,5$	$0,6$
D	$6,0$	$1,1$

1. *Apply the following solution concepts to this game:*

   (a) *Iterated elimination of all weakly dominated strategies;*

   (b) *Iterated elimination of conditionally dominated actions;*

   (c) *Extensive-form Rationalizability*

[8] Order of elimination matters in general in iterative weak dominance.

2. *Redo parts 1a and 1b for the following variation of the above game. At the beginning, Player 1 chooses whether to commit to playing tit-for-tat privately. If she commits, then she only has the action for tit-for-tat available at every history: at the beginning she only has C, and she has the last action played by the other player at other histories. If she does not commit the available actions are as before. Player 2 does not observe whether Player 1 commits.*

Chapter 5

Supermodular Games

A common[1] exercise in economics is to understand how a particular outcome varies with a particular parameter. For example, one may want to know whether a reduction in income tax increases the investment level in equilibrium. When one can answer such a question, this is often driven by a supermodularity (or complementarity) assumption one makes in setting up the game. In this chapter, I will formally introduce supermodular games and present the structure of the solutions and main results for comparative statics. The analyses rely on lattice theory, and I summarize the basic concepts and the tools from lattice theory in Appendix A.5; you should study the appendix before proceeding.

Complementarities are expressed both in terms of constraints and payoff functions. In terms of constraints, two activities are complementary if doing one activity more does not reduce the possible activity level for the other activity. This is mathematically captured by lattices. In terms of payoffs, two activities are complementary if doing one activity more increases the marginal benefit of doing the other. This is mathematically captured by supermodular payoff functions.

The main result will establish the structure of the solution set and monotone comparative statics under complementarity. For individual decision problems, the result establishes that the set of solutions is a lattice and weakly increasing in complementarity parameters—both in terms of constraints and the payoff function. For games, the result establishes that there are extremal equilibria that bound all rationalizable strategies, and the extremal equilibria are weakly increasing in complementarity parameters.

[1]The notes on supermodularity are partly based on lectures by Paul Milgrom.

5.1 Example

The general properties of equilibria under complementarity can be gleaned from Diamond's search model. There is a continuum of players. Each player i exerts effort $a_i \in [0,1]$ which costs him $a_i^2/2$. Let $\bar{a}_{-i}$ be the average effort level for the players other than i. The probability that i finds a match is $a_i g(\bar{a}_{-i})$ for some increasing, continuous function $g : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$. Let the payoff from match be $\theta \geq 0$. Then, the expected payoff of player i is

$$U_i(a) = \theta a_i g(\bar{a}_{-i}) - a_i^2/2.$$

The payoff function exhibits strategic complementarity (i.e. complementarity between the strategies). That is, an increase in $\overline{a}_{-i}$ always results in a (weakly) increase in the marginal utility $\partial U_i/\partial a_i$ of exerting more effort. This leads to an increasing best-response function:

$$B_i(a_{-i}) = \theta g(\bar{a}_{-i}).$$

Note that the level of strategic complementarity depends on θ and the slope of g. Similarly, there is complementarity between the search level a_i and the value θ of match:

$$\partial^2 U_i/\partial a_i \partial \theta = g(\bar{a}_{-i}) \geq 0.$$

Once again the best response is increasing in θ.

Consider the Nash equilibria of the above game. Since the best response function is increasing and the payoffs are symmetric, every equilibrium is symmetric. Equilibria are then characterized by the intersection of the graph of g with the diagonal, as in Figure 5.1. In this figure, there are three equilibria, and all of the equilibria are ordered, where the smallest equilibrium is located at the origin. Among these, the smallest and the largest equilibria are stable, while the middle equilibrium is unstable. While the number of equilibria depends on the shape of g, the equilibria will always be ordered (because g and the diagonal are increasing), and there will exist extremal equilibria. The latter is indeed a general property of supermodular games.

How do the equilibrium search levels vary by θ? To find an answer, increase θ to a higher level θ'. Since this corresponds to scaling up the best response function (by θ'/θ), the new equilibria are formed as in the figure. The smallest equilibrium remains at zero (weakly increasing). The largest equilibrium moves up. These changes are intuitive;

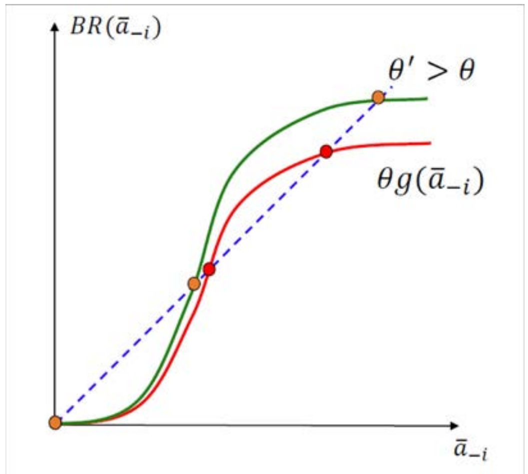

Figure 5.1: Equilibria in Diamond's search model

players search more when the match is more valuable. Note however that the middle equilibrium, which is unstable, decreases, so that the players search less when the match is more valuable. This shows that the intuition is true only for the extreme equilibria, and one should keep this counterexample in mind throughout. This will indeed be generally true for all supermodular games: extremal equilibria are weakly increasing in complementary parameters, while some interior equilibria may exhibit unintuitive comparative statics. Finally, note that the largest equilibrium moves more than individual best responses, i.e., $a_i^*[\theta'] > B_i(a_i^*[\theta], \theta')$, where $a_i^*[\theta]$ and $a_i^*[\theta']$ are the equilibrium strategies under θ and θ', respectively. That is, there is a multiplier effect.

5.2 Supermodular Optimization Problems

In this section, I will present the main result for the individual decision problems, establishing the lattice structure of the optimal solutions and establishing monotonicity of the solution to the complementary payoff parameters.

5.2.1 Monotonicity Theorem

Theorem 5.1 (Topkis's Monotonicity Theorem) *For any lattices* $(X, \geq)$ *and* $(\Pi, \geq)$, *let* $u : X \times \Pi \to \mathbb{R}$ *be a supermodular function (with coordinate-wise order) and define*

$$B(\pi) = \arg\max_{x \in D(\pi)} u(x, \pi).$$

If $\pi \geq \pi'$ *and* $D(\pi) \geq D(\pi')$, *then* $B(\pi) \geq B(\pi')$.

A couple of comments on the statement of the result are in order. First, we use the strong set order in comparing $D(\pi)$ to $D(\pi')$ and $B(\pi)$ to $B(\pi')$. Using such a strong notion to compare the domains make the result weak, but its usage in comparison of the optimal solutions makes the result strong. Second, the supermodularity condition on u here can be weakened as

$$u(x \vee x', \pi) + u(x \wedge x', \pi') \geq u(x, \pi) + u(x', \pi')$$

because we are only interested in the case of π and π'. Finally, the condition that $\pi \geq \pi'$ can always be satisfied as a matter of definition. In general, it suffices to have f supermodular with respect to x and has increasing differences.

Proof. Assuming $\pi \geq \pi'$ and $D(\pi) \geq D(\pi')$, take any $x \in B(\pi)$ and $x' \in B(\pi')$. In order to show that $B(\pi) \geq B(\pi')$, we need to show that $x \vee x' \in B(\pi)$ and $x \wedge x' \in B(\pi')$. For this, it suffices to show that

$$\begin{aligned} x \vee x' &\in D(\pi), \\ u(x \vee x', \pi) &= u(x, \pi), \\ x \wedge x' &\in D(\pi'), \\ u(x \wedge x', \pi) &= u(x', \pi'). \end{aligned}$$

Now, since $x \in B(\pi) \subseteq D(\pi)$, $x \in D(\pi)$. Similarly, $x' \in D(\pi')$. Since $D(\pi) \geq D(\pi')$, we then have $x \vee x' \in D(\pi)$ and $x \wedge x' \in D(\pi')$. To show $u(x \vee x', \pi) = u(x, \pi)$ and $u(x \wedge x', \pi') = u(x', \pi')$, note that since $x \in B(\pi)$ and $x \vee x' \in D(\pi)$,

$$u(x \vee x', \pi) \leq u(x, \pi).$$

Similarly, $u(x \wedge x', \pi') \leq u(x', \pi')$. If either of these inequalities were strict, we would have

$$u(x \vee x', \pi) + u(x \vee x', \pi') < u(x, \pi) + u(x', \pi'),$$

contradicting the supermodularity condition above. Therefore, $u(x \vee x', \pi) = u(x, \pi)$ and $u(x \wedge x', \pi') = u(x', \pi')$. ■

Note that when the domain of optimization is a lattice, the Monotonicity Theorem implies that the optimal solutions form a lattice:

Corollary 5.1 *For any fixed π, if $u(\cdot, \pi) : X \to \mathbb{R}$ is supermodular and $D(\pi)$ is a sublattice of X, then $B(\pi)$ is a sublattice of X.*

Proof. Since $D(\pi)$ is a sublattice, $D(\pi) \geq D(\pi)$. Since $\pi \geq \pi$, Monotonicity Theorem concludes that $B(\pi) \geq B(\pi)$, showing that $B(\pi)$ is a sublattice. ■

Note that Monotonicity Theorem leads to strong comparative statics without making any continuity assumption or any assumption on the domain of the parameters π. For example consider the function u on Figure 5.2, where $\Pi = \{0, 1\}$ and $u(x, 1) - u(y, 1) \geq u(x, 0) - u(y, 0)$ for any $x > y$. Let $D(0) = [0, 2]$ and $D(1) = [a, a + 2]$ for $a \geq 0$. Considering $(\pi = 1, a)$ as the new parameter, note that D is increasing in both π and a. Monotonicity Theorem concludes that B is increasing in π and a. Indeed,

$$B(0) = \{x_0\}$$

$$B(1, a) = \begin{cases} \{a + 2\} & \text{if } a + 2 \leq x_1 \\ \{x_1\} & \text{if } x_1 \leq a + 2 < x_2 \\ \{x_1, x_2\} & \text{if } a + 2 = x_2 \\ \{a + 2\} & \text{otherwise.} \end{cases}$$

Since $a \geq 0$, $B(1, a) \geq B(0)$. This is despite the fact that the solution is discontinuous and u does not satisfy the usual concavity conditions. This example also shows that the assumption that $D(\pi) \geq D(\pi')$ is not superfluous. If $a < x_0 - 2$, so that $D(1, a) \not\geq D(0)$, then $B(1, a) = \{a + 2\} \not\geq \{x_0\} = B(0)$.

5.2.2 Applications

I will illustrate the applications of Monotonicity Theorem on a couple traditional examples next.

Example 5.1 (Pricing) *Consider a monopolist who chooses a price p for its product, facing a demand function $D(p, \theta)$ and marginal cost c, where θ is a demand parameter.*

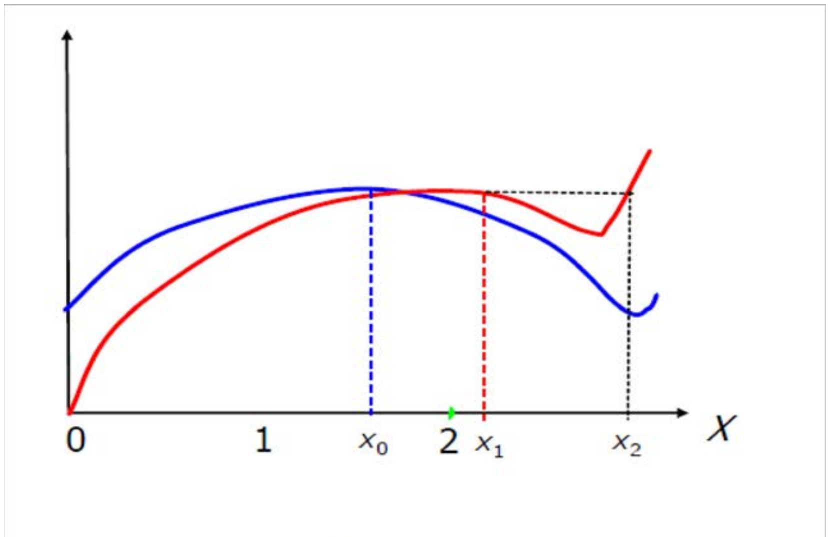

Figure 2:

Figure 5.2:

Write

$$p^*(\theta, c) = \arg\max_{p \geq c'} (p - c) \, D(p, \theta)$$

for the optimal price, where $c' > c$ *is a fixed lower bound for prices. Direct application of Monotonicity Theorem to this problem may not be as useful. Observe however that optimal solution is invariant to monotone transformations of objective functions, and hence*

$$p^*(\theta, c) = \arg\max_{p \geq c'} \log(p - c) + \log D(p, \theta) \, .$$

The new objective function is supermodular with respect to p *and* c*. Hence, Monotonicity Theorem concludes that* p^* *is weakly increasing in* c*. Moreover, the new objective function is supermodular with respect to* p *and* θ *as long as* $\log D(p, \theta)$ *is supermodular (i.e.* $D(p, \theta)$ *is log-supermodular). Hence, Monotonicity Theorem concludes that* p^* *is weakly increasing in* θ *as long as the demand function is log-supermodular.*

Note also that the order on the domain is invariant to monotone transformations on the domain. Hence, the latter condition is equivalent to $\log D(p, \theta)$ being supermodular with respect to $(\log p, \theta)$, i.e. the price elasticity of demand

$$-\frac{\partial \log D(p, \theta)}{\partial \log p}$$

being weakly decreasing in θ.

Example 5.2 (Pricing under Demand Uncertainty) *In the pricing example take $c = 0$ and assume that the monopolist does not know θ and has belief π about θ. Assume also D is weakly increasing in θ and supermodular. Write $\tilde{D}(p, \pi)$ for the expected value of $D(p, \theta)$ under belief π and write*

$$p^*(\pi) = \arg\max_{p \geq 0} p\tilde{D}(p, \pi).$$

Monotonicity Theorem implies that $p^*(\pi)$ is isotone in belief π, in that $p^*(\pi) \geq p^*(\pi')$ whenever π first-order stochastically dominates π', denoted as $\pi \geq_{FOSD} \pi'$. Since the set of probability distributions is a lattice under $\geq_{FOSD}$ (see Exercise A.6), it suffices to show that $p\tilde{D}(p, \pi)$ is supermodular to prove this. But, $p\tilde{D}(p, \pi)$ is supermodular whenever p and $\tilde{D}(p, \pi)$ are isotone in π, non-negative and supermodular (see Exercise A.2). These conditions are satisfied by p trivially and are inherited from D by $\tilde{D}$ (you should prove them).

Example 5.3 (Auction Theory) *Consider a bidder in an auction for an object. The value of winning the object at price/bid p is $U(p, t)$ where t is the type of the player. Suppose that we are only interested in whether the bidder's bid is increasing in his type (which ensures optimality of the auctions under certain conditions). When can we ensure that the bid is indeed increasing in t without computing the solution to the entire auction problem, which is often very difficult? Write $F(p)$ for the probability of winning when he bids p. The optimal bid is*

$$\begin{aligned} p^*(t) &= \arg\max U(p, t)F(p) \\ &= \arg\max \log U(p, t) + \log F(p). \end{aligned}$$

By Monotonicity Theorem, p^* is weakly increasing in t as long as U is log-supermodular.

Exercise 5.1 *Note that the above analysis assumes that the probability of winning is independent of type given the bid, which makes sense only if the types are independent. How would the answer change if F depends on both p and t?*

Example 5.4 (Production) *Suppose that the profit of a firm is*

$$pf(k, l) - L(l, w) - K\ (k, r),$$

where p is the price of the product, k is the capital input, l is the labor input, and w and r are cost parameters for labor and capital, such as wage and interest, respectively. The firm chooses k and l. How would an increase in w affect the optimal labor and capital level? Note that it is natural to assume that L is supermodular (e.g. $L = lw$). This is equivalent to assuming that $-L$ is supermodular with respect to $-l$. It turns out that this suffices to conclude that optimal l is weakly decreasing in w. Note that in order to apply Monotonicity Theorem, we need to ensure that the profit function is supermodular in (k, l, w), i.e., we also need to assume supermodularity with respect to k. We go around this requirement as follows (which is a useful trick). Since we are only interested how l changes with w, the order on k is irrelevant to the end result. Let $l^*(w, r)$ and $l^*(w', r)$ be the solutions at the relevant values where $-l^*(w, r) \geq -l^*(w', r)$. We order k in such a way that the profit function is supermodular at these values:

$$k \geq k' \iff f(k, l^*(w, r)) - f(k, l^*(w', r)) \geq f(k', l^*(w, r)) - f(k', l^*(w', r))$$

Then, Monotonicity Theorem (with the restricted domain) implies that $-l$ is weakly increasing in w, i.e., l is weakly decreasing in w. How does optimal k change in w? In order to answer this question we need to use the original (or reverse) order on k. It is usually assumed that the production function is supermodular (e.g. $f = k^{\alpha} l^{\beta}$). In that case, the profit function is supermodular when we use the reverse order on both k and l. Then, Monotonicity Theorem concludes that the optimal capital k is decreasing in w.

5.2.3 Extensions and Generalizations

There are several generalizations of the Monotonicity Theorem. This section briefly introduces two of them: a general characterization by Milgrom and Shannon (1994) and a special application to expected utility theory by Athey (2002).

Observe that supermodularity is *not* an ordinal property. That is, a function f may be supermodular while a monotone transformation $h \circ f$ of it is not, where h is an increasing function. That is why we transformed the objective function to obtain a supermodular function in previous applications. On the other hand, $\arg\max$ and its comparative statics are ordinal properties in definition. Hence, there must be a weaker ordinal condition than supermodularity to obtain monotone comparative statics. Milgrom and Shannon obtains the weakest such conditions. Clearly, we only need a monotone transformation of the objective function to be supermodular. It turns out that the next property is equivalent to that condition:

Definition 5.1 *A function $f : X \to \mathbb{R}$ on a lattice is said to be* quasi-supermodular *if for any $x, y \in X$,*

$$\begin{aligned} f(x) &\geq f(x \wedge y) \Rightarrow f(x \vee y) \geq f(y) \\ f(x) &> f(x \wedge y) \Rightarrow f(x \vee y) > f(y). \end{aligned}$$

It turns out that this purely ordinal condition is ordinal equivalent to supermodularity in that every quasi-supermodular function becomes supermodular under some monotone transformation. The next result weakens the increasing differences condition, it is widely used in contract theory and mechanism design:

Definition 5.2 *A function $f : X \times \Pi \to \mathbb{R}$ is said to have* single crossing property in (x, π) *if for any $x > x'$ and $\pi > \pi'$*

$$\begin{aligned} f(x, \pi') &\geq f(x', \pi') \Rightarrow f(x, \pi) \geq f(x', \pi) \\ f(x, \pi') &> f(x', \pi') \Rightarrow f(x, \pi) > f(x', \pi). \end{aligned}$$

That is, $f(x, \pi) - f(x', \pi)$ as a function of parameter π crosses the zero at most once and only from below.

Theorem 5.2 (Milgrom and Shannon) *Let $f : X \times \Pi \to \mathbb{R}$, where X is a lattice and Π is a partially ordered set. Then, for all $(\pi, D), (\pi', D') \in \Pi \times 2^X$,*

$$(\pi, D) \geq (\pi', D') \Rightarrow \arg\max_{x \in D} f(x, \pi) \geq \arg\max_{x \in D'} f(x, \pi')$$

if and only if f is quasisupermodular in x and satisfies the single crossing property in (x, π).

The theorem weakens the sufficient conditions supermodularity and increasing differences to their essential ordinal elements for monotonicity; quasisupermodularity and single-crossing property. The theorem also shows that these elements are also necessary to obtain monotone comparative statics for all domains of optimization. One may not need such comparative statics for all domains. For example in game theoretical applications, we often fix the domain as the set of all strategies, X. Quah and Strulovici (2009) obtain a weakening of single crossing property for such problems. In particular, for a subset X of real line, they show that $\arg\max_{x\in D} f(x,\pi)$ is isotone in π for a fixed interval D if f satisfies interval dominance order: for every $\pi > \pi'$, there is an increasing and strictly positive function $\alpha : X \to \mathbb{R}$ such that $f(x,\pi) - f(x',\pi) \geq \alpha(x')(f(x,\pi') - f(x',\pi'))$ for all $x > x'$.

A particular monotone transformation plays a special role in applications (as in the previous sections): the logarithm. The logarithm converts multiplicatively separable functions, such as profit functions, to additively separable functions, for which supermodularity is easy to hold as it is preserved under summation. A function f is said to be log-supermodular if $\log f$ is supermodular. In other words:

$$f(x \vee y) f(x \wedge y) \geq f(x) f(y).$$

Clearly, log-supermodularity is another sufficient condition monotonicity.

In game theoretical applications we typically assume that the decision maker is expected utility maximizer, and players often face uncertainty, especially in games of incomplete information. Athey (2002) studies monotonicity in such optimization problems using log-supermodularity. Log-supermodularity is *not* preserved under addition in general, and hence one would expect it to be preserved under expectation. Nonetheless, using existing results in probability theory, she shows that log-supermodularity is preserved under integration if f satisfies log-supermodularity with respect to all parameters. That is, if $g : X \times \Pi \times \Theta \to \mathbb{R}$ is log-supermodular, then the objective function U, defined by

$$U(x,\pi) = \int g(x,\pi,\theta)\, d\theta$$

is also supermodular. Of course, in expected utility applications g is the product of the utility function and density; and log-supermodularity is preserved under multiplication. Athey (2002) then obtains the following monotonicity theorem for which she also proves a converse:

Theorem 5.3 (Athey) *Consider an expected utility maximizer with utility function* $u : X \times \Pi \times \Theta \to \mathbb{R}$ *and density* $f : \Theta \times \Pi \to \mathbb{R}$*. If both* u *and* f *are log-supermodular, then*

$$B(\pi) = \arg\max_{x\in X} \int u(x,\pi,\theta) f(\theta,\pi)\, d\theta$$

is isotone.

When θ is real-valued and f is independent of π, the function f is trivially log-supermodular. In that case, it suffices to have a log-supermodular utility function. One can also take the utility function independent of π to obtain a comparative static result about beliefs, based on log-supermodularity of f. (Milgrom and Weber show that f is log-supermodular if and only if θ and π are affiliated.)

5.2.4 Monotonicity Theorem with Continuity and Completeness

In application to supermodular games, we will assume that the strategy spaces are complete lattices and the utility functions are continuous with respect to the order topology. In that case the optimal solutions have further properties; this result is due to Milgrom and Roberts (1990).

Theorem 5.4 *Let $(X, \geq)$ be a lattice, and $u : X \to \mathbb{R}$ be supermodular and continuous in the order topology. Then, for any complete sublattice D,*

$$B = \arg\max_{x \in D} u(x)$$

is a complete sublattice, and $\bar{B} \equiv \max B \in B$ and $\underline{B} \equiv \min B \in B$ exist.

Proof. Milgrom and Roberts show that B is non-empty. The fact that B is a lattice follows from Topkis's Monotonicity Theorem as we have seen above. To show that B is complete, one shows that B contains $\sup(A)$ and $\inf(A)$ for every subset A of B. Towards this end, take any $A \subseteq B$. Since D is a complete lattice, $\sup(A) \in D$ exists. Define $\hat{B} = B \cap \{x | \sup(A) \geq x\}$, which is a sublattice. We will show that $\sup(A)$ is the largest element of $\hat{B}$. Since $\hat{B} \subseteq B$, this shows that $\sup(A) \in B$. We will use Zorn's Lemma, which states that, if every chain in a set has a maximal element, then the set also has a maximal element. Take any weakly increasing chain $C \subseteq \hat{B}$, which has $\sup(C) \in D$ by completeness of D. By continuity of u under order topology, $\sup(C)$ is a maximal element in $\hat{B}$. Indeed,

$$u(\sup(C)) = \lim_{x \in C} u(x) = \max_{x \in D} u(x),$$

where the first equality is by definition of continuity for u, and the second equality is by the fact that $C \subseteq B$ and hence $u(x) = \max_{x \in D} u(x)$ for every $x \in C$. This shows that $\sup(C) \in B$. Moreover, since $C \subseteq \hat{B}$, $\sup(A) \geq \sup(C)$, showing that $\sup(C) \in \hat{B}$. Thus, by Zorn's Lemma, $\hat{B}$ has a maximal element $\hat{x} \in \hat{B}$. But $\hat{x} = \sup(A)$ by construction: $\sup(A) \geq \hat{x}$ since $\hat{x} \in \hat{B}$, and $\hat{x} \geq \sup(A)$ because $\hat{x}$ is an upper bound for $A \subseteq \hat{B}$. Therefore, $\sup(A)$ is the largest element of $\hat{B}$. ■

5.3 Supermodular Games

We will now consider games in which the strategy spaces are complete lattices and the utility functions are continuous (with respect to the order topology) and supermodular (we will make a slightly weaker assumption). Such games are called supermodular. For these games we will establish a useful structure of Nash equilibria and rationalizability, showing that the rationalizable strategies are bounded by extremal equilibria, and obtain a useful monotonicity result on extremal equilibria. We will conclude by introducing incomplete information to the analysis.

5.3.1 Formulation

Definition 5.3 *A game (N, S, u) is supermodular if for each player $i \in N$,*

- *strategy space $(S_i, \geq_i)$ is a complete lattice for some order $\geq_i$, and*
- *u_i is continuous, supermodular in s_i and has increasing differences:*

$$u_i(s \vee s) + u_i(s \wedge s') \geq u_i(s) + u_i(s') \qquad (\forall s_i, s_i' \in S_i, \forall s_{-i} \geq s_{-i}^{'} \in S_{-i}). \quad (5.1)$$

Since S is a complete lattice, $\underline{s} = \min S$ and $\overline{s} = \max S$ exist.

Here, u_i is continuous with respect to the order topology under the coordinate-wise order. That is, for every weakly increasing sequence of strategy profiles $s(n)$, $\lim_n u_i(s(n)) = u_i(\sup_n s(n))$, and for every weakly decreasing sequence of strategy profiles $s(n)$, $\lim_n u_i(s(n)) = u_i(\inf_n s(n))$.

Condition 5.1 in the above definition is weaker than full supermodularity because it only considers ordered strategy profiles s_{-i} and s'_{-i} for the other players. When $s_{-i} = s'_{-i}$, the condition reduces to the condition that u_i is supermodular in s_i:

$$u_i\ (s_i \vee s_i', s_{-i}) + u_i\ (s_i \wedge s_i', s_{-i}) \geq u_i\ (s_i, s_{-i}) + u_i\ (s_i', s_{-i}) \qquad (\forall s_i, s_i' \in S_i, \forall s_{-i} \in S_{-i}).$$

When s_i and s'_i are ordered, say $s'_i \geq s_i$, the above condition reduces to the usual increasing differences condition:

$$u_i(s'_i, s_{-i}) - u_i(s_i, s_{-i}) \geq u_i(s'_i, s'_{-i}) - u_i(s_i, s'_{-i}) \qquad (\forall s'_i \geq s_i \in S_i, \forall s_{-i} \geq s'_{-i} \in S_{-i}).$$

These are the only restrictions imposed by the definition. Recall that in product lattices as in here, supermodularity is equivalent to supermodularity with respect to each s_j and increasing differences. Here, we assume supermodularity with respect to s_i and increasing differences, but we do not make any supermodularity assumption with respect to other players' strategies s_j (with $j \neq i$). This is the only weakening of supermodularity.

Example 5.5 (Linear Differentiated Bertrand Oligopoly) *Consider the following price-competition model. There are n players. Each player i faces constant marginal cost c_i and demand function*

$$Q_i(p) = A - a_i p_i + \sum_{j \neq i} b_j p_j,$$

where A, a_i and b_j are all positive numbers. For each i, assume that price p_i is selected from $[c_i, \bar{p}_i]$ for some large $\bar{p}_i$. This yields a supermodular game in the natural order because

$$u_i(p) = (p_i - c_i)\ Q_i(p)$$

is supermodular:

$$\frac{\partial^2 u_1}{\partial p_i \partial p_j} = b_j \geq 0.$$

Example 5.6 (Linear Cournot Duopoly) *Consider a Cournot duopoly model with inverse demand function $P = A - q_1 - q_2$ and cost functions $C_1(q_1)$ and $C_2(q_2)$. Restrict the set of possible production levels to a large compact interval. This leads to a "submodular" game in the natural order because the utility function of firm i is*

$$u_i(q) = q_i P\ (q) - C_i(q_i),$$

yielding

$$\frac{\partial^2 u_1}{\partial q, \partial q_j} = -1 < 0.$$

This is a supermodular game when q_2 is ordered in the reverse order:

$$\frac{\partial^2 u_1}{\partial q_1 \partial(-q_2)} = 1 > 0.$$

In general, submodular two-player games can be made supermodular by reversing the order on one of the strategies. Hence, submodular two-player games exhibit the useful properties of supermodular games. This trick does not work, however, when there are more than two players, and the submodular games with more than two players may exhibit dramatically different properties than the supermodular ones.

Example 5.7 (Linear Cournot Oligopoly) *In the above example, suppose that there are three or more players. Once again, for any $i \neq j$,*

$$\frac{\partial^2 u_1}{\partial q_1 \partial q_j} = -1 < 0.$$

But this game cannot be made supermodular by reversing the orders. Indeed, the relation between rationalizability and Nash equilibria in Cournot oligopoly is quite different than the relation in Cournot duopoly, as we will see later.

5.3.2 Rationalizability and Equilibrium

In this section, we will establish that (i) there exist extremal equilibria and that (ii) all rationalizable strategies are bounded by the extremal equilibria. This is the main result of Milgrom and Roberts (1990). We start with summarizing the useful implications of the monotonicity results in previous section and introduce a couple useful notation.

Lemma 5.1 *For any supermodular game, any $i \in N$,*

1. *for every $s_{-i} \in S_{-i}$,*
$$B_i(s_{-i}) = \arg\max_{s_i \in S_i} u_i(s_{i,} s_{-i})$$
is a complete lattice;

2. *for every s, $\bar{B}_i(s) \equiv \max B_i(s_{-i}) \in B_i(s_{-i})$ and $\underline{B}_i(s) \equiv \min B_i(s_{-i}) \in B_i(s_{-i})$, and*

3. *$\bar{B}_i$ and $\underline{B}_i$ are isotone, i.e., $\bar{B}_i(s) \geq \bar{B}_i(s')$ and $\underline{B}_i(s) \geq \underline{B}_i(s')$ whenever $s \geq s'$.*

Proof. The first two statements are by Theorem 5.4, the monotonicity theorem for complete lattices. But since u_i is supermodular with increasing differences and the domain of optimization is independent of s_{-i}, by Topkis's Monotonicity Theorem, whenever $s \geq s'$,

$B_i(s_{-i}) \geq B_i(s'_{-i})$ in the sense of strong set order. In particular, $\bar{B}_i(s) = \max B_i(s_{-i}) \geq \max B_i(s'_{-i}) = \bar{B}_i(s')$ and $\underline{B}_i(s) = \min B_i(s_{-i}) \geq \min B_i(s'_{-i}) = \underline{B}_i(s)$. ■

The following lemma will be the main step in establishing the extremal rationalizable strategies and equilibria.

Lemma 5.2 *Every s_i with $s_i \ngeq \underline{B}_i(\underline{s})$ is strictly dominated by $s_i \vee \underline{B}_i(\underline{s})$, where $\underline{s} = \min S$.*

Proof. Take any s_i and any s_{-i}. We want to show that

$$u_i\ (s_i \vee \underline{B}_i(\underline{s}), s_{-i}) - u_i(s_i, s_{-i}) > 0. \tag{5.2}$$

Now, since $s_{-i} \geq \underline{s}_{-i}$ and $s_i \vee \underline{B}_i(\underline{s}) \geq s_i$, we have

$$\begin{aligned} u_i(s_i \vee \underline{B}_i(\underline{s}), s_{-i}) - u_i(s_i, s_{-i}) &\geq u_i(s_i \vee \underline{B}_i(\underline{s}), \underline{s}_{-i}) - u_i(s_i, \underline{s}_{-i}) \\ &\geq u_i(\underline{B}_i(\underline{s}), \underline{s}_{-i}) - u_i(s_i \wedge \underline{B}_i(\underline{s}), \underline{s}_{-i}) \end{aligned} \tag{5.3}$$

where the first inequality is by increasing differences and the second inequality supermodularity in own strategy s_i. Hence, to show (5.2), it suffices to show that

$$u_i(\underline{B}_i(\underline{s}), \underline{s}_{-i}) - u_i(s_i \wedge \underline{B}_i(\underline{s}), \underline{s}_{-i}) > 0.$$

But, since $\underline{B}_i(\underline{s}) \in \arg\max_{s_i} u_i(s_i, \underline{s}_{-i})$, $u_i(\underline{B}_i(\underline{s}), \underline{s}_{-i}) \geq u_i(s_i \wedge \underline{B}_i(s), \underline{s}_{-i})$. If it were true that $u_i(\underline{B}_i(\underline{s}), \underline{s}_{-i}) = u_i(s_i \wedge \underline{B}_i\ (\underline{s}), \underline{s}_{-i})$, then we would have $s_i \wedge \underline{B}_i(\underline{s}) \in \arg\max_{s_i} u_i(s_i, \underline{s}_{-i})$, and by definition of $\underline{B}_i(\underline{s})$ we would have $s_i \wedge \underline{B}_i(s) \geq \underline{B}_i(\underline{s})$, showing that $s_i \geq \underline{B}_i(\underline{s})$, contradicting the hypothesis that $s_i \ngeq \underline{B}_i(\underline{s})$. Therefore, $u_i(\underline{B}_i(\underline{s}), \underline{s}_{-i}) > u_i(s_i \wedge \underline{B}_i(\underline{s}), \underline{s}_{-i})$.
■

Iterative application of this lemma leads to the following well-known result, due to Milgrom and Roberts.

Theorem 5.5 *For any supermodular game,*

1. *$\bar{z} \equiv \lim_k \bar{B}^k(\overline{s}) \equiv \inf_k \bar{B}^k(s)$ and $\underline{z} \equiv \lim_k \underline{B}^k(\underline{s}) \equiv \sup_k \underline{B}^k(\underline{s})$ exist, where $\overline{s} = \sup S$ and $\underline{s} = \inf S$;*

2. *for every rationalizable strategy profile s.*

$$\overline{z} \geq s \geq \underline{z},$$

3. and $\overline{z}$ and $\underline{z}$ are (pure strategy) Nash equilibria.

Proof. (Part 1) First note that $\bar{B}^k(\bar{s})$ is weakly decreasing in k.[2] Hence, $\lim_k \bar{B}^k(\bar{s}) = \inf_k \bar{B}^k(\bar{s})$ (existence is by completeness, as seen before). Similarly, $\underline{B}^k(\underline{s})$ is weakly increasing, and therefore $\lim_k \underline{B}^k(\underline{s}) \equiv \sup_k \underline{B}^k(\underline{s})$.

(Part 2) I will show that if $s_i \in S_i^k$, then $s_i \geq \underline{B}_i^k(\underline{s})$. This is true for $k = 0$, by definition. Suppose that $s_j \geq \underline{B}_j^{k-1}(s)$ for all $j \in N$ and for all $s_j \in S_j^k$. Then, by Lemma 5.2, every $s_i \ngeq \underline{B}_i^k(\underline{s})$ is strictly dominated given S_{-i}^{k-1} and is not in S_i^k. Therefore, $s_i \geq \underline{B}_i^k(\underline{s})$ for every $s_i \in S_{i.}^k$.

(Part 3) I will show that $\overline{z}$ is a Nash equilibrium, i.e., $\overline{z}_i \in B_i(\overline{z}_{-1})$. To this end, take any s_i, and consider the weakly decreasing sequences $(s_i, \bar{B}_{-i}^{k-1}(\bar{s}))$ and $(\bar{B}_i^k(\bar{s}), \bar{B}_{-i}^{k-1}(\bar{s}))$. Clearly, $\left(s_i, \bar{B}_{-i}^{k-1}(\bar{s})\right) \to (s_i, \overline{z}_{-i})$ and $(\bar{B}_i^k(\bar{s}), \bar{B}_{-i}^{k-1}(\bar{s})) \to (\overline{z}_i, \overline{z}_{-i})$. Moreover, since $\bar{B}_i^k(\bar{s}) \in B_i(\bar{B}_{-i}^{k-1}(\bar{s}))$, $u_i(\bar{B}_i^k(\bar{s}), \bar{B}_{-i}^{k-1}(\bar{s})) \geq u_i(s_i, \bar{B}_{-i}^{k-1}(\bar{s}))$ for each k. Hence, by continuity of u_i in the order topology,

$$\begin{aligned} u_i(\overline{z}_i, \overline{z}_{-i}) &= u_i(\lim(\bar{B}_i^k(\bar{s}), \bar{B}_{-i}^{k-1}(\bar{s}))) = \lim u_i(\bar{B}_i^k(\bar{s}), \bar{B}_{-i}^{k-1}(\bar{s})) \\ &\geq \lim u_i(s_i, \bar{B}_{-i}^{k-1}(\bar{s})) = u_i(\lim \left(s_i, \bar{B}_{-i}^{k-1}(\bar{s})\right)) \\ &= u_i(s_i, \bar{z}_{-i}). \end{aligned}$$

■

This result establishes several important facts. First of all, it establishes that there exists an equilibrium, indeed, extremal equilibria in pure strategies (Part 3). Second, it establishes a useful procedure to compute these equilibria (Part 1): one iteratively applies extremal best response functions to the largest and smallest strategy profiles. In comparison, finding a fixed point of a function is a computationally hard problem. Finally, it establishes that the rationalizable strategies are bounded by these extremal equilibria (Part 2). This not only relates extreme implications of equilibrium and rationalizability to each other, but also helps in identifying rationalizable strategies. For example, when the extremal best response functions are continuous and strategy sets are convex intervals, the result implies that the rationalizable set is the convex hull of extremal equilibrium strategies. It also implies that uniqueness of Nash equilibrium in pure strategies is equivalent to dominance solvability:

[2] Indeed, $\bar{B}^1(\bar{s}) \leq \bar{B}^0\ (\bar{s}) = \bar{s}$ by definition of $\bar{s}$. If $\bar{B}^k\ (\bar{s}) \leq \bar{B}^{k-1}(\bar{s})$, then by monotonicity of $\bar{B}$ (Lemma 5.1), $\bar{B}^{k+1}(\bar{s}) = \bar{B}(\bar{B}^k(\bar{s})) \leq \bar{B}(\bar{B}^{k-1}(\bar{s})) = \bar{B}^k(\bar{s})$.

Corollary 5.2 *A supermodular game is dominance solvable if and only if there exists a unique Nash equilibrium in pure strategies.*

The following example illustrates the Milgrom-Roberts theorem above and shows that completeness is not superfluous.

Example 5.8 (Partnership Game) *There is an employer, who provides capital K, and a worker, who provides labor L. They share the output, which is $K^\alpha L^\beta$ for some $\alpha, \beta \in (0,1)$ with $\alpha + \beta < 1$. The utility functions of the Employer and the Worker are $K^\alpha L^\beta/2 - K$ and $K^\alpha L^\beta/2 - L$, respectively. The best-response functions K and L are plotted in Figure 5.3. There are two pure strategy equilibria, one at $(0,0)$ and one with positive labor and capital, denoted by $(\hat{K}, \hat{L})$. When all nonnegative inputs are allowed, the strategy sets are not complete lattices. In that case, every strategy is a best response to some other, and hence every strategy is rationalizable, and the bounds of Milgrom and Roberts are not valid. Now suppose that the strategy sets are bounded by above for some large $\bar{K}$ and $\bar{L}$, so that $K \in [0, \bar{K}]$ and $L \in [0, \bar{L}]$. Now, we have a supermodular game (with complete lattices as strategy spaces). Then, as shown in the figure, one can iteratively eliminate all $K > \hat{K}$ and $L > \hat{L}$. Hence $S^\infty \subseteq \left[0, \hat{K}\right] \times \left[0, \hat{L}\right]$, as in the Milgrom-Roberts theorem. Moreover, since the best response functions are continuous, $\left[0, \hat{K}\right] \times \left[0, \hat{L}\right]$ is closed under rational behavior, and hence $S^\infty = \left[0, \hat{K}\right] \times \left[0, \hat{L}\right]$.*

5.3.3 Comparative Statics

The next result, due to Milgrom and Roberts, shows that the extremal equilibria are weakly increasing in complementary parameters, extending the Monotonicity Theorem for optimization in games.

Theorem 5.6 *Consider a family of supermodular games with payoffs parameterized by t. Suppose that for all $i, s_{-i}, U_i(s_i, s_{-i}; t)$ is supermodular in (s_i, t). Write $\overline{z}(t)$ and $\overline{z}(t)$ for the extremal equilibria at t. Then, $\overline{z}(t)$ and $\underline{z}(t)$ are isotone.*

Proof. Take any t, t' with $t \geq t'$, and write $\underline{B}_t$ and $\underline{B}_{t'}$ for the minimal best response function under t and t', respectively. By Topkis's Monotonicity Theorem, $\underline{B}_t(s) \geq \underline{B}_{t'}(s)$

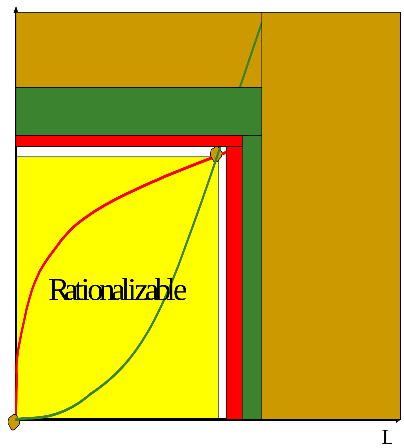

Figure 5.3: Rationalizability and Equilibria in the partnership game

for every s. Since $\underline{B}_t$ and $\underline{B}_{t'}$ are isotone, this further implies that $\underline{B}_t^k(s) \geq \underline{B}_{t'}^k(s)$ for every k. Therefore,

$$\underline{z}(t) = \sup \underline{B}_t^k(\underline{s}) \geq \sup \underline{B}_{t'}^k(\underline{s}) = \underline{z}(t')$$

Similarly,

$$\overline{z}(t) = \inf \bar{B}_t^k(\overline{s}) \geq \inf \bar{B}_{t'}^k(\overline{s}) = \overline{z}(t').$$

■

Example 5.9 *As an illustration of the theorem, take the output function in the partnership game as $tK^\alpha L^\beta$. As it is illustrated in Figure 5.4, an increase in t results in steeper best response functions. This leads the largest equilibrium to increase. On the other hand, the smallest equilibrium remains unchanged (corresponding to a weak increase).*

Monotonicity Theorem established that in single-person decision problems, the entire set of the solutions increase in the sense of strong set order. It is tempting to conjecture that the same is true for Nash equilibria in multi-person decision problems (as in the partnership game above). This is not true in general. Indeed, in Diamond's search model at the beginning, although the extremal equilibria weakly increase, the middle equilibrium actually decreases, as shown in Figure 5.1. Note that S^∞ weakly increases in the sense of set order all of these examples. Indeed, when the best response functions are continuous and strategy spaces are one-dimensional, $S_i^\infty = [\underline{z}_i, \overline{z}_i]$, and the monotonicity

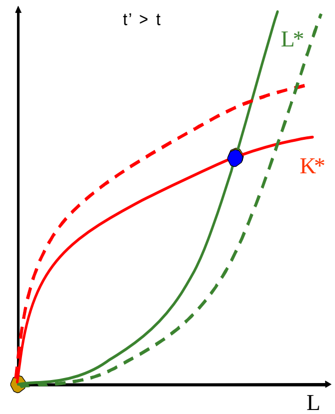

Figure 5.4: Effect of productivity parameter in the partnership game.

result of Milgrom and Roberts already implies that S^{∞} is isotone in the sense of strong set order. It is not clear how general this fact is.

5.3.4 Supermodular Bayesian Games

To some extend, the analyses for complete information games above contains Bayesian games with countable type spaces because any such Bayesian game can be represented by the interim game as a game of complete information. After illustrating this fact, I will introduce monotone supermodular games of incomplete information (due to Vives and van Zandt), in which the type and action spaces are compact subsets of $\mathbb{R}^n$, and the players' beliefs are monotone with respect to their types. These games will then exhibit further monotonicity properties because of belief monotonicity. This analysis will be used later in global games, which are special cases of these games. I will start with illustrating how one can use the above results for Bayesian games with countable (or finite) type spaces.

Definition 5.4 *A countable supermodular Bayesian game is a Bayesian game* $\mathcal{B} = (N, A, \Theta, T, u, p)$ *where*

- *each* A_i *is a complete lattice for some* $\geq_i$*,*
- T *is countable (or finite), and*

- *u_i is measurable, bounded, continuous in a, supermodular in a_i and has increasing differences.*

For any such Bayesian game, one can define the interim game $AG(\mathcal{B})$, by taking $\cup_i T_i$ as the countable set of players, A_i as the action space for each t_i, and

$$u_{t_i}(s) = E[u_i(\theta, s_i(t_i), s_{-i}(t_{-i}))|t_i],$$

where s is taken as a profile of actions (for all types), rather than strategies in the ex-ante sense. Since u_i is supermodular in a_i, $u_{t_i}(s)$ is supermodular in $s_i(t_i)$. Since u_i has increasing differences in a_{-i}, u_{t_i} has increasing differences with respect to all actions other than $s_i(t_i)$ (type t_i puts zero probability on other types of i). Moreover, since u_i is bounded, continuity of u_i implies continuity of u_{t_i}. Since A_i is already a complete lattice, this shows that the interim game is supermodular.

Lemma 5.3 *For any countable supermodular Bayesian game $\mathcal{B}$, the interim game $AG(\mathcal{B})$ is a supermodular game (of complete information).*

Using this observation, one extends the previous results to Bayesian games as follows.

Theorem 5.7 *For any countable supermodular Bayesian game $\mathcal{B}$, the following are true.*

1. *There exist Bayesian Nash equilibria s^* and s^{**} in pure strategies.*
2. *For any interim independent rationalizable action a_i of any type t_i, $s_i^*(t_i) \geq a_i \geq s_i^{**}(t_i)$.*
3. *For any Bayesian Nash equilibrium s, $s^*(t) \geq s(t) \geq s^{**}(t)$ for all $t \in T$.*

Moreover, for any family of countable supermodular Bayesian games $\mathcal{B}^\lambda = \left(N, A, \Theta, T, u^\lambda, p\right)$ with $u_i^\lambda(\theta, a_i, a_{-i})$ supermodular in (a_i, λ), the extremal equilibria $s^{**\lambda}$ and $s^{*\lambda}$ are isotone in λ.

Proof. By Lemma 5.3, $AG(\mathcal{B})$ is a supermodular game. (Part 1) Hence, by Theorem 5.5, $AG(\mathcal{B})$ has Nash equilibria s^{**} and s^* in pure strategies. Of course, s^{**} and s^* are Bayesian Nash equilibria of $\mathcal{B}$. (Part 2) Any interim independent rationalizable action a_i of any type t_i is a rationalizable action of t_i in AG $(\mathcal{B})$ by definition. Hence, by

Theorem 5.5, $s_i^*(t_i) \geq a_i \geq s_i^{**}(t_i)$. Part 3 follows from Part 2. For the last statement, observe that $u_{t_i}^\lambda(s) = E\left[u_i^\lambda(\theta, s_i(t_i), s_{-i}(t_{-i}))|t_i\right]$ is supermodular in $(s_i(t_i), \lambda)$. Hence, by Theorem 5.6, $s^{**\lambda}$ and $s^{*\lambda}$ are isotone in λ. ■

Unfortunately, the above transformation cannot be applied to uncountable type spaces because one needs measurability condition on strategies in order to compute the expectation. (Hence, the space of strategy profiles is not a product set in $AG(\mathcal{B})$.) For such cases, Vives and Van Zandt introduce following class of Bayesian games, which also incorporate useful monotonicity conditions on beliefs.

Definition 5.5 *A monotone supermodular game (of incomplete information) is a Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ with*

- *each A_i is a compact sublattice of $\mathbb{R}^K$;*
- *$\Theta \times T$ is a measurable subset of $\mathbb{R}^M$;*
- *u_i is such that*
  - *$u_i(a, \cdot) : \Theta \to R$ is measurable,*
  - *$u_i(\cdot, \theta) : A \to R$ is continuous, bounded by an integrable function, supermodular in a_i and has increasing differences,*
  - *u_i has increasing differences in (a_i, θ), and*
- *$p(\cdot|t_i)$ is a weakly increasing function of t_i in the sense of first-order stochastic dominance.*

This definition is more general in that type spaces can be any subspace of a $\mathbb{R}^n$, but it is more restrictive in that it restricts the action spaces to be subsets of $\mathbb{R}^n$. Clearly, the continuity and measurability assumptions on u is made in order to ensure the necessary continuity of conditional expected payoffs of types. Finally, the beliefs of types are assumed to be monotone in the sense of first-order stochastic dominance. Together with supermodularity of u, this ensures that the extremal equilibria are monotone (for the same reason behind Theorem 5.6). This leads to the following result.

Theorem 5.8 *Any monotone supermodular game has Bayesian Nash equilibria s^* and s^{**} in pure strategies such that for any for any type t_i and any interim correlated rationalizable action $s_i(t_i)$ of t_i,*

$$s_i^*(t_i) \geq s_i(t_i) \geq s_i^{**}(t_i),$$

and $s_i^*(t_i)$ and $s_i^{**}(t_i)$ are weakly increasing in t_i.

This result is silent about rationalizable strategies, but as we will see in the context of global games they are also bounded by the extremal equilibria as in the previous results. In conclusion, in supermodular games, all rationalizable strategies are bounded by extremal pure strategy equilibria, and these equilibria are weakly increasing with respect to complementary variables, leading to monotone comparative statics.

5.4 Exercises

Exercise 5.2 *Consider a supermodular game with n players.*

1. *Assume that the strategy sets are linearly ordered and there is a unique best response to each pure strategy profile of others. For $n \leq 3$, show that the set of pure Nash equilibria is linearly ordered under the coordinate-wise order on the strategy profiles (i.e., under the order $s \geq s' \iff s_i \geq_i s_i'$ for all i).*

2. *Find a supermodular game in which the set of Nash equilibria is not a lattice under the coordinate-wise order. (Hint: One can find an example with $n > 3$ in which the assumption in part (a) holds but the set of pure Nash equilibria is not a lattice.)*

Exercise 5.3 *Let X be a complete lattice and $T = \mathbb{R}$.*

1. *Let $f : X \times T \to X$ be isotone, and $\overline{x}(t)$ be the highest fixed point of $f(\cdot, t)$ for each t. Show that $\overline{x}$ is isotone. [Hint: First show that $\overline{x}(t) = \sup\{x | f(x,t) \geq x\}$.]*

2. *Let $\bar{B}_i(s,t)$ be the largest best reply to s_{-i} for each i in a supermodular game G_t with a generic strategy profile $s = (s_1, ..., s_n)$. Let also $\bar{B}(s,t) = (\bar{B}_1(s,t), ..., \bar{B}_n(s,t))$. Let $s^*(t)$ be the highest Nash equilibrium of G_t. Show that, if $t \geq t'$, then*

$$s^*(t) \geq \bar{B}(s^*(t'), t).$$

Exercise 5.4 *There is a (large) consumer of a good with integrable, non-increasing demand function* $D_t : \mathbb{R}_+ \to \mathbb{R}_+$ *where* $t \in \mathbb{R}$ *is a demand parameter in which* $D_t\ (q)$ *is increasing for each quantity* q. *Consumer faces an increasing, continuous supply function* $S_\omega : \mathbb{R}_+ \to \mathbb{R}_+$ *where* ω *is an unknown supply parameter (i.e. supply is stochastic). Consumer submits a non-increasing, continuous demand function (or bid)* $x : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $x \le D_t$, *and buys quantity* $q(x, S_\omega)$ *at price* $p(x, S_\omega)$, *where* $p(x, S_\omega) = x(q(x, S_\omega)) = S_\omega(q(x, S_\omega))$ *(i.e. market clearing price and quantity). The payoff of consumer is*

$$u(x, \omega, t) = \int_0^{q(s,S_\omega)} D_t(q)dq - p(x, S_\omega)q(x, S_\omega)$$

His expected utility is $U(x, t) = E[u(x, \omega, t)]$.

1. *Show that* U *is supermodular with respect to the order in Exercise A.3 for functions.*
2. *For any* t, *show that* $B(t)$ *is a sublattice where*

$$B(t) = \arg\max_x U(x, t).$$

3. *Show that* $B(t)$ *is isotone in* t.

Exercise 5.5 *Consider a Cournot duopoly where each firm* i *has a privately known cost function* $c_i : \mathbb{R}_+ \to \mathbb{R}_+$ *and the inverse-demand function* P *is twice differentiable with* $P'' + P' < 0$. *Assume that the set* C_i *of cost functions is finite and each* $c_i \in C_i$ *has positive probability. Putting the order in Exercise A.3 for functions, further assume* C_i *and strategies are restricted in such a way that the strategy space is a complete lattice.*

1. *Show that there exist Nash equilibria* $(\overline{q}_1, \underline{q}_2)$ *and* $(\underline{q}_1, .\overline{q}_2)$ *such that for each interim independent rationalizable strategy* q_i *of each firm* i, $\underline{q}_i \le q_i \le \overline{q_i}$.
2. *Suppose that we add a constant* $\triangle > 0$ *to the inverse demand, so that the new price is* $\tilde{P}(q) = P(q) + \triangle$ *for each* q. *Can you use Milgrom-Roberts theorem to determine how* $\underline{q}_i$ *and* $\bar{q}_i$ *change?*
3. *Suppose that Firm 1 receives a government subsidy, receiving* $s > 0$ *for each unit it sells. Show that* $\underline{q}_1$ *and* $\overline{q}_1$ *are weakly increasing in* s *and* $\underline{q}_2$ *and* $\overline{q}_2$ *are weakly decreasing in* s.

Exercise 5.6 *We have a differentiated Bertrand duopoly in which each firm sells* m *goods,* $k = 1, 2, ..., m$. *Firms 1 and 2 simultaneously set price vectors* p_1 *and* p_2 *and firm* i *gets profit*

$$U_i = \sum_{k=1}^{m} = (p_{i,k} - c_{i,k})Q_{i,k}(p_i, p_j)$$

where $c_{i,k} \in [0, 1]$ *is the constant marginal cost of producing good* k *for* i *and* $Q_{i,k}$ *is the demand for good* k *of* i*; it is continuous, decreasing in* $p_{i,k}$*, increasing in all the other variables (i.e. all* $p_{j,k'}$ *with* $(j, k')6 \neq (i, k)$*) and supermodular. Assume that it must be that* $p_{i,k} \in [c_{i,k}, 1]$ *for all* i, k.

1. *Assuming that each* $c_{i,k}$ *is common knowledge, show that this game is supermodular. (State any additional assumption needed.)*

2. *In part (a), show that there exist Nash equilibria* (p_1^*, p_2^*) *and* (p_1^{**}, p_2^{**}) *such that* $p_{i,k}^* \geq p_{i,k} \geq p_{i,k}^{**}(c_i)$ *for each rationalizable strategy* p_i *and each* (i, k, c_i).

3. *Assume that each* $c_i = (c_{i,1}, ..., c_{i,m})$ *is private information of* i*, coming from a countable subset of [0, 1]. Show that there exist Bayesian Nash equilibria* (p_1^*, p_2^*) *and* (p_1^{**}, p_2^{**}) *such that* $p_{i,k}^*(c_i) \geq p_{i,k}(c_i) \geq p_{i,k}^{**}(c_i)$ *for each rationalizable strategy* p_i *and each* (i, k, c_i).

Exercise 5.7 *Consider a differentiated price competition game with* n *firms and* M *markets. Simultaneously, each firm* i *sets price* $p_{im} \in [0, 1]$ *for each market* m*, and obtains profit*

$$U_i(p) = \sum_m (p_{im} - c_{im}) Q_{im}(p)$$

where $c_{im} \in [0, 1]$ *is a known cost parameter.*

1. *State conditions on functions* Q_{im} *under which the game is supermodular; these conditions are assumed throughout.*

2. *Find conditions on functions* Q_{im} *under which any extremal equilibrium* p^* *is isotone in the vector* c *of cost parameters* c_{im}. *Find an example in which the game is supermodular and* Q_{im} *is decreasing in* p_{im} *but an extremal equilibrium is not isotone.*

3. *Take*

$$Q_{im} = \theta_m - \alpha p_{im} + \beta p_{-im} + \gamma p_{-m}$$

where $\alpha, \beta, \gamma, \theta_m$ *are known parameters,* $p_{-im} = \sum_{j \neq i} p_{jm} / (n-1)$ *and* $p_{-m} = \sum_{j,m' \neq m} p_{jm} / (n(M-1))$*. Take also* $c_{im} \equiv 0$ *everywhere. Compute the unique equilibrium, assuming that it is in the interior. Argue that the game is dominance solvable.*

Exercise 5.8 *Consider a partnership game with two players, who invest in a public good project at each date* $t \in T = \{0, 1, 2, ...\}$ *without observing each other's previous investments. We assume that a strategy of a player* i *is any function* $x_i : T \to [0, 1]$*, where* $x_i(t)$ *is the investment level of* i *at* $t \in T$*. The payoff of a player* i *is*

$$U_i(x_1, x_2) = \sum_{t \in T} \delta^t \left[Af\left(x_1(t), x_2(t)\right) - c_i(x_i(t), t) \right]$$

where $\delta \in (0, 1)$*,* $A \in [0, 1]$ *is a productivity parameter,* $f : [0, 1]^2 \to \mathbb{R}$ *is a supermodular, increasing, and continuous production function, and* c_i *is a time dependent cost function for player* i*. Everything is common knowledge.*

1. *Show that the above game has equilibria* $\underline{x}$ *and* $\overline{x}$ *such that for each equilibrium* x *of this game,*

$$\underline{x}_i(t) \leq x_i\ (t) \leq \overline{x}_i\ (t) \qquad (\forall i, t)$$

2. *Let* X *be the set of all equilibria of this game. Construct an incomplete information model in which (i) it is common knowledge that each player is rational and (ii) a strategy profile* x *is played at some state* ω *if and only if* $x \in X$*.*

3. *Show that, if* $A \geq A'$*, then the extremal equilibria for these parameters satisfy*

$$\underline{x}_i(t; A) \geq \underline{x}_i(t; A') \text{ and } \overline{x}_i(t; A) \geq \overline{x}_i(t; A') \qquad \forall (i, t)\,.$$

4. *Consider a strategy* x_i *with* $x_i(0) > \overline{x}_i(0)$*. Can you construct an incomplete information model such that (i) each player is rational at each state and (ii)* x_i *is played by player* i *at some state?*

Exercise 5.9 *Consider a two-person partnership game. Simultaneously, each player i invests $a_i \in [0,1]$, and the payoff of player i is*

$$u_i(a_1, a_2, \theta) = \theta f(a_1) f(a_2) - c(a_i),$$

where $\theta \geq 0$ is a parameter, and f and c are strictly increasing functions with $f(0) > 0$. Assume that θ is common knowledge.

1. *Show that the game is supermodular.*

2. *Assuming that best-reply correspondence is convex-valued and continuous, compute all rationalizable strategies. How would your answer change without the continuity assumption?*

3. *Show that the minimum and the maximum rationalizable strategies as well as minimum and maximum equilibrium strategies are increasing functions of θ. Give an example, showing that set of Nash equilibria is not increasing in θ in the sense of strong set order.*

Exercise 5.10 *(This question is to illustrate how we can use the ideas in supermodular game literature for structural estimation, where computational costs are very high.) Two discount chains, Walmart and Kmart, are competing for M (geographical) markets. For each market m, decision of a chain i is binary: $D_{i,m} = 1$ if it has a store in market m, and $D_{i,m} = 0$ if it does not have a store in market m. Simultaneously, each chain decides in which markets it will have a store. The profit of chain i is*

$$\Pi_i = \sum_{m \in M} \left[D_{i,m}(\beta_i X_m + \delta_{i,j} D_{j.m}) + \delta_{i,i} \sum_{l \in N_m} D_{i,l} \right]$$

where

- *β is a chain specific constant,*
- *X_m is a market size variable,*
- *$\delta_{i,j} < 0$ is a parameter measuring the competition between the two firms,*
- *$\delta_{i,j} \geq 0$ is a constant measuring the synergy between the neighboring stores, and*

- *N_m is the set of neighboring markets of market m.*

Everything is common knowledge.

1. *Let*
$$B_i(D_j) - \arg\max_{D_i} \Pi_i(D_i, D_j)$$
Show that computing $B_i(D_j)$ by brute force requires at least 2^M utility comparisons. How large is this number if $M = 2065$ (the number of markets in the US)? How many utility comparisons that we would have to make if we want to compute the pure strategy Nash equilibria by brute force? Comment on how long it would take an econometrician to estimate $(\beta_i, \delta_{i,i}, \delta_{i,j})$ this way?

2. *Let $F_i(D_j)$ be the set of D_i that satisfy the first-order conditions in computing $B_i(D_j), i.e., D_i \in F_i\ (D_j)$ if and only if*
$$\begin{aligned} D_{i,m} &= 1 \Longrightarrow \Pi_i\ (D_i, D_j) \geq \Pi_i\ (0, D_{i,-m}, D_j), \\ D_{i,m} &= 0 \Longrightarrow \Pi_i(D_i, D_j) \geq \Pi_i(1, D_{i,-m}, D_j), \end{aligned}$$
where $D_i^{'} = (A, D_{i,-m})$ is the decision where $D_{i,m}^{'} = A$ and $D_{i,m'}^{'} = D_{i,m'}$ for all $m^{'} \neq m$.

   (a) *Show that $\bar{F}_i(D_j) = \max F_i(D_j)$ and $F_i(D_j) = \min F_i(D_j)$ exists.*

   (b) *Show that $B_i(D_j) \subseteq F_i(D_j)$, and for each $D_i \in B_i(D_j)$, $F_i(D_j) \geq D_i \geq \underline{F}_i(D_j)$.*

   (c) *Using the techniques discussed in the class, find a procedure for computing $\bar{F}_i(D_j)$ and $\underline{F}_i(D_j)$ such that each of the computation takes at most M^2 utility comparisons. How large is this number when $M = 2065$?*

3. *Say that $D = (D_1, D_2)$ is a pseudo Nash equilibrium iff $D_i \in F_i(D_j)$ for each i and j. Show that every Nash equilibrium is a pseudo Nash equilibrium. Show that there exists pseudo Nash equilibria D^1 and D^2 such that for each pseudo Nash equilibrium D,*
$$D_1^1 \geq D_1 \geq D_1^2 \text{ and } D_2^2 \geq D_2 \geq D_2^1;$$
in particular, the above bounds apply for each Nash equilibrium.

4. *Find a procedure for computing D^1 (and D^2) such that there are at most $4M^3$ utility comparisons. Briefly discuss this result in comparing with part 1.*

Exercise 5.11 *There is a good with a finite set A^* of attributes α. Simultaneously, the producer (Player 1) selects the quality level $q_\alpha \in [0,1]$ for each attribute $\alpha \in A^*$, and the consumer (Player 2) selects a set $A \subseteq A^*$ of attributes whose qualities to be checked. There is also unknown noise ε_α in the realized quality of each attribute α, where $(\varepsilon_\alpha)_{\alpha \in A^*}$ are i.i.d. with mean 0. The (reduced form) payoffs of the producer and the consumer are*

$$\begin{aligned} U_1(q,A) &= \sum_{\alpha\in A} v_\alpha q_\alpha + b\left|A^*\backslash A\right| - C_1(q,\gamma_1) \\ U_2(q,A) &= Var\left(\exp\left(\sum_{\alpha\in A}(q_\alpha+\varepsilon_\alpha)\right)\right) - C_2(A,\gamma_2), \end{aligned}$$

respectively, where (v_α), b, γ_1, and γ_2 are known variables with $v_\alpha \geq b \geq 0$ for each α, and C_1 and C_2 are continuous functions, weakly increasing in q and A, respectively. (Use set inclusion to order the sets; Var stands for variance.)

1. *Find broad conditions on the above functions and variables under which there are extremal equilibria. Verify that your conditions indeed imply the existence of such equilibria.*

2. *Find conditions under which the extremal equilibria are weakly increasing in (v_α) and weakly decreasing in γ_1 and γ_2.*

Exercise 5.12 *Consider a set N of players, a set of type profiles T, and a set of action profiles A. Let also the payoff function u_i of each player i depend only on action profile a and his own type t_i. Assume that T is a topological space and each $(A_i, \geq_i)$ is a complete lattice. Endow the set S_i of strategies $s_i : T_i \to A_i$ with the product order, i.e., $s_i \geq s_i'$ if and only if $s_i(t_i) \geq s_i'(t_i)$ for each t_i.*

1. *Check if S_i is a complete lattice, and find the join and the meet operators.*

2. *Let X_i be the set of continuous strategies under the order topology on A_i, i.e., the strategies such that $\{t_i : s_i(t_i) \geq a_i\}$ and $\{t_i : a_i \geq s_i(t_i)\}$ are always closed set. Check whether X_i is a sublattice. Check whether it is complete.*

3. *Assume that the type are independently distributed and partially ordered. Find conditions on the payoff functions under which all ex-ante rationalizable strategies and all Bayesian Nash equilibrium strategies are isotone. How would your answer change without independence?*

4. *Find conditions on the payoff functions under which the conclusions of Milgrom-Roberts theorem applies.*

Exercise 5.13 *In a Bayesian supermodular game, show that for any type t_i and any interim correlated rationalizable action a_i for t_i,*

$$s_i^*(t_i) \geq a_i \geq s_i^{**}(t_i)$$

where s^* and s^{**} are the extremal Bayesian Nash equilibria. (You can make any technical assumption, such as finiteness and continuity, you want.)

Exercise 5.14 *Consider the following coordination problem over a network (such as adopting a new technology). The set of players is $N = \{1, \ldots, n\}$. Simultaneously, each player i chooses an action $a_i \in \{0, 1\}$. The payoff from action 0 is normalized to zero. The payoff from action 1 is*

$$u_i = v_i + x_i + \sum_j \lambda_{ij} a_j$$

where

- *v_i is a known parameter;*
- *x_i is privately known (the type of player i in the Bayesian game) with*

  $$x_i = \theta + \varepsilon_i$$

  for independently and uniformly distributed random variables $\theta \in [-L, L]$ and $\varepsilon_1, \ldots, \varepsilon_n \in [-\epsilon, \epsilon]$, where L and ϵ are positive numbers;
- *$\lambda_{ij} \in [0, 1]$, $\lambda_{ij} = \lambda_{ji}$, and $\lambda_{ii} = 0$ for all $i, j \in N$, where $\Lambda = [\lambda_{ij}]$ is the known interaction network.*

1. *Impose assumption on the parameters above and apply Frankel-Morris-Pauzner theorem to show that there is a unique rationalizable strategy profile (except for multiplicity at the cutoffs where a player changes her action) at the limit $\epsilon \to 0$.*

2. *Determine how the unique limiting solution varies with respect to the parameters in the game, including the network.*

3. *Take* $v_1 = \cdots = v_n = 0$. *For each network below, compute the limiting solution; comparing your answers to different network briefly discuss your finding. (You may also want to plot the networks for visualization.)*

   (a) *A star network:*

$$\Lambda = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

   (b) *A core-periphery network (missing entries are zero):*

$$\Lambda = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & & & \\ 1 & 0 & 1 & 1 & & 1 & & \\ 1 & 1 & 0 & 1 & & & 1 & \\ 1 & 1 & 1 & 0 & & & & 1 \\ 1 & & & & 0 & & & \\ & 1 & & & & 0 & & \\ & & 1 & & & & 0 & \\ & & & 1 & & & & 0 \end{bmatrix}$$

Exercise 5.15 *This question asks you to investigate how one can or cannot use non-linear monotone transformations of utility functions to apply the main results we have learned for supermodularity.*

1. *Find an optimization problem*

$$B(\pi) = \arg\max_{x \in X} E\left[u(x, \pi, \theta)\right],$$

   where $x \in X$ *is the choice variable,* $\pi \in \Pi$ *is a known parameter and* $\theta \in \Theta$ *is an unknown state, such that* B *is decreasing in* π *but* $f \circ u$ *is a supermodular function for some strictly increasing function* f.

2. *Find supermodular games* $G = (N, S, u)$ *and* $\tilde{G} = (N, S, \tilde{u})$ *such that (i) for each* $i \in N$, $\tilde{u}_i = f_i \circ u_i$ *for some strictly increasing function* f_i*, and (ii) the sets of rationalizable strategy profiles for games* G *and* $\tilde{G}$ *are different from each other.*

3. *Consider game* $G = (N, S, u)$ *and a supermodular game* $\tilde{G} = (N, S, \tilde{u})$ *such that, for each* $i \in N$, $\tilde{u}_i = f_i \circ u_i$ *for some strictly increasing function* f_i.[3] *Let* $\bar{z}, \underline{z} \in S$ *be the extremal equilibria of the supermodular game* $\tilde{G}$*. Show that*

   (a) $\bar{z}$ *and* $\underline{z}$ *are Nash equilibria of game* G*, and*

   (b) $\bar{z} \geq s \geq \underline{z}$ *for every rationalizable strategy profile* s *of game* G*.*

Exercise 5.16 *Consider a partnership game with two players, who invest in a public good project at each date* $t \in T = \{0, 1, 2, ...\}$ *without observing each other's previous investments. We assume that a strategy of a player* i *is any function* $x_i : T \to [0, 1]$*, where* $x_i(t)$ *is the (incremental) investment of* i *at* $t \in T$*. The payoff of each player* i *is*

$$U_i(x_1, x_2) = \sum_{t \in T} \delta^t \left[Af\left(\sum_{s \leq t} x_1(s), \sum_{s \leq t} x_2(s) \right) - c_i(x_i(t), t) \right]$$

where $\delta \in (0, 1)$, $A \in [0, 1]$ *is a productivity parameter,* $f : [0, 1]^2 \to \mathbb{R}$ *is a supermodular, increasing, and continuous production function, and* c_i *is a time dependent cost function for player* i*. Everything is common knowledge. (You can make differentiability assumptions to simplify the exposition.)*

1. *Show that the above game has equilibria* $\underline{x}$ *and* $\overline{x}$ *such that for each equilibrium* x *of this game,*

$$\underline{x}_i(t) \leq x_i\ (t) \leq \overline{x}_i\ (t) \qquad (\forall i, t)$$

2. *Show that, if* $A \geq A'$*, then the extremal equilibria for these parameters satisfy*

$$\underline{x}_i(t; A) \geq \underline{x}_i(t; A') \text{ and } \overline{x}_i(t; A) \geq \overline{x}_i(t; A') \qquad \forall\, (i, t)\,.$$

[3] Notice that the game G need not be supermodular.

Chapter 6

Global Games

Casual observations suggest that the outcomes are drastically different in many situations that appear to be quite similar. For example, two countries with similar natural resources and geographical attributes may have quite different wealth levels. Similarly, one partnership may breakdown, while another partnership with similar fundamentals thrives. In a less obvious example, a financial system that has been thriving based on the mutual trust may suddenly crash as the market participants lose their trust while underlying fundamentals remain similar. Traditional economic explanation for such divergence is multiple equilibria. It is presumed that the game played in these situations have multiple equilibria. Then, in one county, firm, or market, players may coordinate on a good equilibrium that is beneficial for everyone, while in another country, firm, or market with similar fundamentals, the less fortunate players coordinate on a bad equilibrium. They may be "trapped" in poverty because the unilateral deviations to good behavior only hurt individuals. Formally, the coordination games that model the above situations are supermodular, and the smallest and the largest equilibria differ for relevant parameter values under complete information.

In its simplistic form, multiple-equilibrium explanation above is too fragile to explain the high variability of the outcomes in similar situations. Carlsson and van Damme (1993) show that 2×2 games with multiple equilibria become dominance solvable when one introduces a small amount of additive independent noise in players' observations of the fundamentals. Moreover, the unique solution selects the "risk-dominant" equilibrium.

Applying the ideas of Carlsson and van Damme to a general currency-attack problem with continuum of players, Morris and Shin (1998) demonstrated that not only there is a unique solution in the incomplete information version, but also one can make sensible predictions about how the fundamentals (and policy) affects the outcome using the unique solution. This led to a large applied literature, called Global Games literature, in which one applies the methodology of Carlsson and van Damme and Morris and Shin to various economic problems.

As observed by Frankel, Morris, and Pauzner (2003), the global games are a special class of Bayesian supermodular games, studied in the previous chapter, and the monotone comparative statics one obtains from the global games is simply the monotone comparative statics for the extremal equilibria in the supermodular game. The extremal equilibria happen to be the same due to the additional structure imposed in global games. Global games structure allows one to derive monotone comparative statics implied by the underlying supermodular structure as the property of the unique solution. In contrast, in general supermodular games, one can consider selections from equilibrium set that do not exhibit the above comparative statics.

This chapter is devoted to illustration of the main ideas in global games literature, as an application of supermodular games. The first section illustrates the main ideas on a simple partnership game. The next section presents the results of Frankel, Morris, and Pauzner (2003), showing the uniqueness of rationalizable strategies in general global games with explicit supermodular structure. While the unique solution is increasing as in the monotone supermodular games, it may depend on the details of the distribution of the additive noise. The following section follows Carlsson and van Damme to show that in 2×2 games risk dominant equilibrium is selected independent of the distribution of the small additive noise. The next sections present applications with continuum of players, such as the currency attack problem, and dynamics of global games.

6.1 Risk Dominance

As discussed in the introduction, the main result of global games literature yields an equilibrium selection on the basis of a concept called, risk dominance. In this section, I formally describe the concept of risk-dominance in 2×2 games. Under symmetry, an equilibrium is risk dominant if each player's strategy is a best response against randomization with equal probabilities.

Consider an arbitrary 2×2 complete-information game (N, A, u) with strategy set $A = \{a, b\}$ for each player. Assume that both (a, a) and (b, b) are Nash equilibria. We want to select between the two equilibria. One standard selection is called *Pareto-efficient selection*. It corresponds to selecting the dominating equilibrium if one of the equilibrium Pareto-dominates the other. For example, consider the following partnership game

	a	b
a	θ, θ	$\theta - 1, 0$
b	$0, \theta - 1$	$0, 0$

(6.1)

which we will study in detail throughout this chapter. Now, when $\theta \in (0, 1)$, both (a, a) and (b, b) are Nash equilibria, and $(\theta, \theta) > (0, 0)$. Hence, Pareto-efficient selection selects equilibrium (a, a). Such a selection may not be possible; for example, no equilibrium dominates the other in the Battle of the Sexes game. While Pareto-efficient selection is often utilized in game theory and its applications (e.g. in repeated games), such selection may be highly sensitive to incomplete information, as it will be clear below.

Risk-dominance is based on a different idea and may select a Pareto-dominated equilibrium. It is related to another concept, called *p-dominance*. In a general game, an equilibrium s^* is said to be $(p_1, \ldots, p_n)$-*dominant* if s_i^* remains a best response whenever i assigns at least probability p_i on s_{-i}^* for each player i. In a 2×2 game, an equilibrium is said to be *risk dominant* if it is (p_1, p_2)-dominant for some

$$p_1 + p_2 < 1.$$

That is, it remains equilibrium even if other players may tremble with high probabilities. In particular, in symmetric games, risk dominance selects the equilibrium that is a best response to uniform distribution: it selects (a, a) when $\theta > 1/2$ and selects (b, b) when $\theta < 1/2$. In general, it is defined as follows.

When there are two strict Nash equilibria, the risk dominance can also be defined as follows. Define the players' loss from equilibrium (a, a) by

$$g_1^a = u_1(a, a) - u_1(b, a) \text{ and } g_2^a = u_2(a, a) - u_2(a, b).$$

Define the players' loss from equilibrium (b, b) similarly. Risk-dominant equilibrium is the equilibrium in which the product of those losses is greater. That is, (a, a) is said to be *risk dominant* if

$$g_1^a \cdot g_2^a > g_1^b \cdot g_2^b;$$

(b, b) is said to be *risk dominant* if

$$g_1^a \cdot g_2^a < g_1^b \cdot g_2^b.$$

For example, in the partnership game above, $g_i^a = \theta$ while $g_i^b = 1 - \theta$. Hence, (a, a) is risk dominant when $\theta < 1/2$, and (b, b) is risk dominant when $\theta > 1/2$.

Now, in symmetric 2×2 games, the latter condition is equivalent to equilibrium strategy is being a best response to the uniform distribution on other players' actions. Sometimes, risk-dominance is defined by that condition in more general games. The player may hold such a uniform belief on the basis of the "principle of insufficient reason", and as in the case of the principle itself, it is not clear on what set one should have uniform distribution: each opponent's strategy set or on the aggregate outcome.

6.2 Example: A Partnership Game

Consider a two player Bayesian game with the payoff matrix

	a	b
a	θ, θ	$\theta - 1, 0$
b	$0, \theta - 1$	$0, 0$

Each player $i \in N = \{1, 2\}$ chooses between the actions a, which corresponds to investing in a project, and b, which corresponds to not investing. If both players invest, each receives θ; if only one of them invests, he gets only $\theta - 1$. The utility function above is denoted by u_i. Due to the cost of investing when the other party does not invest, this game can be viewed as a coordination game.

When θ is in $(0, 1)$ and common knowledge, there are two equilibria in pure strategies and one equilibrium in mixed strategies. In the "good" equilibrium, anticipating that the other player invests, each player invests in the project, and each gets the positive payoff of θ. In this equilibrium, the players "coordinate" on the good outcome. In the "bad" equilibrium, each player correctly anticipates that the other party will not invest. Consequently, neither of them invests, receiving zero payoff. When $\theta < 0$, not investing strictly dominates investing, leading to (b, b) as the only rationalizable outcome.

Beliefs Following Carlsson and van Damme, we will assume that players do not know θ. Instead, each player gets a signal about θ with an additive noise. That is, each player i observes

$$x_i = \theta + \sigma\varepsilon_i, \tag{6.2}$$

where ε_i is a noise term with distribution function F and density function f, and $\sigma \in (0, 1)$ is a scalar that measures the level of uncertainty players face (namely, *shock size*). Without loss of generality, assume that

$$E(\varepsilon_1) = E(\varepsilon_2) = 0.$$

Ex ante, θ is distributed by G with density function g. Assume that $(\theta, \varepsilon_1, \varepsilon_2)$ are stochastically independent. In order to relax the common knowledge assumptions about the value of θ, we will assume that the support of G contains an interval $[-L, L]$ for some $L > 1$ so that, ex ante it is possible that either action is dominant; investment is dominant when $\theta > 1$ and not investing is dominant when $\theta < 0$. These are called the *dominance regions*, and it is this assumption that makes the game "global".

Write $F(\theta|x_i)$ for the distribution of θ conditional on x_i, which represents the interim belief of type x_i about θ. Assume further that these interim beliefs are weakly increasing in the sense of first-order stochastic dominance (henceforth FOSD):

Assumption 6.1 *$F(\theta|x_i)$ is decreasing in x_i.*

That is, as he observes a higher signal about θ, a player becomes more optimistic about the return from investment, expecting higher value for any increasing function of θ. This intuitive property is exhibited under usual (thin tailed) distributions:

Fact 6.1 *Assumption 6.1 holds whenever the noise distribution F is log-concave (i.e. $\log f$ is concave).*

Many distributions used in economic theory and applications, such as uniform, exponential and normal distributions, are log-concave, and hence the monotonicity condition here is satisfied under such noise distributions. On the other hand, log-concavity rules out fat-tailed noise distributions, such as Pareto and t distributions, and the monotonicity assumption may fail when the noise terms have fat tails.

Supermodularity Under Assumption 6.1, the Bayesian game above is monotone supermodular with order

$$a \geq b$$

on the actions. To see this, observe that the utility function

$$U_i(a_i, a_{-i}, \theta, x_1, x_2) = u_i(a_i, a_{-i}, \theta)$$

is supermodular and satisfies the continuity and measurability assumptions. Moreover, the interim belief of type x_i about (θ, x_j) is increasing in x_i in the sense of FOSD, i.e., the conditional distribution $F(\theta, x_j|x_i)$ of (θ, x_j) given x_i is decreasing in x_i. Indeed, since $x_j = \theta + \sigma\varepsilon_j$,

$$F(\theta, x_j|x_i) = \int 1_{\{\theta' \leq \theta\}} F((x_j - \theta')/\sigma)\, dF(\theta'|x_i),$$

where $1_{\{\theta' \leq \theta\}}$ is the characteristic function of the set $\{\theta' \leq \theta\}$, taking value of 1 on the set and zero outside. To see the formula, note that given θ', the signal of player j is below the fixed x_j if and only if $\varepsilon_j \leq (x_j - \theta')/\sigma$, which happens with probability $F((x_j - \theta')/\sigma)$. By integrating this under the distribution $F(\theta'|x_i)$, one obtains the above formula. Note that the function $1_{\{\theta' \leq \theta\}} F((x_j - \theta')/\sigma)$ is decreasing in θ'. Since $F(\theta'|x_i)$ is decreasing in x_i, this implies that $F(\theta, x_j|x_i)$ is decreasing in x_i.

Extremal Equilibria and Rank Beliefs Since the game is monotone supermodular game and symmetric, there exist symmetric and monotone extremal equilibria s^* and s^{**} such that

$$s^* \geq s \geq s^{**}$$

for all rationalizable strategy profiles s. Note also that $s_i^*(-L) = s_i^{**}(-L) = b$ and $s_i^*(L) = s_i^{**}(L) = a$ as b and a are strictly dominant at $x_i = -L$ and $x_i = L$, respectively. Hence, the monotone symmetric equilibria s^* and s^{**} are in cutoff strategies, where a type plays a above the cutoff and b below the cutoff.

The equilibrium cutoffs are determined by players' *rank belief function*:

$$R(x_i) \equiv \Pr(x_j < x_i | x_i),$$

probability a type x_i assigns to that the other player has a lower type. In any symmetric equilibrium with a cutoff $\hat{x}$, a cutoff type $x_i = \hat{x}$ assigns probability $R(x_i)$ to the event that the other player does not invest. Hence, his expected payoff from a is

$$E[\theta|\hat{x}] - R(\hat{x}).$$

But the cutoff type must be indifferent between a and b, which gives 0.[1] Therefore, the cutoff must satisfy

$$R(\hat{x}) = E[\theta|\hat{x}]. \tag{6.3}$$

By Assumption 6.1, this condition is also sufficient for having an equilibrium with cutoff $\hat{x}$, and thus the solutions to (6.3) coincide with the equilibrium cutoffs. The following result then follows from Theorem 5.8.

Proposition 6.1 *Let x^* and $x^{**} \in (0,1)$ be the smallest and the largest solutions to equations $R(\hat{x}) = E[\theta|\hat{x}]$. Then, a is the unique (interim correlated) rationalizable action for any type $x_i > x^{**}$, and b is the unique rationalizable action for any type $x_i < x^*$. Conversely, there exist Bayesian Nash equilibria s^* and s^{**}, defined by*

$$s_i^*(x_i) = \begin{cases} a & \text{if } x_i \geq x^*, \\ b & \text{if } x_i < x^*, \end{cases} \quad \text{and } s_i^{**}(x_i) = \begin{cases} a & \text{if } x_i > x^{**}, \\ b & \text{if } x_i \leq x^{**}. \end{cases}$$

The proposition states that, under rationalizability, a player must invest if his signal about the return from investment is above a threshold x^{**} and he must not invest if his signal is below a threshold x^*. There is multiplicity in between the threshold in that either action could be played in some symmetric monotone equilibrium. The key here is that, under incomplete information, the above thresholds are distinct from the thresholds for dominance regions. In particular, in the region $(x^{**}, 1)$, investment is uniquely rationalizable although it is not dominant. In contrast, with complete information, players could coordinate on either actions $(x^{**}, 1)$.

[1] Afterall, because of the dominance regions, there are nearby types $x_i > \hat{x}$ who choose a and there are nearby types $x_i < \hat{x}$ who chose b. Since the beliefs are payoffs are continuous, $\hat{x}$ must be indifferent.

Contagion Multiplicity disappears under incomplete information because of the *contagion* from the dominance regions. A player invests if his signal is above the dominance cutoff

$$\bar{x}^0 = 1$$

for investment. Indeed, the lowest strategy $\underline{s}_j$ is defined by

$$\underline{s}_j(x_j) = b \text{ for all } x_j,$$

and the lowest best response to $\underline{s}_j$ is

$$\underline{B}_i(\underline{s}_j) = \begin{cases} a & \text{if } x_i > \bar{x}^0, \\ b & \text{if } x_i \leq \bar{x}^0. \end{cases}$$

Hence, a rational player must invest above $\bar{x}^0$ and may choose not to invest below $\bar{x}^0$. If he knows that the other player is also rational, then he assigns probability 1 on strategies that are weakly greater than $\underline{B}_j$ (i.e., the other player invests whenever above $\bar{x}^0$). Hence, he invests if his signal is above the cutoff $\bar{x}^1$ where

$$\Pr\left(x_j < \bar{x}^0 | x_i = \bar{x}^1\right) = E\left[\theta | x_i = \bar{x}^1\right].$$

Indeed,

$$\underline{B}_i^2(\underline{s}_j) = \begin{cases} a & \text{if } x_i > \bar{x}^1, \\ b & \text{if } x_i \leq \bar{x}^1. \end{cases}$$

Iterating this argument, as in the proof of Theorem 5.5, one can conclude that under kth-order knowledge of rationality, a player must invest if his signal is above $\bar{x}^k$ where

$$\Pr\left(x_j < \bar{x}^{k-1} | x_i = \bar{x}^k\right) = E\left[\theta | x_i = \bar{x}^k\right].$$

The above sequence converges to x^{**}.

The key observation of Carlsson and Van Damme is that, when the shock size σ is sufficiently small vis a vis the ex-ante distribution G of θ, the cutoffs x^* and x^{**} are close to each other and it is uniquely rationalizable to play the risk dominant action. This is because, conditional on x_i, θ has an approximately uniform distribution, and those cutoffs coincide with the risk dominance threshold in the limit—as we will see next.

6.2.1 The case of uniform prior and small shock sizes

Assume that θ is distributed uniformly on a large interval $[-L, L]$ where σ/L is sufficiently small and $L \gg 1+\sigma$. Assume also that ε_i takes values in $[-1,1]$. Then for any $\hat{x} \in (0,1)$, the rank belief is

$$R(\hat{x}) \equiv \Pr(x_j < x_i | x_i = \hat{x}) = \Pr(\varepsilon_j < \varepsilon_i) = 1/2,$$

and the expected return from investment is

$$E[\theta | x_i = \hat{x}] = \hat{x}.$$

Hence, the indifference condition $R(\hat{x}) = E[\theta|\hat{x}]$ reduces to

$$\hat{x} = 1/2.$$

Therefore, by Proposition 6.1, the game is dominance solvable:[2]

Proposition 6.2 *For any type $x_i \neq 1/2$, the unique (interim correlated) rationalizable action is*

$$s_i^*(x_i) = \begin{cases} a & \text{if } x_i > 1/2, \\ b & \text{if } x_i < 1/2. \end{cases}$$

Excluding the cutoff value $\hat{x} = 1/2$, the proposition states that the resulting game is dominance-solvable. Under the unique solution, a is played if and only if $x_i > 1/2$. When σ is close to zero, x_i is approximately equal to θ, and a is played if and only if

$$\theta > 1/2.$$

A comparison to the limit case $\sigma = 0$ is useful. For $\sigma > 0$, the extremal equilibrium strategies differ only at the cutoff $1/2$ as in Proposition 6.1. For $\sigma = 0$, the extremal equilibria are also in cutoff strategies, but they are distinct on $[0,1]$:

$$s_i^*(x_i) = \begin{cases} a & \text{if } x_i \geq 0, \\ b & \text{if } x_i < 0, \end{cases} \quad \text{and } s_i^{**}(x_i) = \begin{cases} a & \text{if } x_i > 1, \\ b & \text{if } x_i \leq 1. \end{cases}$$

While the sets of equilibria and rationalizable strategies are upper-hemicontinuous with respect to σ; they are not lower-hemicontinuous. There are many more equilibria in the

[2] Throughout this chapter, rationalizability refers to interim correlated rationalizability.

limit game $\sigma = 0$. Indeed, each type profile $x_1 = x_2 = \theta$ corresponds to a separate complete information game, which has three equilibria when $\theta \in (0,1)$. Any selection from that correspondence, while θ varies, yields a Bayesian Nash equilibrium. In particular there are many non-monotone Bayesian Nash equilibria.

What if θ is not uniformly distributed? In that case the cutoffs x^* and x^{**} can be quite different and one can select any action outside of the dominance region, depending on the prior distributions, as we will see below in the example with Normal distributions. However, when σ is small, the ex-ante distribution is approximately uniform from the interim perspective, leading to selection of risk-dominant action:

Proposition 6.3 *Assume that f and g are Lipschitz continuous and g is positive on a closed interval that contains $[0,1]$. Then, for any $\epsilon > 0$, there exists $\bar{\sigma} > 0$ such that, for all $\sigma \in (0, \bar{\sigma})$, a is uniquely rationalizable whenever $x_i > 1/2 + \epsilon$ and b is uniquely rationalizable whenever $x_i < 1/2 - \epsilon$.*

To see the proof of proposition, consider the case with bounded distribution for ε_i, and restrict ε_i to $[-1,1]$. Then, $F(\theta|x_i)$ puts only positive probability on $[x_i - \sigma, x_i + \sigma]$. On that region, by Lipschitz continuity, $g(\theta) \in [g(x_i) - \sigma c, g(x_i) + \sigma c]$ for some constant c. When $\sigma c \ll \min_{\theta \in [0,1]} g(\theta)$, g is approximately uniform, yielding

$$R(x_i) \cong 1/2.$$

Moreover, as $\sigma \to 0$, for every $\hat{x} \in [0,1]$, $E[\theta|x_i = \hat{x}]$ approaches $\hat{x}$, showing that the extremal cutoffs approach $1/2$.

The risk dominant selection above is partly due to the common prior assumption. The next exercise shows that either equilibrium can be selected once the common prior assumption is dropped.

Exercise 6.1 *In the above game drop the common-prior assumption about the distribution of $(\varepsilon_1, \varepsilon_2)$ while keeping the rest of the model as is. Assume instead that, according to each player i, the probability that the signal of the other player is higher with probability q for some arbitrary $q \in (0,1)$:*

$$\Pr_i(\varepsilon_j > \varepsilon_i) = q.$$

Exercise 6.2 *In the original model, assume that the common distribution of* ε_1 *and* ε_2 *is discrete:*

$$\Pr(\varepsilon_i = -1) = \Pr(\varepsilon_i = 0) = \Pr(\varepsilon_i = 1) = 1/3.$$

6.2.2 The Case of Normal Distributions

Assume that the return from investment and the idiosyncratic noise terms in players signals have standard Normal distributions:

$$\theta = y + \tau\eta \text{ and } x_i = \theta + \sigma\varepsilon_i \text{ where } (\eta, \varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0,1) \tag{6.4}$$

where parameters y, $\sigma > 0$ and $\tau > 0$ are known. Here, y represents the ex-ante expected return; $\tau\eta$ represents the common shock to the fundamentals, and $\sigma\varepsilon_i$ is an idiosyncratic noise term. Observe that type x_i expects that the return from investment is

$$E[\theta|x_i] = y + \alpha(x_i - y)$$

where

$$\alpha = \frac{\tau^2}{\sigma^2 + \tau^2}.$$

Intuitively, as he observes a deviation $x_i - y$ between his signal and the ex-ante mean, he attributes α fraction of it to a shock $\tau\eta$ to fundamentals and $1 - \alpha$ fraction of it to a noise in his signal, where the fraction α is determined by the relative sizes of the variances of the shock and the noise. In this model, signal x_i has no intrinsic value. The economically relevant variable is the conditional expectation $E[\theta|x_i]$ of investment return. The next result presents rank belief function in terms of $E[\theta|x_i]$.

Lemma 6.1 *In the normal model (6.4), the rank belief of any type* x_i *is*

$$R(x_i) = \Phi(\alpha\lambda(x_i - y)) = \Phi(\lambda(E[\theta|x_i] - y))$$

where Φ *is the cumulative distribution function of* $N(0,1)$ *and*

$$\lambda = \frac{\sigma}{\tau^2}\frac{1}{\sqrt{\alpha+1}} \in \left(\frac{1}{\sqrt{2}}\frac{\sigma}{\tau^2}, \frac{\sigma}{\tau^2}\right)$$

is the sensitivity of rank beliefs to fundamental expectations.

Proof. Conditional on x_i, ε_i is normally distributed with

$$\begin{aligned} E[\varepsilon_i|x_i] &= (1-\alpha)(x_i - y)/\sigma \\ Var_{\varepsilon_i|x_i} &= \alpha. \end{aligned}$$

Hence, the difference $\varepsilon_j - \varepsilon_i$ is normally distributed with mean $-E[\varepsilon_i|x_i]$ and variance $\alpha + 1$. Therefore, the rank belief is

$$R(x_i) = \Pr(\varepsilon_j - \varepsilon_i \leq 0|x_i) = \Phi\left(\frac{E[\varepsilon_i|x_i]}{\sqrt{\alpha+1}}\right) = \Phi\left(\frac{1-\alpha}{\sigma\sqrt{\alpha+1}}(x_i - y)\right) = \Phi(\alpha\lambda(x_i - y)).$$

■

The above lemma establishes that, under normal distributions, the difference $E[\theta|x_i] - y$ between the interim and ex-ante expectations translates into rank beliefs via a sensitivity coefficient λ. Since $\alpha \in (0,1)$, the sensitivity coefficient λ is in the order of σ/τ^2. Higher deviations between the interim and ex-ante expectations lead to higher rank belief, which vary between 0 and 1. The other parameters of the models, namely σ and τ, affect the rank beliefs only through their effect on λ and the interim expectation, which will be taken as the main ingredient of the analysis by itself. By Lemma 6.1, the equation $R(\hat{x}) = E[\theta|\hat{x}]$ for equilibrium cutoffs can be written in terms of cutoff for expectations at which the players start investing:

$$\Phi(\lambda(E[\theta|\hat{x}] - y)) = E[\theta|\hat{x}] \tag{6.5}$$

The equilibrium cutoff here depends only on λ and y.

Figure 6.1 plots the equilibrium cutoffs $E[\theta|\hat{x}]$ for expectations as functions of y for various values of λ. For small $\lambda = 0.1$, there is a unique equilibrium cutoff and the cutoff is near $1/2$ in the region plotted. Thus, the players play the risk dominant action under rationalizability, as in the case of uniform priors. Indeed, for any fixed y, and any expectation $E[\theta|x_i]$, as $\lambda \to 0$, rank belief R converges to $\Phi(0) = 1/2$, and thus the equilibrium cutoff $E[\theta|\hat{x}]$ converges to $1/2$. A special case of this limit is the Carlsson and van Damme limit above: $\sigma \to 0$ for a fixed (τ, y).

For any positive value of λ, the rank belief $\Phi(\lambda(E[\theta|x] - y))$ varies from 1 to 0 as y varies from $-\infty$ to ∞ for any fixed $E[\theta|x]$. Hence, there is a unique equilibrium cutoff $E[\theta|\hat{x}]$ for small and large values of y. Moreover, the unique cutoff is nearly 1 when y is small. That is, when the players are highly pessimistic ex-ante, they do not invest unless

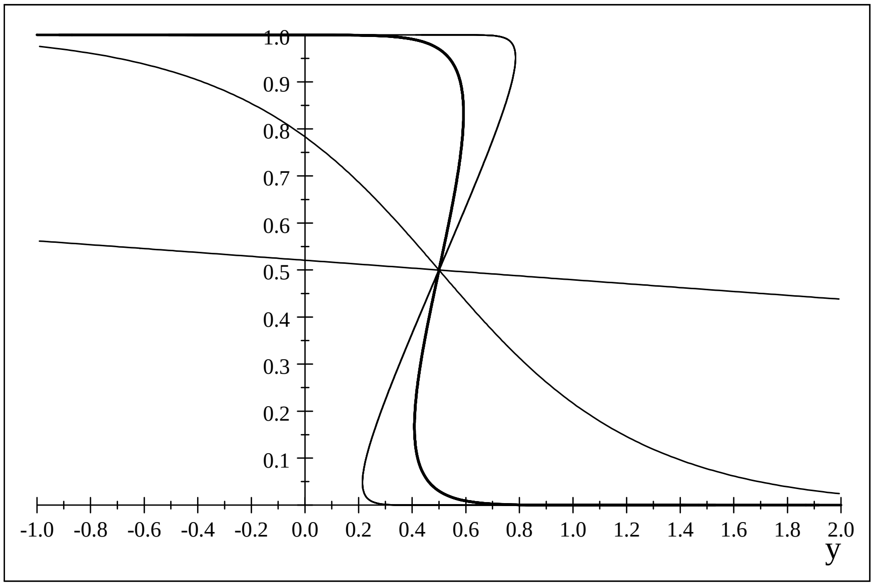

Figure 6.1: The equilibrium cutoff $E[\theta|\hat{x}]$ in Normal example as a function of y for $\lambda = 0.1, 1, 4, 10$.

it is nearly dominant to do so. This is because they do not think that the other players will be as optimistic (i.e. they are pessimistic about the other players). Symmetrically, for large values of y, the unique cutoff is nearly 0 and ex-ante optimistic players invest unless not investing is nearly dominant. This is true for all positive λ. For small λ, such as 0.1, ex-ante beliefs must be highly extreme in order to have such an impact. For larger values of λ, the effect of ex-ante mean is more prominent. For example, in the figure, for $\lambda = 1$ (e.g. $\sigma^2 = \sqrt{2}$ and $\tau^2 = 1$), there is a unique equilibrium cutoff where the required expected return is above 0.9 for $y = -1$ and it is below 0.1 for $y = 2$.

For larger values of λ, the rank beliefs become highly sensitive to expectations of the fundamentals near the ex-ante mean, resulting in multiple equilibria. This can be seen in the figure for $\lambda = 4$, plotted in thicker lines. In that case, there are multiple equilibria for values of y between 0.4 and 0.6. In one equilibrium, the cutoff is nearly 0, so that the players invest so long as it is efficient to do so, and in another equilibrium, the cutoff is nearly 1, so they do not invest unless it is nearly dominant. Outside of the multiplicity region above, there is a unique rationalizable action where the cutoff is nearly 0 for relatively high values of y and 1 for relatively low values of y. As λ gets larger, the region of multiplicity grow and the extremal cutoffs become more extreme—

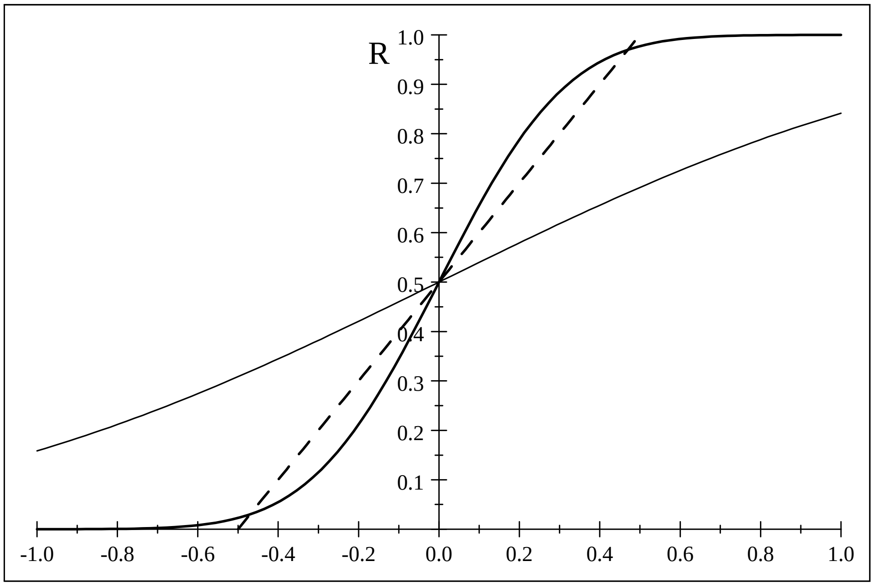

Figure 6.2: Equilibrium Cutoffs under Normal Distribution–in terms of the difference $E[\theta|x] - y$, plotted on the horizontal axis. (The rank beliefs are for $\lambda = 1$—in thin lines–and for $\lambda = 4$—in thick lines.)

both in multiplicity and in uniqueness regions. This can be seen for $\lambda = 10$ in the figure.

When does multiplicity arise? Towards an answer, consider Figure 6.2, where rank belief function is plotted for $\lambda = 1$ and $\lambda = 4$ (as a function of $E[\theta|x_i]-y$). The expected return $E[\theta|x_i]$ for $y = 1/2$ is plotted as the dashed line with slope 1. For $\lambda = 4$, the slope of R is larger than 1 at the origin. Consequently, it intersects $E[\theta|x_i]$ from below. This results in two additional cutoffs—one near 1 and one near 0 as discussed above. For $\lambda = 1$, the slope of R is smaller than 1 at the origin. Since the slope of Φ is highest at the origin, this implies that its slop is below 1 everywhere. In that case, there exists a unique equilibrium for all values of y. More broadly, there is a unique equilibrium for all values of y if and only if the slope of R is less than or equal to 1 at the origin, i.e.,

$$\lambda \leq \sqrt{2\pi}.$$

There is a unique rationalizable strategy for all values of y when $\lambda \leq \sqrt{2\pi}$, and there are multiple equilibria near $y = 1/2$ otherwise. These multiplicity and uniqueness regions are plotted in Figure 6.3.

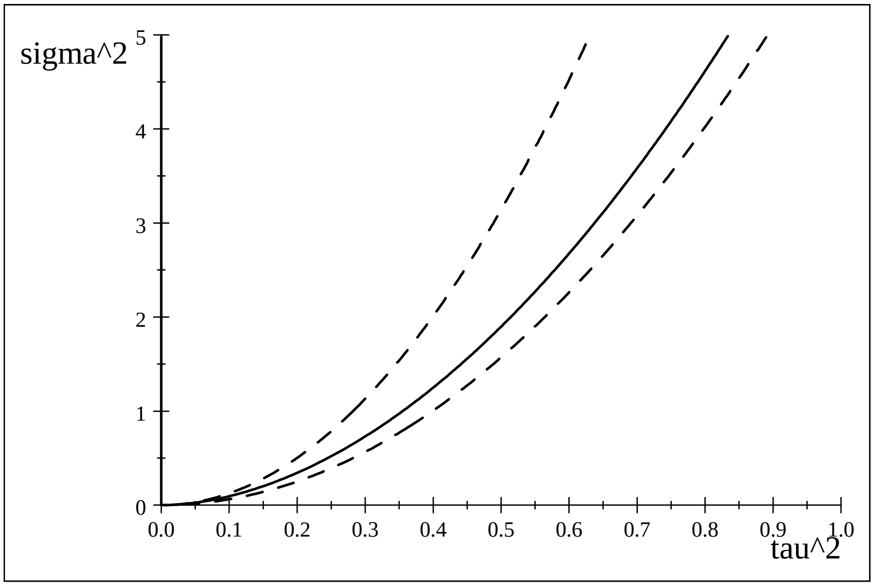

Figure 6.3: Multiplicity and uniqueness regions in Normal case. There is multiplicity above the solid curve and uniqueness below the solid curve. The dash curves are isocurves of σ/τ^2 at $\sqrt{2\pi}$ and $2\sqrt{\pi}$. (The figure is taken from Morris and Shin.)

6.2.3 Model Uncertainty

Imagine that the players do not know the underlying statistical model that generates the fundamentals. In particular, they think that the common shock η is normally distributed with mean zero and some unknown variance ξ^2 where $1/\xi^2$ has χ^2 distribution with k degrees of freedom. That is,

$$\eta = \xi\bar{\eta} \text{ where } \bar{\eta} \sim N(0,1) \text{ and } 1/\xi^2 \sim \chi^2(k). \tag{6.6}$$

Overall, the common shock η has t-distribution with k degrees of freedom. Such beliefs arise often when the population mean and variance is estimated from some normally distributed sample. In this example, the sample statistics are computed from publicly available data (possibly by a third party) and the players also observe additional private signal about the fundamentals with additive noise as before. For simplicity, assume that shock sizes for the common shock and the noise are identical (i.e. $\tau = \sigma$):

$$\theta = y + \sigma\eta \text{ and } x_i = \theta + \sigma\varepsilon_i \text{ where } \varepsilon_i \sim N(0,1). \tag{6.7}$$

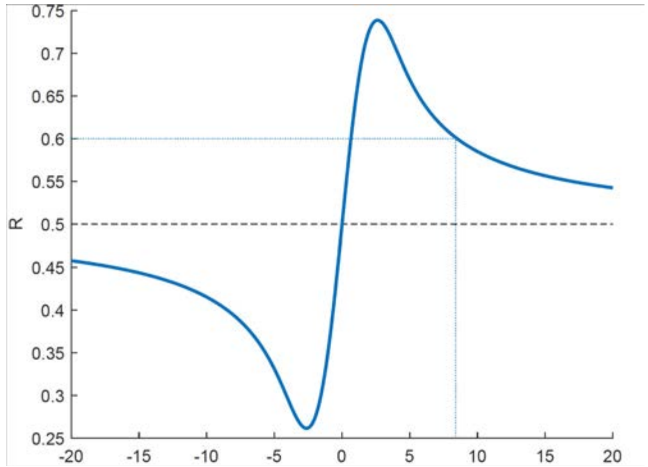

Figure 6.4: Rank belief function under model uncertainty (as a function of $z = (x_i - y)/\sigma$).

How does such model uncertainty impact the rationalizable behavior in our ongoing example?

Towards answering this question, under (6.6-6.7), the rank beliefs can be written as

$$R(x_i) = E\left[\Phi\left(\alpha(\xi)\lambda(\xi)(x_i - y)\right)|x_i\right].$$

That is, observing x_i, player i updates his beliefs about the variance ξ^2 and uses these updated beliefs to compute the expectation of rank beliefs under normal model with variance ξ^2. An increase in $x_i > y$ has two impact on rank beliefs. First, the deviation $(x_i - y)$ increases, increasing R. Second, player i comes to believe that the variance ξ^2 of the common shock is larger. This decreases R because

$$\alpha(\xi)\lambda(\xi) = \frac{1}{\sigma\sqrt{(1+\xi^2)(1+2\xi^2)}}$$

is decreasing in ξ. Intuitively, with larger variance in common shock, the player attributes a larger fraction of the deviation to the common shock (i.e. α is larger) and thereby holds more optimistic beliefs about the other players. The latter indirect effect dominates when $|x_i - y|$ is large, leading to a decreasing rank belief function near the tails.

This is vividly illustrated in Figure 6.4. Note that R depends only on the normalized shock $z = (x_i - y)/\sigma$. At $z = (x_i - y)/\sigma = 0$, the direct effect dominates and the rank

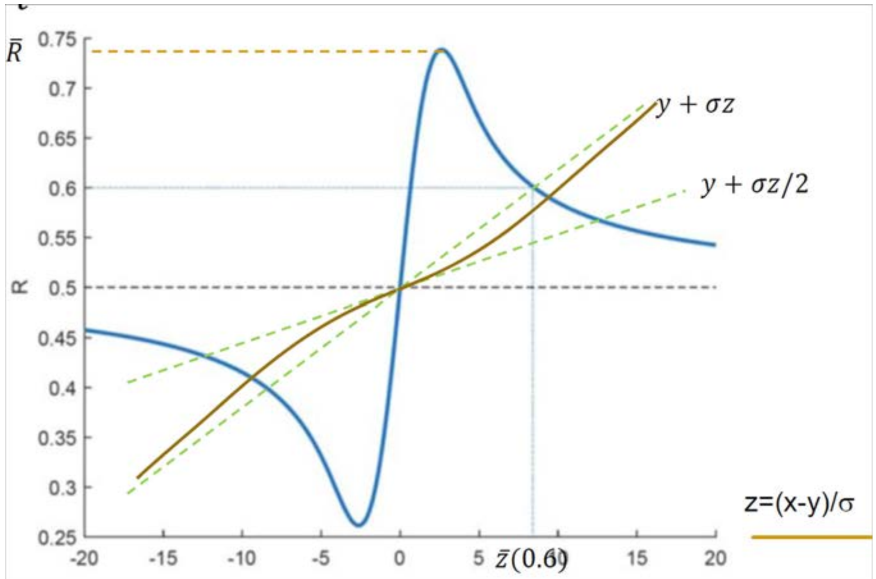

Figure 6.5: Extremal equilibrium cutoffs under model uncertainty ($\sigma = 0.01$).

belief function is increasing. As x_i gets larger, the player updates his beliefs more dramatically offsetting some of the direct effect. Eventually, the indirect effect overwhelms the direct effect and the rank belief function start decreasing. As the player observes very large deviations $x_i - y$ that are highly unlikely under small ξ, he attributes nearly all of the deviation to the large variance in the common shock, concluding that the other player is nearly as optimistic as he is (i.e. R is nearly 1/2). Such a non-monotonicity of rank beliefs have important strategic implications.

Towards exploring the strategic implications of model uncertainty, the extremal equilibria are plotted in Figure 6.5. Note that the expected return $E[\theta|x_i]$ remains in between $x_i = y + \sigma z$ and $y + \sigma z/2$, which is the expected return under normal distribution without model uncertainty. When z is near 0, $E[\theta|x_i]$ remains near the bound $y + \sigma z/2$, as the player attributes nearly half of the deviation to the common shock, as in the case without model uncertainty ($\alpha \cong 1/2$). After $z \cong 5/2$, $E[\theta|x_i]$ rises somewhat sharply towards $y + \sigma z$ because the player updates his belief about the variance of common shock substantially, leading to a sharp increase in α. This is when R starts decreasing. The expected return $E[\theta|x_i]$ approaches the upper bound $x_i = y + \sigma z$, as α approaches 1. The extremal equilibrium cutoffs x^* and x^{**} are approximately when $x_i = y + \sigma z$

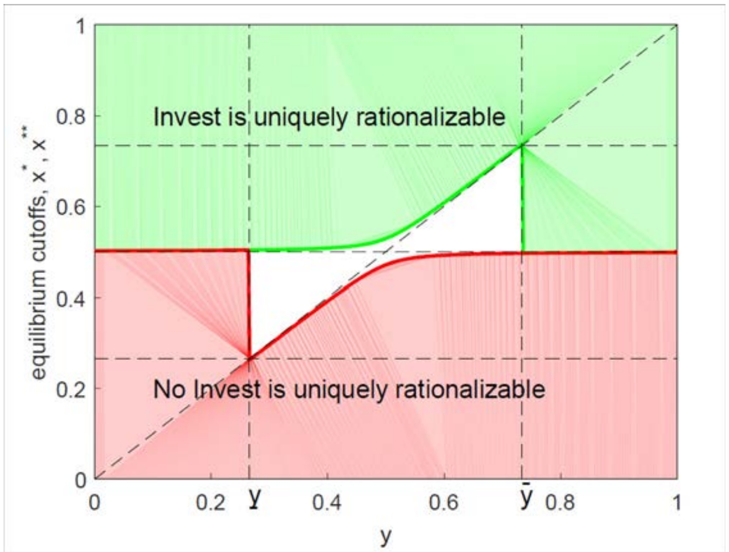

Figure 6.6: Equilibrium cutoffs and rationalizability under model uncertainity, as a function of prior belief y.

intersect R; i.e., $R(x_i) = x_i$. The equilibrium cutoffs are plotted in Figure 6.6 as a function of y.

The first implication of non-monotonicity is that the rank beliefs remain bounded away from 0 and 1:

$$\bar{R} = \sup_z R(z) \cong 0.74.$$

Since $E[\theta|x_i] < x_i$, this implies that $x^{**} < \bar{R}$ for all values of y (and σ). Hence, invest is uniquely rationalizable whenever the expected return $E[\theta|x_i]$ exceeds $\bar{R}$. This leads to a large region of interim expectations for which invest is uniquely rationalizable for all prior beliefs although it is not dominant—as in Figure 6.6. In contrast, without model uncertainty, the equilibrium cutoff approaches 1 for highly negative y—as shown in the previous section. Second, invest is uniquely rationalizable whenever it is risk dominant and there is a large shock. To state this formally, for $\theta > 1/2$, define

$$\bar{z}(\theta) = \max\{z > 0 | R(y + \sigma z) = \theta\},$$

observing that $\bar{z}(\theta) = -\infty$ when $\theta > \bar{R}$. Moreover, since R approaches $1/2$ as $z \to \infty$, $\bar{z}(\theta)$ is decreasing and finite for all $\theta \in (1/2, \bar{R})$, approaching ∞ as θ approaches $1/2$. Since R is independent of σ and y, so is $\bar{z}(\theta)$. Since $E[\theta|x_i] < x_i$ for $z > 0$, the following is immediate:

Proposition 6.4 *Under model uncertainty, invest is uniquely rationalizable whenever it is risk dominant (i.e.* $E[\theta|x_i] > 1/2$*) and there is a large shock:*

$$E[\theta|x_i] - y \geq \sigma\bar{z}\left(E[\theta|x_i]\right).$$

The same result holds for not investing (with negative large shock). One can also prove a converse of this result for $E[\theta|x_i] \ll \bar{R}$. This can be viewed also from Figure 6.6, where x^{**} is near $\max\{y, 1/2\}$ and increasing in y until it gets close to $\bar{R}$. In contrast, without model uncertainty, large shocks do not play an important role. On the contrary, the equilibrium cutoff x^{**} is a decreasing function of prior expectation, and hence for a given value $E[\theta|x_i]$, a larger shock $E[\theta|x_i] - y$ will only prevent invest being uniquely rationalizable. Indeed, for any $E[\theta|x_i] < 1$, invest is *not* uniquely rationalizable when the shock $E[\theta|x_i] - y$ is sufficiently large. This is because, when they face a large shock, the players evaluate their rank quite differently when there is model uncertainty.

6.3 Global Supermodular Games

More generally, in a monotone supermodular game with one dimensional action and type spaces, if the players' private information is generated by signals with small additive noise, the game is approximately dominance solvable, in that all rationalizable strategies converge to a unique strategy almost everywhere as the size of the additive noise goes to zero. In the 2×2 example above, the limit solution was independent of the noise distribution, selecting the risk-dominant action everywhere. This is no longer true.

Formally, consider a Bayesian game G^{σ}, indexed by the shock size σ:

- the set of players is $N = \{1, \ldots, n\}$;
- the set of payoff parameters is a closed interval $\Theta \subseteq \mathbb{R}$;
- for each player i, the set of actions A_i is a countable union of closed intervals within $[0, 1]$ where $0, 1 \in A_i$;
- the payoff function $u_i : A \times \Theta \to \mathbb{R}$ is continuous with bounded derivatives;

- each player observes a signal

$$x_i = \theta + \sigma\varepsilon_i$$

  where $(\theta, \varepsilon_1, \ldots, \varepsilon_n)$ are stochastically independent with atomless densities, θ has full support, and the noise terms $\varepsilon_1, \ldots, \varepsilon_n$ are bounded.

Towards imposing the supermodularity assumptions, write

$$\Delta u_i\left(a_i, a_i', a_{-i}, \theta\right) = u_i\left(a_i, a_{-i}, \theta\right) - u_i\left(a_i', a_{-i}, \theta\right)$$

for the marginal contribution of increasing one's action from a_i' to a_i for a fixed (a_{-i}, θ). The next assumption states the supermodularity assumptions, implying that the game is monotone supermodular.

Assumption 6.2 *Each payoff function u_i satisfies the following conditions:*

Strategic Complementarity *For any $a_i \geq a_i'$ and $a_{-i} \geq a_{-i}'$,*

$$\Delta u_i\left(a_i, a_i', a_{-i}, \theta\right) \geq \Delta u_i\left(a_i, a_i', a_{-i}', \theta\right).$$

Dominance Regions *There exist $\underline{\theta}$ and $\bar{\theta}$ in the interior of Θ such that 0 is strictly dominant whenever $\theta < \underline{\theta}$, and 1 is strictly dominant whenever $\theta > \bar{\theta}$.*

State Monotonicity *There exists $K > 0$ such that for all $a_i \geq a_i'$, a_{-i}, and $\bar{\theta} > \theta \geq \theta' > \underline{\theta}$,*

$$\Delta u_i\left(a_i, a_i', a_{-i}, \theta\right) - \Delta u_i\left(a_i, a_i', a_{-i}, \theta'\right) \geq K\left(a_i - a_i'\right)\left(\theta - \theta'\right).$$

Here, strategic complementarity condition is the standard supermodularity condition, making G^σ supermodular. Existence of dominance regions is what makes the game "global" in that certain preference are not excluded a priori. This is a weak richness assumption on preferences, requiring existence of preferences under which the extreme actions are dominant, but not making any further assumption. In particular, there need not be preferences under which a non-extreme action is dominant. Finally, state monotonicity requires that G^σ is a monotone supermodular game. Monotone supermodularity would simply require the above condition for $K = 0$. State monotonicity further requires that the sensitivity of marginal contribution with respect to state θ is bounded away from 0.

Since G^σ is a monotone supermodular game, by Theorem 5.8, it has extremal Bayesian Nash equilibria $\bar{s}^\sigma$ and $\underline{s}^\sigma$ in pure strategies where $\bar{s}^\sigma$ and $\underline{s}^\sigma$ are weakly increasing and

$$\bar{s}_i^\sigma(x_i) \geq s_i(x_i) \geq \underline{s}^\sigma(x_i) \qquad (\forall i, x_i)$$

for every (interim correlated) rationalizable strategy s_i. As $\sigma \to 0$, the extremal equilibria $\bar{s}^\sigma$ and $\underline{s}^\sigma$ converge to weakly increasing strategy profiles $\bar{S}$ and $\underline{S}$, respectively, everywhere. The next result shows that in the limit the extremal equilibria coincide, in that $\bar{S}(x) = \underline{S}(x)$ at every x where both functions are continuous. Consequently, all rationalizable actions are within a small neighborhood of $\bar{S}(x)$ when σ is small. This is stated in the next result, where $S_i^\infty[x_i|G^\sigma]$ is the set of interim correlated rationalizable actions of type x_i in game G^σ.

Theorem 6.1 *There exists a weakly increasing strategy profile s^* such that, for every $x = (x_1, x_1, \ldots, x_1)$ at which s^* is continuous,*

1. *for every $\epsilon > 0$, there exists $\bar{\sigma} > 0$ such that*

$$S_i^\infty[x_i|G^\sigma] \subset (s^*(x_i) - \epsilon, s^*(x_i) + \epsilon) \qquad (\forall \sigma < \bar{\sigma});$$

2. *$s^*(x)$ is a Nash equilibrium of the complete information game in which it is common knowledge that $\theta = x_1$.*

Note that since s^* is weakly increasing, it is continuous everywhere except for a countably many points at which it jumps up. Clearly, when it jumps up, by upper-hemicontinuity of rationalizability, the extremal rationalizable actions are distinct and remain bounded away from each other. In that case, $s^*(x)$ can be selected from that set arbitrarily. The theorem is stated for remaining types. In that case, the first part states that there is a unique rationalizable action for vanishingly small noise, in that all rationalizable actions converge to $s^*(x_i)$ as $\sigma \to 0$. The second part states that the unique solution in the limit is a Nash equilibrium of complete information game. This part is a upper-hemicontinuity property that immediately follows from continuity of payoffs.

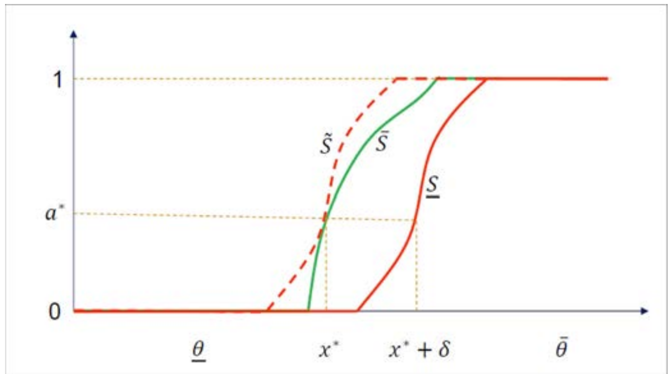

Figure 6.7: An illustration for (a failure) of Theorem 6.1.

Exercise 6.3 *Show that $\bar{S}(x)$ is a Nash equilibrium of game at which it is common knowledge that $\theta = x_1$ for any $x = (x_1, x_1, \ldots, x_1)$ at which $\bar{S}$ is continuous, assuming $\bar{s}^\sigma$ converges uniformly over a neighborhood of x.*

For an intuition for the first part, consider a symmetric game where θ is uniformly distributed, $A_i = [0,1]$, and there is a unique best response to every strategy (e.g. payoffs are strictly quasi-concave in own action). In that case, the extremal equilibria are also symmetric, and I will write drop the subscript, writing $\bar{S}(x)$ for the action played at signal x, where $\sigma > 0$ is fixed. Towards a contradiction, suppose that $\bar{S}(x) > \underline{S}(x)$ on an open interval. Then there exist some $\delta > 0$, a signal value x^*, and a strategy $\tilde{S}$, such that $\tilde{S}(x^*) = \underline{S}(x^* + \delta) = \bar{S}(x^*)$ and $\tilde{S}(x) = \underline{S}(x + \delta) \geq \bar{S}(x)$ for all x, as in Figure 6.7. Assuming $\tilde{S}(x^*) > 0$, this leads to a contradiction:

$$B\left(\tilde{S}|x^*\right) \geq B\left(\bar{S}|x^*\right) = \bar{S}(x^*) = \underline{S}(x^* + \delta) = B\left(\underline{S}|x^* + \delta\right) > B\left(\tilde{S}|x^*\right)$$

where $B(S|x)$ stands for the unique best response to S for type x. Here, the first equality follows from strategic complementarity and $\tilde{S} \geq \bar{S}$; the next equalities hold because $\bar{S}(x^*) = \underline{S}(x^* + \delta)$ and $\bar{S}$ and $\underline{S}$ are Bayesian Nash equilibria. The key strict inequality follows from uniform prior and state monotonicity as follows. Because of the uniform prior, the beliefs $p_{\theta,x_{-i}|x^*}$ and $p_{\theta,x_{-i}|x^*+\delta}$ of types $x_i = x^*$ and $x_i = x^* + \delta$, respectively, on (θ, x_{-i}) satisfy

$$p_{\theta,x_{-i}|x^*+\delta} = p_{\theta,x_{-i}|x^*} \circ \Delta^{-1}$$

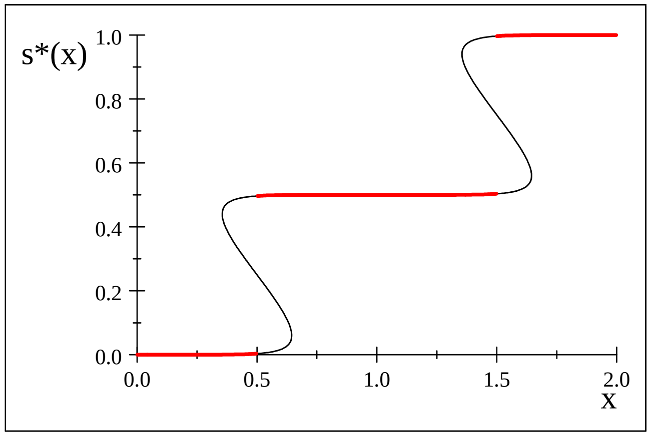

Figure 6.8: An illustration of Theorem 6.1

where $\Delta(\theta, x_{-i}) = (\theta + \delta, x_1 + \delta, \ldots, x_n + \delta)$. That is, type $x^* + \delta$ simply adds δ to the fundamental states and the other players' types. Then, by definition, under $\underline{S}$, beliefs of type $x^* + \delta$ on (θ, a_{-i}) are as in the beliefs of type x^* under $\tilde{S}$, except that he thinks that $\delta > 0$ is added to θ. Then, by state monotonicity, his best response must be higher than $B\left(\tilde{S}|x^*\right)$.

The equilibrium selection from Nash equilibrium is illustrated in Figure 6.8, where once again the game is symmetric and $A_i = [0, 1]$. Under complete information, by supermodularity and monotonicity, the symmetric equilibrium correspondence as in the figure. For example, at $x = \theta = 1/2$, there are three symmetric Nash equilibria, one is near 0, one is near $1/2$, and one is in between. In this graph, the equilibria where the graph is increasing are stable in that the best response dynamics would take us back to equilibrium if we perturb the strategies. For example, the extreme equilibria are stable at $\theta = 1/2$. The equilibria where the graph is decreasing are unstable, in that the best response dynamics would take us further away from the equilibrium if we perturb the strategies. The middle equilibrium at $\theta = 1/2$ is unstable. Since s^* is weakly increasing and a selection from the above Nash equilibrium graph, it selects a stable equilibrium for each complete information game at which s^* is continuous, as in the figure. (Of course, at singular points where s^* jumps, one can select an unstable equilibrium or any strategy within a range.)

Here, the selected equilibrium s^* can depend on the distributions of the noise terms $(\varepsilon_1, \ldots, \varepsilon_n)$ (see Frankel, Morris and Pauzner (2003)). One can find conditions under which the selected equilibrium does not depend on the distribution of the noise. The seminal work of Carlsson and van Damme (1993) provides an example of noise-independent equilibrium selection for 2×2 games.

6.4 Risk-dominant Equilibrium Selection in 2x2 Games

For general 2×2 games and general type spaces with additive noise, Carlsson and Van Damme (1993) show that risk-dominant action is uniquely rationalizable for small noise, provided that the game can be connected to dominance region (as it will be clearer momentarily). As such they provide a noise-independent equilibrium selection—by risk dominance.

Consider a 2-player, 2-action Bayesian game $\hat{G}^{\sigma}$, indexed by the shock size σ:

- the set of players is $N = \{1, 2\}$;
- the set of payoff parameters is an open set $\Theta \subseteq \mathbb{R}^m$;
- for each player i, the set of actions is $A_i = \{a, b\}$;
- the payoff function $u_i : A \times \Theta \to \mathbb{R}$ is continuously differentiable with bounded partial derivatives with respect to θ on Θ;
- each player observes a signal

  $$x_i = \theta + \sigma \varepsilon_i$$

  where θ has a continuously differentiable and bounded density that is strictly positive on Θ; the noise terms $(\varepsilon_1, \varepsilon_2)$ are bounded (taking values in $[-1, 1]^m$), stochastically independent from θ, and have a continuous joint density.

In comparison to the monotone supermodular game G^{σ} in the previous section, this model allows more general types and payoff function at the expense of restricting the attention to 2 player and 2 action case. It allows multidimensional payoff parameters and type spaces, by allowing both θ and x_i to be m-dimensional for any $m \geq 1$. It allows players' signals to be correlated conditional on θ, although it keeps assuming that they are independent from θ. It rules out the problematic cases, such as perfect correlation, by assuming that they have a joint density. In terms of payoffs, the above definition does not impose any non-technical restriction; the result will assume dominance regions and some form of supermodularity in the relevant region (implicitly) as I will explain below. Towards stating the result, recall that (a, a) is *risk-dominant equilibrium of the complete information game at* θ if (a, a) is (p_1, p_2)-dominant at the complete information game with θ for some $p_1 + p_2 < 1$, which is equivalent to

$$g_1^{a,\theta} \cdot g_2^{a,\theta} > g_1^{b,\theta} \cdot g_2^{b,\theta}$$

when there are multiple equilibria where

$$g_1^{a,\theta} = u_1(a, a, \theta) - u_1(b, a, \theta) \text{ and } g_2^{a,\theta} = u_2(a, a, \theta) - u_2(a, b, \theta)$$

and the players' loss $g_i^{a,\theta}$ from equilibrium (b, b) is defined similarly. Carlsson and van Damme (1993) show that risk dominant equilibrium is selected at θ by the incomplete information game $\hat{G}^\sigma$ with small σ whenever θ is connected to a dominance region without changing the risk dominant action:

Theorem 6.2 *Consider any* $x_i \in C \subset \Theta$ *for some continuous curve* C *such that*

1. *a is strictly dominant for bath players at some* $\theta \in C$*, and*

2. (a, a) *is a risk-dominant equilibrium of the complete information game at each* $\theta \in C$*.*

Then, there exists $\bar{\sigma} > 0$ *such that a is uniquely (interim correlated) rationalizable for type* x_i *in every game* G^σ *with* $\sigma < \bar{\sigma}$*.*

That is, if a type x_1 is connected to a dominance regions with a path along which a pair (a_1, a_2) of strategies remain risk dominant equilibrium or a strictly dominant-strategy equilibrium under complete information, then (a_1, a_2) is the unique rationalizable outcome at (x_1, x_1) when the noise term is vanishingly small. For example, in the partnership game above, (a, a) is dominant-strategy equilibrium when $\theta > 1$ and risk-dominant when $1 \geq \theta > 1/2$. Hence, a type $x_i = 3/4$ is connected to dominance region because we can draw a continuous curve from 3/4 to dominance region as above, where the curve is simply the interval $\left[3/4, \hat{\theta}\right]$ for some $\hat{\theta} \in \Theta$ with $\hat{\theta} > 1$. Then, the above theorem concludes that a is uniquely rationalizable for type $x_i = 3/4$. In the partnership game, except for the knife-edge case $x_i = 1/2$, all types are connected to a dominance region as above, yielding unique rationalizable action at all such cases. In higher dimensional spaces, while the dominant-strategy equilibrium remains to be risk dominant near the dominance regions, that equilibrium can be risk dominant at regions that are not connected to the dominance region, in that one cannot draw a curve from that region to the dominance region without getting out. The theorem predicts risk dominant selection in the regions near the dominance region and is silent for disconnected regions. Intuitively, contagion from the dominance region spreads so long as the strategy profile is risk dominant. This covers the regions connected to the dominance regions. Such a contagion is contained in the region where the original equilibrium remains risk dominant and therefore cannot spread to disconnected regions.

In the above formulation, actions are labeled symmetrically for clarity, but of course the result applies to asymmetric games. For example, consider the following variation of hawk-dove game

	Hawk	*Dove*
Hawk	$\theta_1 - 1, \theta_2 - 1$	$\theta_1, 0$
Dove	$0, \theta_2$	v, v

(6.8)

where $0 \leq v < 1$ is fixed; the states are pairs $(\theta_1, \theta_2) \in \mathbb{R}^2$. When each $\theta_i \in (v, 1)$, there are two asymmetric Nash equilibria, $(Hawk, Dove)$ and $(Dove, Hawk)$, and a symmetric Nash equilibrium in mixed strategies. When $\theta_i > 1$, Hawk is dominant, and when $\theta_i < 0$, Dove is dominant. Hence, $(Hawk, Dove)$ is a strictly dominant strategy equilibrium on

$$D_1 = \{(\theta_1, \theta_2) \, | \theta_1 > 1, \theta_2 < 0\} .$$

Moreover, $(Hawk, Dove)$ is a risk-dominant equilibrium on

$$R_1 = \{(\theta_1, \theta_2) \, | \theta_1 \geq v, \theta_2 \leq 1, \theta_1 > \theta_2\} .$$

Clearly, pairs in R_1 are connected to the dominance region D_1. Hence, in the region $M = (v, 1)^2$ with multiple equilibria, $(Hawk, Dove)$ equilibrium is selected when $\theta_1 > \theta_2$, and $(Dove, Hawk)$ is selected when $\theta_1 < \theta_2$.

Although Theorem 6.2 does not make a supermodularity assumption explicitly, it is somewhat implied by the assumption that a fixed strategy pair remains an equilibrium throughout the relevant region. A strategically equivalent game will be supermodular under an order on actions throughout that region.

6.5 Applications with Continuum of Players

Applications of global games often assume a continuum of players and binary actions. This section is devoted to the analysis of such commonly used models, including the currency attack problem studied by Morris and Shin (1998), which popularized the global games.

There is a continuum of players $i \in N = [0,1]$. Simultaneously, each player i selects an action $a_i \in \{a, b\}$. His payoff from playing action b is normalized to zero, while his payoff from a is

$$U(\alpha, \theta)$$

where α is the fraction of players who play action a and $\theta \in \mathbb{R}$ is the state. Moreover, U is weakly increasing in a and in θ. (In many applications, such as the currency attack problem, U is weakly decreasing in θ in the natural order—and hence we will be using the reverse order.) Once again, the players do not know θ but observe a noisy private signal $x_i = \theta + \sigma\varepsilon_i$ where the noise terms are i.i.d. Write F and f for the distribution and density functions of the noise terms, respectively, and G and g for the distribution and the density functions of θ, respectively, where f is assumed to be symmetric around zero. Action a is dominant when θ is above a cutoff $\bar{\theta}$ and action b is strictly dominant when θ is below a cutoff $\underline{\theta}$. Hence in the extremal equilibria, players play action a above a cutoff $\hat{x}$ and b below the cutoff $\hat{x}$ where $\underline{\theta} < \hat{x} < \bar{\theta}$. Note that, given any θ, the fraction of players who play a is

$$\alpha(\theta) = 1 - F((\hat{x} - \theta)/\sigma) = F((\theta - \hat{x})/\sigma).$$

Hence, the cutoff is determined by the indifference condition

$$\int U(F((\theta - \hat{x})/\sigma), \theta)\, dG(\theta|\hat{x}) = 0 \tag{6.9}$$

where $G(\cdot|\hat{x})$ is the conditional distribution of θ given $\hat{x}$. In two common cases the indifference condition simplifies dramatically:

Linear Payoffs The payoff function U is linear, and is normalized as

$$U(\alpha, \theta) = \theta + \alpha - 1, \tag{6.10}$$

so that $\underline{\theta} = 0$ and $\bar{\theta} = 1$. In that case, the indifference condition above reduces to

$$R(\hat{x}) = E[\theta|\hat{x}]$$

as in the partnership game above. The analysis of the linear model reduces to the two-player game studied extensively in Section 6.2.

Games of Regime Change There are two regimes: status quo, and a new regime. The regime changes to a new regime if the fraction α of players who play a ("attack") exceeds some threshold $\bar{\alpha}(\theta)$ where $\bar{\alpha}$ is a decreasing, continuous function; the status quo prevails otherwise. The payoff function is as in the following table

	Regime Change $(\alpha > \bar{\alpha}(\theta))$	Status quo $(\alpha \leq \bar{\alpha}(\theta))$
a	$V(\theta) - C(\theta)$	$-C(\theta)$
b	0	0

where V is the benefit of a regime change for an attacker, weakly increasing in θ, positive, and Lipschitz continuous, and C is the cost of attacking, weakly decreasing and Lipschitz continuous. A typical case is $V(\theta) = 1$ and $C(\theta) = c \in (0,1)$. Assume that there exists $\bar{\theta}$ such that $\bar{\alpha}(\theta) < 0$ and $V(\theta) > C(\theta)$ for $\theta > \bar{\theta}$, and there exists $\underline{\theta}$ such that $\bar{\alpha}(\underline{\theta}) > 1$ and $C(\underline{\theta}) > 0$. Thus, a is strictly dominant for $\theta > \bar{\theta}$, and b is strictly dominant for $\theta < \underline{\theta}$. Assume that $\bar{\alpha}$ is strictly decreasing on $[\underline{\theta}, \bar{\theta}]$.

Note that, in an equilibrium with cutoff $\hat{x}$, there exists a unique solution $\hat{\theta}$ to

$$\bar{\alpha}\left(\hat{\theta}\right) = \alpha\left(\hat{\theta}\right) \equiv F\left(\left(\hat{\theta} - \hat{x}\right)/\sigma\right). \tag{6.11}$$

The regime changes if and only if $\theta > \hat{\theta}$, because both the threshold $\bar{\alpha}$ and the attack size α are deterministic functions of θ. Thus, the indifference condition reduces to

$$\int_{\hat{\theta}}^{\infty} V(\theta)\, dG(\theta|\hat{x}) = E[C|\hat{x}]. \tag{6.12}$$

In the limit $\sigma \to 0$, the condition further simplifies as follows.

Proposition 6.5 *In a game of regime change, assume the noise terms are bounded, and g is continuous. Then, for any $x_i \neq \hat{x}$, there exists a unique rationalizable action for* *sufficiently small σ where the unique rationalizable action is*

$$s_i^*(x_i) = \begin{cases} a & \text{if } x_i > \hat{x}, \\ b & \text{if } x_i < \hat{x} \end{cases}$$

and the cutoff $\hat{x}$ is the unique solution to

$$V(\hat{x})(1 - \bar{\alpha}(\hat{x})) = C(\hat{x}). \tag{6.13}$$

Proof. Clearly, as $\sigma \to 0$, $E[C|\hat{x}] \to C(\hat{x})$. On the other side, since V is Lipschitz continuous and the error terms are bounded, $V(\theta)$ is uniformly within a small neighborhood of $V(\hat{x})$ for sufficiently small σ. Hence,

$$\begin{aligned} \lim_{\sigma\to 0} \int_{\hat{\theta}}^{\infty} V(\theta)\, dG(\theta|\hat{x}) &= V(\hat{x}) \lim_{\sigma\to 0} \Pr\left(\theta \geq \hat{\theta}|\hat{x}\right) \\ &= V(\hat{x}) \lim_{\sigma\to 0} \Pr\left(\varepsilon_i \leq \left(\hat{x} - \hat{\theta}\right)/\sigma | x_i = \hat{x}\right) \\ &= V(\hat{x}) \lim_{\sigma\to 0} F\left(\left(\hat{x} - \hat{\theta}\right)/\sigma\right), \text{ because } \theta \text{ is approximately uniform}^3 \\ &= V(\hat{x}) \left(1 - \lim_{\sigma\to 0} \bar{\alpha}\left(\hat{\theta}\right)\right), \text{ by (6.11)} \\ &= V(\hat{x})(1 - \bar{\alpha}(\hat{x})). \end{aligned}$$

Since the left-hand side is increasing and the right-hand side is decreasing there is a unique solution to (6.13). By upperhemicontinuity, the solutions for small σ are nearby the unique solution $\hat{x}$ to the limit case. ■

Proposition 6.5 establishes a noise-independent selection for global games of regime change. The selection criterion, identified by the indifference condition (6.13), is a generalization of risk dominance to continuum of players. Suppose that the fraction of α of players who take action a is uniformly distributed; such a belief is called *Laplacian*. This is quite different from other players randomizing between a and b uniformly, which would lead to $\alpha = 1/2$ with probability 1. Now, under the Laplacian belief, at $\theta = \hat{x}$, the probability of a regime change is $1 - \bar{\alpha}(\hat{x})$, when a player who plays a gets $V(\hat{x})$. Of course, he also incurs a cost of $C(\hat{x})$ regardless of a regime change. Hence, (6.13) states that he is indifferent under the Laplacian beliefs. At $x_i \neq \hat{x}$, he is not indifferent, and selects the strict best response to the Laplacian belief, showing that "risk-dominant equilibrium" is selected, where "risk dominance" is defined as playing a best response to Laplacian belief under continuum of players.

A couple of intuitive comparative statics immediately follow, thanks to the supermodular structure. First, if regime change becomes more beneficial, in that the function V gets weakly higher everywhere, then by (6.13) $\hat{x}$ decreases, inducing more types to attack. This is because $V(\hat{x})(1-\bar{\alpha}(\hat{x}))-C(\hat{x})$ is increasing in $\hat{x}$. Similarly, when changing regime becomes more difficult—in that either the cost function C or the threshold function $\bar{\alpha}$ become higher everywhere, then $\hat{x}$ increases, inducing fewer types to attack. A special example is the currency attack problem.

Currency Attacks A government has pegged the exchange rate for local currency at some e^*, in terms of a foreign currency such as the US dollars. If the government gives up and floats the currency, the exchange rate would be $\tilde{f}(\theta)$ for some increasing function $\tilde{f}$ where $\theta \in [0,1]$ measures the strength of the local currency. Government knows θ, and $\tilde{f}(1) < e^*$, i.e., the local currency will be devaluated at all states. The players are a continuum of speculators, who do not know θ and observe noisy private signals as above. A speculator can attack the local currency by borrowing a unit in local currency at cost t and buy the foreign currency at exchange rate e^*. After each speculator decides whether attack, observing the fraction α of speculators who attack (and observing θ), the government decides whether to keep the peg (by absorbing the attack through open market operation) or float the currency. For the government, the value of peg is $v > 0$ and the cost of keeping the peg is $c(\alpha,\theta)$ where c is strictly increasing in α and strictly decreasing in θ. The government keeps the peg if $v > c(\alpha,\theta)$, and floats it otherwise. Assume $c(0,0) > v$, so that the government floats the currency even without any attack when $\theta = 0$. Thus, there exists $\bar{\alpha}(\theta)$, defined by $v = c(\bar{\alpha}(\theta),\theta)$ when in the interior, such that government floats the currency if and only if $\alpha > \bar{\alpha}(\theta)$ where $\bar{\alpha}$ is strictly increasing, continuous, and $\bar{\alpha}(0) < 0$. Given the government's behavior, which is fixed mechanically, the payoff of a speculator from attacking is $e^* - \tilde{f}(\theta) - t$ if the currency is floated and $-t$ otherwise; the payoff from not attacking is zero. Assume also that $\tilde{f}(0) < e^* - t < \tilde{f}(1)$.

Note that the incentive to attack is decreasing with θ, and hence we use the reverse order on θ and x to apply the general results above. Observe that $V(\theta) = e^* - \tilde{f}(\theta)$ and $C(\theta) = -t$, satisfying the assumptions above (with the reverse order). Since $e^* - t < \tilde{f}(1)$, we have $V(1) < C(1)$, and hence not attack is strictly dominant near $\theta = 1$, and attack is strictly dominant near $\theta = 0$ (where $V(0) > C(0)$ and $\bar{\alpha}(0) < 0$). Thus, by Proposition 6.5, for small σ, the players attack if and only if $x_i < \hat{x}$ where

$$\left(e^* - \tilde{f}(\hat{x})\right)(1 - \bar{\alpha}(\hat{x})) = t.$$

Using the general comparative statics above, one can conclude that the size of attack at a given θ is increasing in the peg level e^*, decreasing in the transaction coast t of attacking, increasing in the coast function c of defending the peg, and decreasing in the function $\tilde{f}$—when the functions are ordered by $f \geq g \iff f(x) \geq g(x)$ for all x.

This has some interesting policy implications. For example, suppose that the government sets e^* before observing θ and the value $v(e)$ of peg level e is increasing in e. Then its payoff is

$$U(e) = \int_{\hat{x}(e)}^{1} (v(e) - c(\alpha(x, e), x))\, dG(x)$$

where G is the distribution of θ. Now, the government faces the following trade off. On the one hand, higher e has higher value. On the other hand, higher e leads to higher attack size $\alpha(x, e)$ resulting in higher cost of defending when the attack fails and also resulting in less stable currency as the government floats at a lower level $\hat{x}(e)$. Thus, the government typically chooses an interior exchange rate e^*. Applying Topkis's monotonicity theorem, one can further conclude that the optimal exchange rate is increasing in the government's belief G in the strength of local currency as well as the value v of the peg (as a function), while decreasing in the cost c of defending the currency (as a function once again).

Now, since higher G leads to higher exchange rate e^*, the speculator can learn about the government belief about the strength of local currency from the exchange rate e^*. If the government does have private information about θ, then this leads to a signaling game between the government and the speculators, where the government does no longer choose the optimal e^* naively (see Angeletos, Hellwig, and Pavan (2006)).

6.6 Dynamic Global Games

In dynamic games, incomplete information can affect the strategies in many distinct ways. For example, a central bank may set the target exchange rate high in order to signal the speculators that the fundamentals of the local currency are strong. Likewise, an employer may want to screen the workers with low incentive to work. Having seen a currency survived a possible attack in the first day, the speculators may conclude that the fundamentals are strong enough to rule out the possibility that attack is a dominant strategy. Such effects often lead to multiple equilibria. In dynamic games, long-run incentives may overwhelm short-run incentives leading to excessive number of equilibria, as in folk theorems. This section is devoted to exploring such well-understood affects as well as some novel effects.

6.6.1 Dynamic Contagion

An important, novel effect arises when the fundamentals are stochastic. When the fundamentals are in a dominance region, the players may find playing the dominant action as the only best response. Then, even outside of the dominance regions, when the state is near a dominance region, players may anticipate that the state will enter the dominance region very soon, playing the associated action as the only best response—although it is no longer dominant. As in the static global games this leads to a contagion of a unique best response, this time through dynamic concerns rather than incomplete information. As shown by Burdzy, Frankel, and Pauzner (2001) this may lead to risk dominant selection. I will next present their result on the partnership game; their result applies to general 2×2 supermodular games and allows the fundamentals to have drift.

Formally, consider a continuum $[0, 1]$ of players i. Time is discrete, $t = 0, \tau, 2\tau, 3\tau, \ldots$ for some small but positive τ. At each t a subset of players are selected and randomly matched to play a 2×2 game

	a	b
a	θ_t, θ_t	$\theta_t - 1, 0$
b	$0, \theta_t - 1$	$0, 0$

where the state θ_t follows a random walk: $\theta_t = \theta_{t-\tau} + \sigma\sqrt{\tau}$ with probability 1/2 and $\theta_t = \theta_{t-\tau} - \sigma\sqrt{\tau}$ with the remaining probability for some small $\sigma > 0$. The state θ_t is publicly observed at time t. The probability that a player i is selected to play at time t is $m\tau$, and he is equally likely to be matched to any other player.

The key element of the model is that the actions are somewhat sticky. At each t, before θ_t is realized, each player i gets a chance—with probability $k\tau$—to revise his action, in which case he can select any action $a_{it} \in \{a, b\}$; otherwise $a_{it} = a_{i(t-\tau)}$. Everybody selects a new action a_{i0} at 0. If selected, player i plays a_{it} at time t.

Write α_t for the fraction of players j with $a_{jt} = a$. At each time t, the history of past states $(\theta_0, \ldots, \theta_{t-\tau})$ and aggregate actions $(\alpha_0, \ldots, \alpha_{t-\tau})$ is publicly known; each player also can remember his own and previous partners actions privately, but the latter feature is not relevant because of continuum of players. A players' payoff is the discounted sum of his payoffs from the stage games:

$$\sum_t e^{-rt} u_i(a_{it}, a_{jt})$$

where the sum is taken over dates at which i is selected, a_{jt} is the action of his match at t, and $r > 0$ is known. Now, if a player gets to revise his action at t, his belief about the other players' future play is independent of his own past actions because only a negligible set of players have seen them. Hence, his optimal action is independent of his past actions—except for the cases of indifference. Hence, the expected payoff difference between having $a_{it} = a$ and $a_{it} = b$ at t is

$$\begin{aligned} \Delta u(\theta_{t-\tau}, \alpha_{t-\tau}) &= \sum_{n=0}^{\infty} e^{-rn\tau} m\tau (1 - k\tau)^n \left(E[\theta_{t+n\tau} | \theta_{t-\tau}] - E[1 - \alpha_{t+n\tau} | \theta_{t-\tau}, \alpha_{t-\tau}] \right) \\ &= \sum_{n=0}^{\infty} e^{-rn\tau} m\tau (1 - k\tau)^n \left(\theta_{t-\tau} - E[1 - \alpha_{t+n\tau} | \theta_{t-\tau}, \alpha_{t-\tau}] \right) \end{aligned}$$

where expectation is taken according to his belief about how people will revise their future actions. To see the formula, observe that the players' current action is relevant for time $t + n\tau$ only if he does not get a chance to revise his action until then, and this happens with probability time $(1 - k\tau)^n$, and hence his effective discount factor is $\delta = (1 - k\tau) e^{-rn}$. On that event, he will be selected with probability $m\tau$, and his payoff will be $\theta_{t+n\tau}$ if his match plays a, with probability $\alpha_{t+n\tau}$, and $\theta_{t+n\tau} - 1$ otherwise. The above formula is obtained by taking expectation conditional on $(\theta_{t-\tau}, \alpha_{t-\tau})$. The player plays a if $\Delta u(\theta_{t-\tau}, \alpha_{t-\tau})$ is positive and plays b if it is negative.

The solution concept is *iterated conditional dominance*. At any given history, for any player who gets to revise his action, we eliminate action b for that player at that history if playing a yields higher payoff at all beliefs, this is the case when $\Delta u(\theta_{t-\tau}, \alpha_{t-\tau}) > 0$ for all beliefs about $\{\alpha_t, \alpha_{t+\tau}, \ldots\}$. Likewise, we eliminate a a player at a history if $\Delta u(\theta_{t-\tau}, \alpha_{t-\tau}) < 0$ for all his beliefs about $\{\alpha_t, \alpha_{t+\tau}, \ldots\}$. We iterate this elimination procedure for all histories and players indefinitely. Clearly, only $(\theta_{t-\tau}, \alpha_{t-\tau})$ is relevant for this procedure.

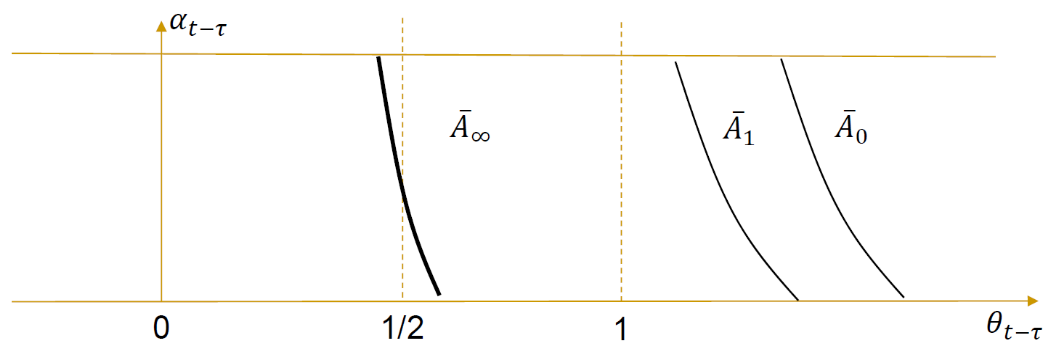

Figure 6.9: Dynamic Contagion

The main result applies for the continuous-time limit for a fixed but sufficiently large k (i.e. for sufficiently small stickiness in strategies):

Theorem 6.3 *For any $\theta_{t-\tau} \neq 1/2$, there exist $\bar{k} < \infty$ and $\bar{\tau} : \mathbb{N} \to (0,1)$ such that, when $k > \bar{k}$ and $\tau < \bar{\tau}(k)$, a unique action $s^*(\theta_{t-\tau})$ survives the iterated conditional dominance at any $(\theta_{t-\tau}, \alpha_{t-\tau})$ where*

$$s^*(\theta_{t-\tau}) = \begin{cases} a & \text{if } \theta_{t-\tau} > 1/2, \\ b & \text{if } \theta_{t-\tau} < 1/2. \end{cases}$$

This is quite remarkable: there is a unique solution in a repeated/stochastic game with arbitrarily patient forward looking players. Here, the long-run incentives that lead to a folk theorem in repeated games are muted because a player's current action has a negligible impact on the future play of other players as it can be observed only by a negligible set of players. His action affects only his own future actions, due to stickiness, and that friction is assumed to arbitrarily small, although it is that friction that leads to the unique solution. Since $\tau \to 0$, the time preferences do not play a role and one can take $e^{-rn\tau} = 1$.

The main insight is a contagion argument, induced by backward induction in a stochastic environment. Consider a history $(\theta_{t-\tau}, \alpha_{t-\tau})$ with large $\theta_{t-\tau} \gg 1$. Consider a player i who gets to revise his action at t. If he chooses $a_{it} = a$, he will get a positive profit at the current period, but there is a chance that he is stuck with action a in the future and the state moves below 1 while other players switching to b, yielding a negative payoff in the future. But when k is sufficiently large, he will get to change his action with high probability before the state goes below 1, and he should choose a. This yields a set $\bar{A}_0$ of states at which a is dominant as in Figure 6.9, assuming that the other players will switch to b whenever they can. (The figure is plotted for general payoff functions; the boundary of A_0 lies outside of the dominance region in the specific game here.) Now in the second round one gets a larger set $\bar{A}_1$ of states at which a is dominant assuming that the other players choose a on $\bar{A}_0$ and b outside of $\bar{A}_0$ when they revise. Iterating this argument, in the limit, one gets a set $\bar{A}_\infty$ of states at which only a survives iterated conditional dominants where a is the best response to the belief that the others will choose a on $\bar{A}_\infty$ and b outside of $\bar{A}_\infty$ when they revise. Note however that, when k is very large, the boundary of $\bar{A}_\infty$ cannot depend on $\alpha_{t-\tau}$ much, i.e., the curves in Figure 6.9 are almost vertical lines. Indeed, when k is very large, regardless of $\alpha_{t-\tau}$, almost everybody gets to revise his action very quickly. Moreover, the revised action is a on $\bar{A}_\infty$, and in the worst-case scenario it is b outside of $\bar{A}_\infty$. This implies that $\bar{A}_\infty$ consists of the states at which a is risk dominant as follows. When the state is on the boundary of $\bar{A}_\infty$, the players think that with probability $1/2$, θ_t will move up and everybody will choose a, and with the remaining probability, the state will move down and everybody will choose b. In order to be indifferent, it must be that $\theta_{t-\tau} = 1/2$. More generally, the players must play the risk dominant action at every state.

Proof. For any set $A \subset \mathbb{R} \times [0,1]$, define

$$\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau} | A\right) = \sum_{n=0}^{\infty} e^{-rn\tau} m\tau \left(1 - k\tau\right)^n \left(\theta_{t-\tau} - 1 + E\left[\alpha_{t+n\tau} | \theta_{t-\tau}, \alpha_{t-\tau}\right]\right)$$

as the payoff difference between choosing a and b at history $(\theta_{t-\tau}, \alpha_{t-\tau})$ assuming that the other players choose $s^*\left(\theta_{t'-\tau}, \alpha_{t'-\tau} | A\right)$ whenever they get to revise their action at any future date t' where

$$s^*\left(\theta_{t'-\tau}, \alpha_{t'-\tau} A\right) = \begin{cases} a & \text{if } \left(\theta_{t'-\tau}, \alpha_{t'-\tau}\right) \in A, \\ b & \text{otherwise.} \end{cases}$$

Clearly, $\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau} | A\right)$ is increasing in all arguments $(\theta_{t-\tau}, \alpha_{t-\tau}, A)$. Action a conditionally dominates b if $\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau} | \varnothing\right) > 0$. Observe that the set

$$\bar{A}_1 = \left\{\left(\theta_{t-\tau}, \alpha_{t-\tau}\right) | \Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau} | \varnothing\right) > 0\right\}$$

of states at which a conditionally dominates b is non-empty. In particular, it contains any $(\theta_{t-\tau}, \alpha_{t-\tau})$ with $\theta_{t-\tau} > 1$:

$$\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau}|\varnothing\right) \geq \Delta u\left(\theta_{t-\tau}, 0|\varnothing\right) = \textstyle\sum_{n=0}^{\infty} e^{-rn\tau} m\tau\left(1-k\tau\right)^{n}\left(\theta_{t-\tau}-1\right) > 0.$$

Writing $\bar{A}_l$ for the set of states at which a is the unique action that survives l rounds of conditional dominance, observe that

$$\bar{A}_l = \left\{\left(\theta_{t-\tau}, \alpha_{t-\tau}\right)|\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau}|\bar{A}_{l-1}\right) > 0\right\},$$

and

$$\bar{A}_1 \subset \bar{A}_2 \subset \cdots \subset \bar{A}_l \subset \cdots.$$

The limit set

$$\bar{A}_\infty = \textstyle\bigcup_l \bar{A}_l$$

is a fixed point:

$$\bar{A}_\infty = \left\{\left(\theta_{t-\tau}, \alpha_{t-\tau}\right)|\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau}|\bar{A}_{\infty}\right) > 0\right\}.$$

Since b conditionally dominates a whenever $\theta_{t-\tau} < 0$, $\bar{A}_\infty$ contains only non-negative states $\theta_{t-\tau}$, and in particular it has a lower boundary $\partial\bar{A}_\infty$ defined by the indifference condition

$$\Delta u\left(\theta_{t-\tau}, \alpha_{t-\tau}|\bar{A}_{\infty}\right) = 0.$$

Note however that when $(\theta_{t-\tau}, \alpha_{t-\tau}) \in \bar{A}_\infty$,

$$E\left[\alpha_t|\theta_{t-\tau}, \alpha_{t-\tau}\right] = \alpha_{t-\tau} + \left(1-\alpha_{t-\tau}\right)k\tau = \left(1-k\tau\right)\alpha_{t-\tau} + k\tau.$$

Hence, if it is known that $(\theta_{t-\tau}, \alpha_{t-\tau}), \ldots, (\theta_{t+n\tau-\tau}, \alpha_{t+n\tau-\tau}) \in \bar{A}_\infty$, then

$$E\left[\alpha_{t+n\tau}|\theta_{t-\tau}, \alpha_{t-\tau}\right] = 1 + \left(1-k\tau\right)^{n+1}\left(\alpha_{t-\tau}-1\right),$$

which converges to 1 at an exponential rate. Burdzy, Frankel, and Pauzner (2001) show that for large k and small τ, if $(\theta_{t-\tau}, \alpha_{t-\tau}) \in \bar{A}_\infty$, the state remains in $\bar{A}_\infty$ for arbitrarily many dates and $\alpha_{t'}$ approaches 1 arbitrarily quickly in real time. Similarly, under $s^*(\theta_{t'-\tau}, \alpha_{t'-\tau}A)$, if $(\theta_{t-\tau}, \alpha_{t-\tau}) \notin \bar{A}_\infty \cup \partial\bar{A}_\infty$, $\alpha_{t'}$ approaches 0 arbitrarily quickly in real time. Hence, on the boundary, with probability 1/2, the state moves in $\bar{A}_\infty$ and $\alpha_{t'}$ approaches 1 in the near future; and state moves outside of $\bar{A}_\infty$ and $\alpha_{t'}$ approaches 0 with probability 1/2. Indifference condition then implies that $\theta_{t-\tau}$ is approximately equal to 1/2. ■

6.6.2 Dynamic Games of Regime Change

6.6.3 Dynamics under Model Uncertainty

6.6.4 Supermodularity with Dynamics

6.7 Concluding Remarks

6.8 Exercises

Exercise 6.4 *Consider the following differentiated Bertrand oligopoly with demand uncertainty. Each firm i can choose a price p_i from a set $[0, \bar{p}]$ for some $\bar{p} > 0$ and has zero marginal cost. The demand for the product of firm i is $Q_i(\theta, p)$ where Q_i is a differentiable function, decreasing in p_i and increasing in (θ, p_{-i}), and θ is a real-valued unknown parameter, common to all firms. Each player i observes a signal*

$$x_i = \theta + \varepsilon \eta_i$$

where $(\theta, \eta_1, \ldots, \eta_n)$ are independently distributed with atomless densities and η is bounded.

1. *Find sufficient conditions under which there is a unique rationalizable strategy profile p^* in the limit $\varepsilon \to 0$—as in Frankel, Morris and Pauzner.*

2. *Take $n = 2$ and $Q_i = \theta - ap_i + bp_j$ where $a \geq b > 0$. Check what further parameter restrictions, if any, are needed to ensure the existence of p^* as above. Under these restrictions, compute p^*.*

Exercise 6.5 *Consider the partnership game with normal distributions as in Section 6.2.2, where $\varepsilon_i \sim N(0, 1)$. Now assume instead that $\varepsilon_i \sim N(0, v_i^2)$ where the precision $1/v_i^2$ of player i's information is privately known by player i and v_1 and v_2 are independently distributed. Compute the extremal equilibria. Describe the extremal equilibria when σ and τ are small. (You can make any simplifying assumption about the distribution of variances if it helps.)*

Exercise 6.6 *Consider the following version of the partnership game with model uncertainty in Section 6.2.3. As in Section 6.2.3, $\varepsilon_i \sim N(0, 1)$ while $\eta \sim N(0, \tau^2)$ for some* *unknown variance* τ^2. *Now assume that* τ^2 *can be either* τ_H^2 *or* τ_L^2 *with probabilities* p *and* $1-p$, *respectively, for some* $\tau_H^2 \gg \tau_L^2 > 0$. *For simplicity let the payoff from* a *for player* i *be* x_i *if the other player plays* a *and* $x_i - 1$ *if the other player plays* b.

1. *Compute the rank belief function and discuss how the rank belief function varies as* p *varies from 0 to 1, focusing on values near 0 and 1. (For illustration, you may fix* $\tau_H = 5$ *and* $\tau_L = 1$ *and plot the rank belief function for* $p = 0.9, 0.5, 0.1, 10^{-3}, 10^{-6}$.*)*

2. *Characterize the extremal equilibrium cutoffs and briefly discuss how they vary with* p*—building on your answer for the first part.*

3. *Now suppose that there are two periods:* $t = 0$ *and* $t = 1$. *In each period, the distribution of the shocks* η, ε_1, *and* ε_2 *are as above, where the unknown variance* τ^2 *for the common shock is the same in both periods. At the end of period* $t = 0$, *the common shock at period 0 becomes publicly observable but players do not observe each others' actions in that period. Characterize the extremal equilibria and briefly discuss how the equilibrium cutoffs at the last period vary with the realization of the common shock at the initial period.*

Exercise 6.7 *Consider a coordinated investment problem with a continuum of players* $i \in [0,1]$. *Simultaneously, each* i *chooses between Invest and Not Invest. The payoff from Invest is*

$$x_i + \alpha - 1,$$

and the payoff from Not Invest is 0, where x_i *is a productivity parameter for player* i, *who knows* x_i *privately, and* α *is the measure of the players who invest. Assume that*

$$x_i = y + \sigma\eta + \sigma\varepsilon_i$$

where y *and* $\sigma > 0$ *are known parameters and* $(\eta, \varepsilon_i)_{i\in[0,1]}$ *are i.i.d. with* $N(0,1)$.

1. *Argue that there exist extremal Bayesian Nash equilibria and these equilibria are in cutoff strategies. Write* x^* *and* x^{**} *for the smallest and the largest equilibrium cutoffs and characterize these cutoffs; find an equation.*

2. *How do* x^* *and* x^{**} *vary with* σ *and* y.

3. *Find* $\lim_{\sigma\to 0} x^*$ *and* $\lim_{\sigma\to 0} x^{**}$, *and plot* x^* *and* x^{**} *as a function of* y *for small* σ. *Briefly discuss which types have unique rationalizable actions (in the limit* $\sigma \to 0$*).*

Exercise 6.8 *Consider a two-person partnership game. Simultaneously, each player* i *invests* $a_i \in [0,1]$, *and the payoff of player* i *is*

$$u_i(a_1, a_2, \theta) = \theta f(a_1) f(a_2) - c(a_i),$$

where $\theta \geq 0$ *is a parameter, and* f *and* c *are strictly increasing functions with* $f(0) > 0$. *Assume that* θ *is common knowledge.*

1. *Show that the game is supermodular.*

2. *Assuming that best-reply correspondence is convex-valued and continuous, compute all rationalizable strategies.*

3. *Show that the minimum and the maximum rationalizable strategies as well as minimum and maximum equilibrium strategies are increasing functions of* θ.

Exercise 6.9 *In the previous question, assume that* θ *is not common knowledge. Instead,* θ *is distributed with a continuous probability density function that is strictly positive at each* $\theta \geq 0$, *and each player observes a signal*

$$x_i = \theta + \varepsilon\eta_i,$$

where $\varepsilon \in (0,1)$, (θ, η_1, η_2) *are independent and each* η_i *is bounded and has a continuous density. All of these are common knowledge. Make any other technical assumptions on* f *and* c *under which the statements 1-3 below are true (prove that the statements are true under the assumptions you made):*

1. *All rationalizable strategies converge to a unique strategy,* $S : \mathbb{R}_+ \to [0,1]$.

2. S *is a non-decreasing function of* x_i. *Briefly discuss this in comparison to Part 3 of the previous question.*

3. $(S(x_i), S(x_i))$ *is a Nash equilibrium of the complete-information game in which it is common knowledge that* $\theta = x_i$ *at any* x_i *at which* $\bar{S}(x_i) = \max\{a \mid a \text{ is rationalizable at } x_i\}$ *is continuous.*

4. *Compute S for $f(a) = a + \alpha$ and $c(a) = a^2/2 + \gamma a$ where α and γ are positive.*

Exercise 6.10 *Consider a two-player incomplete information game with types*

$$x_i = \theta + \varepsilon_i$$

where real-valued random variables θ, ε_1, and ε_2 are independently distributed. Each noise ε_i has a symmetric, bounded distribution F on $[-1, 1]$ with density f, and θ has a symmetric distribution G with density g such that there exists $\bar{\theta}$ such that

$$g(\theta) = |\eta|^{-\alpha} \qquad (\forall \theta > \bar{\theta})$$

for some $\alpha > 1$.

1. *Assume F is the uniform distribution on $[-1, 1]$. For $x_i \gg \bar{\theta}$, compute the conditional distributions of θ and x_{-i} and the rank belief $R(x_i)$. What happens as $x_i \to \infty$?*

2. *Show that for every $\epsilon > 0$, there exists $\bar{x} < \infty$ such that*

$$\left| F(\theta, x_{-i}|x_i) - \bar{F}(\theta, x_{-i}|x_i) \right| < \epsilon \qquad \forall x_i > \bar{x}, x_{-i}, \theta$$

   where $F(\cdot|\cdot)$ and $\bar{F}(\cdot|\cdot)$ conditional distributions of (θ, x_{-i}) under (F, G) and $(F, \bar{G})$ where $\bar{G}$ is the improper flat prior on the real line. (Make any technical assumption that you may need to make.)

3. *Bonus: For 2×2 games where payoffs depend on θ, find conditions under which the players must play a risk-dominant action (under rationalizability) whenever they observe a large shock.*

Exercise 6.11 *Members of a political party is voting between two candidates, a and b, in a primary election. There are $2n + 1$ members. Simultaneously, each member votes for one of the two candidates, and the candidate who gets at least $n + 1$ votes wins and goes on to compete against a candidate from another party in the general elections. Each member cares only about which party wins the general election. In particular, she gets 1 if the candidate from her own party wins in the general election and 0 otherwise. If candidate a competes in the general election, the probability of winning for her depends* *on some unknown real-valued parameter* $\theta \in [0,1]$*; the probability is* $\alpha(\theta)$ *for some increasing and continuous function* α *with* $\alpha(0) = 0$ *and* $\alpha(1) = 1$*. If candidates* b *competes in the general election, she wins with probability* β *for some known* $\beta \in (0,1)$*. The parameter* θ *is uniformly distributed on* $[0,1]$*, and each member* i *observes a signal*

$$x_i = \theta + \varepsilon_i$$

with additive noise where $(\varepsilon_1, \ldots, \varepsilon_n)$ *is independent from* θ *and independently and identically distributed with uniform distribution on* $[-\epsilon, \epsilon]$ *for some known small but positive* ϵ*.*

1. *Compute a symmetric Bayesian Nash equilibrium in cutoff strategies, where each member* i *votes for* a *if and only if* $x_i \geq x^*$ *for some known cutoff* x^**. (Write the condition the cutoff* x^* *must satisfy.)*

2. *Compute the limit of* x^* *as* $\epsilon \to 0$ *and briefly discuss the limit.*

3. *Challenge (not for grade): In some political parties there are "super delegates" who may be committed to voting for a candidate (as in the case of the Democratic Party in 2016 presidential elections). Towards an analysis of the effect of these super delegates, assume that candidate* a *wins if and only if* k *members votes for her for some known* $k \in (0, 2n+1)$*. Write a model (e.g. of information, preferences and equilibrium behavior) and analyze the impact of super delegates on the outcome of the general election. (Feel free to relax any of the assumptions I made.)*

Exercise 6.12 *In the currency attack model, assume that the government has a debt* D *denominated in the foreign currency. The value of the peg comes from lowering the debt payment in the local currency. In particular, the value of the peg and the cost of defending the peg are*

$$\begin{aligned} v(\theta) &= D/f(\theta) - D/e^* \\ c(\alpha, \theta) &= \alpha(e^* - f(\theta)), \end{aligned}$$

respectively, where $f(\theta) = \theta$ *is the floating rate,* e^* *is the pegged exchange rate, and* α *is the fraction of speculators who attack.*

1. *Check whether this is a monotone supermodular game.*

2. *Compute the limit of the symmetric monotone Bayesian Nash equilibria as noise converges to 0.*

3. *Determine how each of the (limit) equilibria vary with respect to the transaction cost t, D, and e^*; briefly discuss your results.*

4. *Bonus (open ended, no credit): develop a model that also takes the exports and the imports into account, and determine the optimal exchange rate e^* for the government.*

Exercise 6.13 *Consider a large number of depositors, denoted by $N = [0, 1]$. The bank has 1 unit of funds, accounting for the deposits by the depositors. Each depositor simultaneously chooses between "Withdraw Now" (action 0) and "Keep until maturity" (action 1). If the bank has sufficient funds, it pays $r > 1$ to those who withdraw and $R > r$ to those who keep. But the bank may not have enough funds, in which case it divides its existing funds equally among those who demand their money at the moment as follows. Write w for the fraction of depositors who withdraws. If $w < 1/r$, then the bank pays r to each depositor who withdraws, and invests the remaining funds $1 - rw$ to a project that yields $(1 - rw)\,\theta$ and pays*

$$u_K = \min\left\{\frac{1 - rw}{1 - w}\theta, R\right\}$$

to each depositor who keeps. If $w \geq 1/r$, then the bank pays $1/w$ to each depositor who withdraws, giving 0 to the depositors who keep.

1. *Assume that θ is known. Determine the values of θ for which there is a dominant strategy, and identify the dominant strategy equilibrium for those values. Which strategies can be dominant (briefly discuss your finding)? For the remaining values θ, compute the set of symmetric Nash equilibria in pure strategies.*

2. *Now assume that θ is not known and each depositor i gets a signal $x_i = \theta + \varepsilon_i$ where θ is uniformly distributed on $[0, R]$ and ε_i is an idiosyncratic noise term, uniformly distributed on $[-\epsilon, \epsilon]$ for some small but positive ϵ. Compute the monotone symmetric Bayesian Nash equilibria for the limit $\epsilon \to 0$. Plot the equilibrium cutoffs as a function of r and briefly discuss your findings.*

Hint: Observe—and verify—that w is uniformly distributed on $[0,1]$ according to the cutoff type.

3. *Now imagine that the deposits up to some D is insured. That is, a depositor gets at least* $\min\{D,r\}$ *if she withdraws and at least* $\min\{D,R\}$ *if she keeps. Compute the set of monotone symmetric Bayesian Nash equilibria in pure strategies for the limit $\epsilon \to 0$, and briefly discuss how D affects the equilibrium strategies.*

Exercise 6.14 *This question asks you to analyze hoarding of supplies in times of crises. There are a countinuum of players $i \in [0,1]$ and a consumption good of total size $Q \in (0,1)$. The value of the good for each player i is x_i where*

$$x_i = \theta + \varepsilon_i$$

for some common shock θ with improper uniform prior and idiosyncratic shock ε_i with distribution function F and density function f. Each player privately knows the value of the good for herself. There are two periods, $t = 0, 1$. The price of the good at period $t = 0$ is p_0, a known number, and the price of the good at $t = 1$ is determined by the market-clearing condition (as below). Each player can choose to buy the good at $t = 0$ or at $t = 1$. If she chooses to buy at $t = 0$, and the fraction A of the players who chooses to buy at $t = 0$ is less than or equal to Q, then she buys it at price p_0 obtaining a payoff of $x_i - p_0$. If $A > Q$, then the buyers are rationed so that each of them buy with probability Q/A, obtaining a payoff of $(x_i - p_0) Q/A$; those who do not buy at $t = 0$ obtain a zero payoff in that case. If $A \leq Q$, then the price p_1 of the good at $t = 1$ is such that the fraction of players i with $x_i > p_1$ who have not bought at $t = 0$ is equal to the remaining amount of good, $Q - A$. Compute the set of monotone symmetric Bayesian Nash equilibria in pure strategies. Hint: Define rank r of a player i as the fraction of players j with $x_j > x_i$. Observe that each player believes that her rank is uniformly distributed on $[0,1]$.

Challenge: More generally, compute the symmetric Bayesian Nash equilibria for a given rank belief function R that maps x_i to the rank belief of the type x_i. When θ has fat tails, the rank belief becomes uniform after a large shock; and when θ has light tails, the rank belief becomes concentrated at one of the ends after a large shock. How does a shock affect the hoarding behavior under such two specifications?

Exercise 6.15 *Consider the Currency Attack problem in Section 6.5. Assume that θ and each x_i can take any real value. Suppose that*

$$\begin{aligned} \theta &= y + \sigma\eta \\ x_i &= \theta + \sigma\varepsilon_i \end{aligned}$$

for some known parameters y and σ with $\sigma \in (0,1)$ and some independently distributed random variables η, ε_i (for $i \in [0,1]$) where ε_i and η have densities f and g, respectively:

$$\begin{aligned} f(\varepsilon) &= \min\left\{\bar{f}, |\varepsilon|^{-\alpha}\right\} \\ g(\eta) &= \min\left\{\bar{g}, |\eta|^{-\beta}\right\} \end{aligned}$$

for some known positive numbers $\bar{f}, \bar{g}, \alpha, \beta$. For each specification below compute the limit of the unique equilibrium as $\sigma \to 0$ and briefly discuss your findings by comparing them to each other and to that of Morris and Shin.[4]

1. $\alpha \gg \beta$
2. $\alpha \ll \beta$

Exercise 6.16 *Consider a regime change game in which there is a unit mass of citizens who choose between revolting (action $a_i = 1$) and not revolting (action $a_i = 0$), and there is regime change if the fraction of citizens who revolt exceeds θ. Ex-ante θ is uniformly distributed over a large interval $[-L, L]$, where $L \gg 1$, and each citizen observes a signal $x_i = \theta + \varepsilon_i$ where idiosyncratic noise ε_i is distributed uniformly on some small interval $[-\epsilon, \epsilon]$. The game is so far as in Section 6.5. It differs from that framework as follows. A player receives payoff 0 if they do not revolt. If a player revolts, their payoff is $v_i - c$ if there is regime change and $-c$ otherwise, for some $c > 0$. Here, v_i is private information of player i, and it is independently and identically distributed with*

$$v_i = \begin{cases} v_H & \text{with probability } q \\ v_L & \text{with probability } 1-q \end{cases}$$

for some $q \in (0,1)$. Moreover, $v_H > c$ and $v_L < v_H$, but v_L can be larger or smaller than c.

[4] Hint: Suppose X and Y have power distributions as in this question. Conditional on $X+Y$ being on the tail, Bayes' rule attributes all of the variation to one of the addends (e.g., puts nearly probability 1 on $X = E[X]$ and $Y = X + Y - E[X]$), attributing all of the variation for the random variable with lower index.

1. *Compute a symmetric Bayesian Nash equilibrium.*

2. *Now introduce the government explicitly. Before the citizens move, the government observes* θ*, and can then decide to arrest a fraction of its citizens. These arrested citizens cannot choose to revolt (action* $a = 1$*). The payoff of the government is* $V - C(p)$ *if the regime does not change and* $-C(p)$ *if the regime changes where* p *is the fraction of players who are arrested. Here,* $V > 0$*, and* C *is a convex increasing function with* $C(0) = 0$ *and* $C(1) \gg V$*. For this part, assume that* $v_L < c$*. Answer the following question for two separate cases:*

   Privacy *The government cannot observe* v_i *for any player* i*.*

   No Privacy *The government observes* v_i *for each player* i*.*

   (a) *If the players who are not arrested were not aware of this possibility (and played according to the equilibrium in the first part), what would be the optimal arrest policy for the government?*

   (b) *Briefly discuss the role of privacy.*

Bibliography

[1] * Angeletos, G. M., Hellwig, C., & Pavan, A. (2007). Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. Econometrica, 75(3), 711-756.

[2] Angeletos, G.M. and A. Pavan (2007): "Efficient use of information and social value of information." Econometrica 75.4 1103-1142.

[3] * Carlsson, H. and E. van Damme [1993] "Global Games and Equilibrium Selection" Econometrica 61, 989-1018.

[4] Chassang, S. [2010] "Fear of Miscoordination and the Robustness of Cooperation in Dynamic Global Games with Exit," Econometrica 78, 973 – 1006.

[5] Chen, H., & Suen, W. (2016). "Falling dominoes: A theory of rare events and crisis contagion." American Economic Journal: Microeconomics, 8(1), 228-255.

[6] Frankel, D. and A. Pauzner (2000): "Resolving indeterminacy in dynamic settings: the role of shocks." The Quarterly Journal of Economics 115.1 285-304.

[7] * Frankel, D. M., S. Morris and A. Pauzer (2003) "Equilibrium Selection in Global Games with Strategic Complementarities," Journal of Economic Theory 108, 1-44.

[8] Izmalkov, S. and M. Yildiz (2010). Investor sentiments. American Economic Journal: Microeconomics, 2(1), 21-38.

[9] * Morris, S. and H. S. Shin [1998] "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks" American Economic Review 88, 587-597.

[10] Morris, S. and H. S. Shin "Social Value of Public Information," American Economic Review 92, 1521-1534.

[11] * Morris, S. H.S. Shin, and M. Yildiz (2016): "Common Belief Foundations of Global Games" (with Stephen Morris and Hyun Shin), Journal of Economic Theory, 163, 826-848.

[12] * Morris, S. and M. Yildiz, "Crises: Equilibrium Shifts and Large Shocks."

Chapter 7

Potential Games

In some games, players may act as if each player tries to maximize a common payoff function, in that the payoff differences from unilateral deviations reflect the changes in a common payoff function from such unilateral deviations. Such games are called potential games. An obvious example of a potential game would be a literal common-interest game. However, there are many more potential games—with substantial conflict between the players, such as the Prisoners' Dilemma game, Cournot competition, and many of the coordination games studied in previous chapters. Global games exhibit many useful properties. This chapter is devoted to the introduction of potential games and illustration of some of these proprerties. It mainly follows Monderer and Shapley (1996), who introduced potential games.

7.1 Ordinal Potential Games

This section introduces introduces the weakes notion of a potential games: ordinal potential games.

Definition 7.1 *For a game $G = (N, S, u)$, a function $P : S \to \mathbb{R}$ is said to be an* ordinal potential function *for G if*

$$u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i}) > 0 \iff P(s_i, s_{-i}) - P(s'_i, s_{-i}) > 0 \qquad (\forall i, s_i, s'_i, s_{-i}).$$

A game G is said to be an ordinal potential game *if it admits an ordinal potential function.*

Observe that when S_i is an interval of real numbers and u_i and P are differentiable the above condition can be written as

$$sign\left(\frac{\partial u_i}{\partial s_i}\right) = sign\left(\frac{\partial P}{\partial s_i}\right).$$

The next example establishes that Cournot competition is an ordinal potential function.

Example 7.1 (Cournot Competition) *Consider an n-firm Cournot oligopoly with inverse-demand function Π and constant marginal cost c, where each firm i must produce a positive quantity $q_i > 0$. The payoff function of a firm is*

$$u_i(q_1, \ldots, q_n) = q_i(\Pi(Q) - c)$$

where $Q = q_1 + \cdots + q_n$. Consider the function P with

$$P(q_1, \ldots, q_n) = q_1 \cdot \cdots \cdot q_n (\Pi(Q) - c).$$

Observe that P is an ordinal potential function for Cournot oligopoly. Indeed,

$$\frac{\partial u_i}{\partial q_i} = \Pi(Q) - c - q_i \Pi'(Q).$$

Hence,

$$\begin{aligned} \frac{\partial P}{\partial q_i} &= \prod\nolimits_{j \neq i} q_j (\Pi(Q) - c) - \prod\nolimits_j q_j \Pi'(Q) \\ &= \left(\prod\nolimits_{j \neq i} q_j\right) \frac{\partial u_i}{\partial q_i}. \end{aligned}$$

Ordinal potential games are strategically equivalent to common-interest games under ordinal solution concepts, such as pure strategy Nash equilibrium, inheriting their properties.

Proposition 7.1 *Let $G = (N, S, u)$ be an ordinal potential game with ordinal potential function P. Then, the following are true.*

1. *Pure strategy Nash equilibria of game G coincide with the pure strategy Nash equilibria of game $(N, S, (P, \ldots, P))$.*

2. *Every $s \in \arg\max_{s \in S} P$ is a pure strategy Nash equilibrium of G.*

3. *If S is compact and P is continuous, then G has a pure strategy Nash equilibrium.*

The first part follows immediately from the definitions; Part 2 immediately follows from Part 1, and Part 3 immediately follows from Part 2. Ultimately it establish an exsiting theorem for pure strategy Nash equilibrium.

For general games, adjustment processes, such as best response dynamics, can exhibit cycles and may not converge to any equilibrium point. (They converge eventually get close to the set of rationalizable strategies.) Ordinal potential games behave much better. In particular, define an *improvement path* as a sequence $s^0, s^1, \ldots$ of strategy profiles such that there exists a sequence of players $i_0, i_1, \ldots$ with

- $s^m_{-i_m} = s^{m-1}_{-i_m}$ for each m
- $u_{i_m}(s^m) > u_{i_m}(s^{m-1})$.

If a game G is an ordinal potential game, then the ordinal potential function P increases along any improvement path. If the game is finite, such a path can consists of only finitely many elements, and it must eventually reach to a local maximum of the ordinal potential function (when only unilateral deviations are allowed). Thus, such adjustment processes must eventually converge to a Nash equilibrium. In particular, one can compute a Nash equilibrium (quickly) by using finite improvement paths.

Proposition 7.2 *Let P be an ordinal potential function for a game G. Then, P increases along every improvement path. If G is finite, then every improvement path $s^0, s^1, \ldots$ must be finite: $|s^0, \ldots, s^m| \leq |S|$. Moreover, every maximal improvement path terminates at a pure strategy Nash equilibrium of game G. (That is, for an improvement path $s^0, s^1, \ldots, s^m$ for which there is no improvement path $s^0, s^1, \ldots, s^m, s^{m+1}$, s^m is a pure strategy Nash equilibrium.)*

General solution concepts use also cardinal properties of the utility functions. Hence, in order to study general solution concepts, the next section introduces a stronger concept of potential.

7.2 Potential Games

A potential function P is such that the payoff from unilateral deviations is equal to the change in P with respect to such deviations. Since the preferences are invariant under affine transformation of utility functions, one only needs to find a potential after weighting the payoffs of individual players, as follows.

Definition 7.2 *For a game $G = (N, S, u)$ and $w = (w_1, \ldots, w_n) \in \mathbb{R}^N_+$, a function $P : S \to \mathbb{R}$ is said to be a w-*potential function *for G if*

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}) = w_i(P(s_i, s_{-i}) - P(s_i', s_{-i})) \qquad (\forall i, s_i, s_i', s_{-i}).$$

A game G is said to be a w-potential game if it admits an w-potential function. A function $P : S \to \mathbb{R}$ is said to be a (exact) potential function *for G if*

$$u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}) = P(s_i, s_{-i}) - P(s_i', s_{-i}) \qquad (\forall i, s_i, s_i', s_{-i}).$$

A game G is said to be a potential game if it admits an exact potential function.

Since one can always transform utility functions by dividing each u_i with w_i, the weighted potential games can be viewed as potential games so long as each w_i is positive. Weighted potential games may include dummy players whose payoffs do not depend on their own strategies. Indeed, the payoff function for a w-potential game can be written as

$$u_i(s) = w_i P(s) + d_i(s_{-i}),$$

where $d_i(s_{-i})$ is a "dummy" payoff that does not depend on player's own strategy. We will focus on potential games (for which the weights must be posititive and taken to be 1 for each player). When the strategy sets are intervals in real line and the payoff functions are differentiable, the above condition can be written as

$$\frac{\partial u_i}{\partial s_i} = \frac{\partial P}{\partial s_i} \qquad (\forall i).$$

Observe that when exists, potential function is unique up to a constant:

Proposition 7.3 *If P and $\tilde{P}$ are potential functions for a game G, then*

$$P(s) = \tilde{P}(s) + c \qquad (\forall s)$$

for some constant c.

I will next present some examples of potential games.

Example 7.2 (Investment Game) *Consider the investment game*

	a	b
a	θ, θ	$\theta - 1, 0$
b	$0, \theta - 1$	$0, 0$

(7.1)

This is a potential game with potential function:

	a	b
a	θ	0
b	0	$1 - \theta$

(7.2)

Observe that each pure strategy Nash equilibrium is a "local maximum" when only unilateral deviations are allowed. The potential function has a unique maximizer, and it coincides with risk-dominant equilibrium.

Example 7.3 (Prisoners Dilemma) *Consider the Prisoners' Dilemma game for which a potential function is given on the right:*

	C	D
C	$5, 5$	$0, 6$
D	$6, 0$	$1, 1$

	C	D
C	0	1
D	1	2

(7.3)

The Prisoners' Dilemma game shows that the potential function does not necerrarily reflect the players' common preferences. Although players prefer (C, C) to (D, D), we must have $P(C, C) < (D, D)$.

Of course, every ordinal potential game is a potential game. The next example illustrates that the converse is not true.

Example 7.4 (Modified Prisoners Dilemma) *Consider the following variation of the Prisoners' Dilemma game for which an ordinal potential function is given on the right:*

	C	D
C	$5, 5$	$0, 6$
D	$6, 0$	$1, 2$

	C	D
C	0	1
D	1	2

(7.4)

However, this version is not a potential game. Indeed, any potential function P must satisfy

$$\begin{aligned} P(C,C) - P(D,C) &= -1 \\ P(D,C) - P(D,D) &= -3, \end{aligned}$$

showing that

$$P(C,C) - P(D,D) = -3.$$

On another path, it must satisfy

$$\begin{aligned} P(C,C) - P(D,C) &= -1 \\ P(D,C) - P(D,D) &= -1, \end{aligned}$$

showing that

$$P(C,C) - P(D,D) = -2,$$

a contradiction.

When the inverse-demand function is linear, Cournot oligopoly is a potential function.

Example 7.5 (Linear Cournot Competition) *Consider an n-firm Cournot oligopoly with linear inverse-demand function $\Pi(Q) = a - bQ$ and any marginal cost function c_i. The payoff function of a firm is*

$$u_i(q_1, \ldots, q_n) = q_i(a - bQ) - c_i(q_i)$$

where $Q = q_1 + \cdots + q_n$. Consider the function P with

$$P(q_1, \ldots, q_n) = a \sum_i q_i - b \sum_i q_i^2 - b \sum_{i<j} q_i q_j - \sum_i c_i(q_i).$$

Observe that P is a potential function for Cournot oligopoly. Indeed,

$$\frac{\partial P}{\partial q_i} = a - 2bq_i - bq_j - c_i'(q_i) = \frac{\partial u_i}{\partial q_i}.$$

The next result, due to Monderer and Shapley, provides a useful characterization of potential games in many economic applications, and provides a formula to compute the potential function. (One can vary the potential function by varying the fixed strategy profile $\hat{s}$, at which the constructed potential function is zero. Of course, they vary by a constant.)

Proposition 7.4 *Let G be a game in which the strategy sets are intervals of real numbers. Suppose the payoff functions are twice continuously differentiable. Then, G is a potential game iff*

$$\frac{\partial^2 u_i}{\partial s_i \partial s_j} = \frac{\partial^2 u_j}{\partial s_i \partial s_j} \qquad (\forall i, j \in N).$$

Moreover, if the payoff functions satisfy the above equality, then a potential function is given by

$$P(s) = \sum\nolimits_{i \in N} \int_0^1 \frac{\partial u_i}{\partial s_i}(x(t))\, x_i'(t)\, dt$$

where $x : [0,1] \to S$ is a continuously differentiable path with $x(0) = \hat{s}$ and $x(1) = s$ for some arbitrarily fixed strategy profile $\hat{s} \in S$.

For example, in the Cournot oligopoly with linear inverse demand function, we have

$$\frac{\partial^2 u_i}{\partial q_i \partial q_j} = -b,$$

showing that it is a potential game. When the game inverse-demand function is not linear, we have

$$\frac{\partial^2 u_i}{\partial q_i \partial q_j} = \Pi'(Q) + q_i \Pi''(Q).$$

Hence, whenever $q_i \neq q_j$ and $\Pi''(Q) \neq 0$, we have

$$\frac{\partial^2 u_i}{\partial q_i \partial q_j} \neq \frac{\partial^2 u_j}{\partial q_i \partial q_j}.$$

The Cournot oligopoly is not a potential game whenever the inverse demand function is non-linear (and twice continuously differentiable) although it is an ordinal potential game.

For general games, one can use payoff variations along paths to check whether a game is a potential game as follows. A finite path is a finite sequences $\gamma = (s^0, \ldots, s^m)$ such that there exists a unique sequence $\iota = (i_0, \ldots, i_m)$ such that $s^k_{-i_k} = s^{k-1}_{-i_k}$ for all $k > 0$. A finite path is *closed* if $s^0 = s^m$; it is *simple* if $s^k \neq s^l$ for all other distinct pairs (k, l). For any finite path $\gamma = (s^0, \ldots, s^m)$ define

$$I(\gamma) = \sum\nolimits_{k=1}^{m} u_{i_k}\left(s^k\right) - u_{i_k}\left(s^{k-1}\right).$$

The game is a potential game if the above sum is zero for all closed paths.

Proposition 7.5 *For any game G with finite set of players, the following are equivalent.*

- *G is a potential game.*
- *$I(\gamma) = 0$ for all finite closed paths.*
- *$I(\gamma) = 0$ for all finite, closed, simple paths of length 4.*

The above proposition establishes that it suffices to check violation of a potential function for only on paths with 4 elements, along which only two players vary their strategies. The next section provides an important class of potential games (in fact, they exhoust the set of potential games).

7.3 Congestion Games

A congestion game is a game similar to a strategic situation in which a set of drivers independently choose their routes to commute and have externalities (introduced by Rosenthal (1973), who defined a potential function for them). It turns out that all potential games can be represented as a congestion game. This characterization will also show that many games considered in global games literature are indeed potential games.

Definition 7.3 *A* congestion model *is a tupple $C = (N, M, S, c)$ where*

- *$N = \{1, \dots, n\}$ is the set of players*
- *$M = \{1, \dots, m\}$ is the set of facilities (or resources)*
- *$S = S_1 \times \cdots \times S_n$ is the set of strategy profiles where $S_i = 2^M \setminus \{\emptyset\}$ for each i, so that each $s_i \in S_i$ is the set of facilities used by player i, and*
- *for each $j \in M$, $c_j \in \mathbb{R}^N$, so that $c_j(k)$ is the payoff from using facility j when exactly k players use the facility j (including the player herself).*

Here, the players could be a set of drivers, each facility could be a segment of roads, so that each subset $s_i \in S_i$ is a route for player i. Clearly, a player should get a payoff from a route if the route takes her from her current location to her destination (in addition to the cost associated with them). The next definition absracts away from such practical issues.

Definition 7.4 *A* congestion game *associated with a congestion model* $C = (N, M, S, c)$ *is a game* $G = (N, S, u)$ *where*

$$\begin{aligned} u_i(s) &= \sum\nolimits_{j \in s_i} c_j(k_j(s)) \\ k_j(s) &= \#\{i' \in N | j \in s_{i'}\} \qquad (\forall j \in M). \end{aligned}$$

The following theorem states that every congestion game is a potential game (Rosenthal, 1973), and conversely every potential game can be represented as a congestion game (Monderer and Shapley). Two games $G = (N, S, u)$ and $G' = (N, S', v)$ are said to be *isomorphic* if there exists bijections $g_i : S_i \to S'_i$, $i \in N$, such that for every $i \in N$ and $s \in S$, $u_i(s) = v_i(g(s))$.

Theorem 7.1 *Every congestion game is a potential game, where a potential function is obtained by*

$$P(s) = \sum_{j \in \cup_i s_i} \sum_{l=1}^{k_j(s)} c_j(l).$$

Coversely, every finite potential game G is isomoprhic to a congestion game.

The above result points to an important class of games that we considered in global games chapter. Consider a binary action game with $A_i = \{0, 1\}$, and let the payoff from action $a_i = 1$ be a function $v(\alpha)$ of the number α of players who take action 1 and the payoff from action o be zero. This game can be represented as a congestion game with two facilities where $c_0(k) = 0$ and $c_1(k) = v(k)$. The potential function for such a game is

$$P(s) = \sum_{l=1}^{\alpha(s)} v(l),$$

where $\alpha(s)$ denotes the number of players who take action 1 under strategy profile s. The applictions in the previous chapter mainly focused on games with continuum of players. Sandholm (2000) studies potential games with continuum of players. For such games a potential function can be computed as

$$P(s) = \int_0^{\alpha(s)} v(x)\, dx,$$

where $\alpha(s)$ is the fraction of players who take action 1 under s.

Now imagine that there is strategic complementarity, i.e., the function v is increasing. In that case, the potential function is maximized either at $(0, \ldots, 0)$ or at $(1, \ldots, 1)$. Since $P(0, \ldots, 0) = 0$, the maximizer is determined whether

$$P(1, \ldots, 1) = v(1) + \cdots + v(n)$$

is positive or negative. But observe that $P(1, \ldots, 1) > 0$ and hence the potential function is maximized at $(1, \ldots, 1)$ if and only if 1 is a best response to the Laplacian belief that puts probability $1/n$ on each number l of players taking action 1. That is, potential function is maximized at the equilibrium in which each player plays the risk-dominant action.

7.4 Bayesian Potential Games

Consider a Bayesian game $\mathcal{B} = (N, \Theta, T, A, u, p)$ where $N = \{1, \ldots, n\}$ is the set of players, Θ is the set of fundamental payoff parameters, $T = T_1 \times \cdots \times T_n$ is the set of type profiles, $A = A_1 \times \cdots \times A_n$ is the set of action profiles, $u_i : A \times \Theta \times T \to \mathbb{R}$ is the payoff function of player i, and $p_i(\cdot|t_i) \in \Delta(\Theta \times T_{-i})$ is the interim belief of type t_i. When $\mathcal{B}$ admits a common pror, we will simply write $p \in \Delta(\Theta \times T)$ for the common prior. Assume for simplicity that B is finite. A Bayesian game $\mathcal{B} = (N, \Theta, T, A, u, p)$ is said to be a potential game if there exists a function $P : A \times \Theta \times T \to \mathbb{R}$ such that

$$u_i(a_i, a_{-i}, \theta, t) - u_i(a'_i, a_{-i}, \theta, t) = P(a_i, a_{-i}, \theta, t) - P(a'_i, a_{-i}, \theta, t).$$

Under a common prior, a Bayesian game is a potential game as it is represented as an ex-ante game. Therefore, it inherits the properties of potential games in the previous sections.

Proposition 7.6 *Let $\mathcal{B} = (N, \Theta, T, A, u, p)$ be a Bayesian potential game with potential function $P : A \times \Theta \times T \to \mathbb{R}$ and a common prior p. Then, the ex-ante game $G(\mathcal{B})$ is a potential game with potential function $\mathcal{P} : A^T \to \mathbb{R}$,*

$$\mathcal{P}(s) = E[P(s, \theta, t)]$$

where E is the expectation operator under p. If in addition $p(t_i) > 0$ for each $t_i \in T_i$ and $i \in N$, then each $s \in \arg\max \mathcal{P}(s)$ is a Bayesian Nash equilibrium of game $\mathcal{B}$.

The common prior assumption and ex-ante formulation is important for this result. If the game does not have a common prior, the ex-ante game may not be a potential game even if the Bayesian game admits a potential. Likewise, the interim game can fail to be a potential game (as a player's belief and incentives may depend on her type). Using the above result for the ex-ante game, one can use improvement paths to compute a pure strategy Bayesian Nash equilibrium. Moreover, the adjustment processes would converge to an equilibrium when ex-ante formulation makes sense for the adjustment process.

7.5 Noise-independent Selection in Global Games

As discussed in Section 6.3, in global supermodular games, the game becomes dominance-solvable when players observe the state with vanishingly small additive noise. In the limit, at each state, the players play a Nash equilibrium of the complete information game in which the state is known. Unfortunately, in general, the equilibrium selected depends on the distribution of the additive noise in players' observation. In contrast, in 2 x 2 games studied by Carlsson and van Damme, the risk-dominant equilibrium was selected regradless of the noise distribution. Frankel, Morris, and Pauzner (2003) also study such noise-independent selection. They show that, in Bayesian potential games, the equilibrium that maximizes the potential function will be selected independent of the noise. In particular, in binary action congestion games (widely considered in global games literature), risk-dominant action will be selected independent of the noise distribution.

Formally, as in Section 6.3, consider a Bayesian game G^σ, indexed by the shock size σ:

- the set of players is $N = \{1, \ldots, n\}$;
- the set of payoff parameters is a closed interval $\Theta \subseteq \mathbb{R}$;
- for each player i, the set of actions A_i is a countable union of closed intervals within $[0, 1]$ where $0, 1 \in A_i$;
- the payoff function $u_i : A \times \Theta \to \mathbb{R}$ is continuous with bounded derivatives;

- each player observes a signal

$$x_i = \theta + \sigma\varepsilon_i$$

  where $(\theta, \varepsilon_1, \ldots, \varepsilon_n)$ are stochastically independent with atomless densities, θ has full support, and the noise terms $\varepsilon_1, \ldots, \varepsilon_n$ are bounded.

Under Assumption 6.2, Theorem 6.1 shows that, as $\sigma \to 0$, the extremal equilibria of game G^σ converge to a unique weakly increasing strategy profile s^*, where $s^*(x)$ is a Nash equilibrium of the complete information game in which it is common knowledge that $\theta = x_1$.

Theorem 7.2 *Assume, in addition, that G^σ is a Bayesian potential game with potential function $P : A \times \Theta \to \mathbb{R}$. Then,*

$$s^*(\theta, \ldots, \theta) \in \arg\max_{a \in A} P(a, \theta)$$

at each θ where each $u_i(a_i, a_{-i}, \theta)$ is quasi-concave in a_i.

The last condition requires that the sets of the form $\{a_i | u_i(a_i, a_{-i}, \theta) \geq c\}$ are convex. In that case, $P(a, \theta)$ is maximized at a unique $a^*(\theta) \in A$, which is a strict Nash equilibrium of the complete information game at θ. Theorem establishes that equilibrium $a^*(\theta)$ is selected as the noise vanishes, independent of the distribution of ε_i. This of course implies risk-dominant selection in games discussed above. Frankel, Pauzner, and Morris (2003) establish this for a broader classes of games, requiring only that $a^*(\theta)$ is the unique maximizer of a "Local potential function".

Part II

Advanced Topics

Chapter 8

Interactive Epistemology

This chapter introduces the basic notions in interactive epistemology, mainly focusing on Aumann's partition model of knowledge. It builds on Aumann's classic lecture notes with the same title; see Aumann (1999a,1999b) for a published version. Here, I will add some new concepts and notation that will be used in later chapters. I will first present the standard model of knowledge for one-person case. I will then present the basic concepts of the interactive epistemology, such as common knowledge, common certainty, and public events. I will conclude with Aumann's celebrated Agreeing to Disagree Theorem.

8.1 Standard Model of Knowledge—One-Person Case

In this section I will present the canonical partition model of knowledge, due to Aumann (1974). Of course, epistemology is a vibrant area in philosophy and there are many alternative models of knowledge. The partition model is a canonical model that incorporates several stringent assumptions. These assumptions reflect a highly clear-minded and intelligent individual who knows what she knows and what she does not and can figure out all the logical implications of her knowledge. There are several equivalent representation of knowledge in this model, and each representation is useful in its own way.

Consider a state space Ω. Each state $\omega \in \Omega$ is a full description of the world, listing all things that are true. Only one of these states is the true state of the world. The true state is not necessarily known. All the other states are hypothetical; they are often introduced to describe players' beliefs. Consider also an individual who has some knowledge about the true state $\omega \in \Omega$ but may still face some uncertainty about the state. Any subset $E \subseteq \Omega$ is called an *event*, and E and F are designated as generic events. The complement of an event E is denoted by $\sim E \equiv \Omega \backslash E$ (is read: event E does not occur).[1] I will next present alternative—but equivalent—formulations of knowledge.

8.1.1 Knowledge Function

Knowledge function is the canonical model of knowledge in Economics: the individual observes the value of a function

$$\kappa : \Omega \to X$$

from the state space to some abstract space X. The function κ is often called a *signal* in applications. I will use the next example throughout the chapter to illustrate the concepts introduced.

Example 8.1 *Let $\theta \in \mathbb{N} = \{0, 1, \ldots\}$ be the price of bread. Ann does not know the price, but she observes the price with some additive noise*

$$\kappa_A = \theta + \varepsilon$$

where $\varepsilon \in \{-1, 0, 1\}$ is the noise. Here, the state space is $\Omega = \mathbb{N} \times \{-1, 0, 1\}$, where a state of the world is a price-noise pair (θ, ε). The knowledge function $\kappa_A : \Omega \to \mathbb{R}$ is defined by

$$\kappa_A (\theta, \varepsilon) = \theta + \varepsilon.$$

Example 8.2 *In the previous example, consider another individual, Bob, who knows the price θ of a bread, but he does not know what Ann observes. Formally, Ω is defined as above, but Bob's knowledge function $\kappa_B : \Omega \to \mathbb{R}$ is given by*

$$\kappa_B (\theta, \varepsilon) = \theta.$$

Recall that an event E is any subset of Ω. For example, event

$$E_1 = \{(\theta, \varepsilon) \in \Omega | \theta \geq 1\}$$

[1] When there is a continuum of states, one may impose a measurability restriction, as in Probability Theory.

represents the statement that "the price of a bread is at least 1 dollar". On the other hand, the event

$$E_{A,1} = \{(\theta, \varepsilon) \, |\theta + \varepsilon \geq 1\}$$

represents the statement that "Ann's signals is at least 1".

8.1.2 Information Function

The information content of a knowledge function is limited to the restriction it imposes on the whereabouts of the true state. By observing the value $\kappa(\omega)$, individual learns that ω is within the set

$$I(\omega) = \{\omega' | \kappa(\omega') = \kappa(\omega)\}. \tag{8.1}$$

This is the entire informational content of the signal she observes.

In the ongoing example, the informational content of Ann's signal is

$$I_A(\theta, \varepsilon) = \{(\theta', \varepsilon') \, |\theta + \varepsilon = \theta' + \varepsilon'\}.$$

Ann learns which $-45°$-line the true state lies, but she does not know where it is in that line. The informational content of Bob's signal is

$$I_B(\theta, \varepsilon) = \{(\theta', \varepsilon') \, |\theta' = \theta'\} = \{\theta\} \times [-1, 1].$$

For example, suppose that the true state is $\omega_0 = (1, 0)$, i.e., the price of a bread is 1 dollars and Ann observes a signal 1, so that her signal tells exactly what the price is. In that case, Ann will observe $\kappa_A(\omega_0) = 1$ and will learn that the true state is in the set $I_A(\omega_0) = \{(\theta, \varepsilon) \, |\theta + \varepsilon = 1\}$. She learns that the price of bread is in between 0 and 2, i.e., $\theta \in [0, 2]$, and the noise in her signal is $1 - \theta$. On the other hand, Bob will observe $\kappa_B(\omega_0) = 1$ and will learn that the true state is in the set $I_B(\omega_0) = \{\theta\} \times [-1, 1]$. He learns that the price of a bread is 1, but he does not learn anything about the noise in Ann's observation.

The set $I(\omega)$ is called the *information set*, and the function

$$I : \Omega \to 2^{\Omega} \backslash \{\varnothing\}$$

is called the *information function*. Here, $I(\omega)$ is interpreted as the set of states that one cannot rule out when the true state is ω. In other words, this is the set of states that

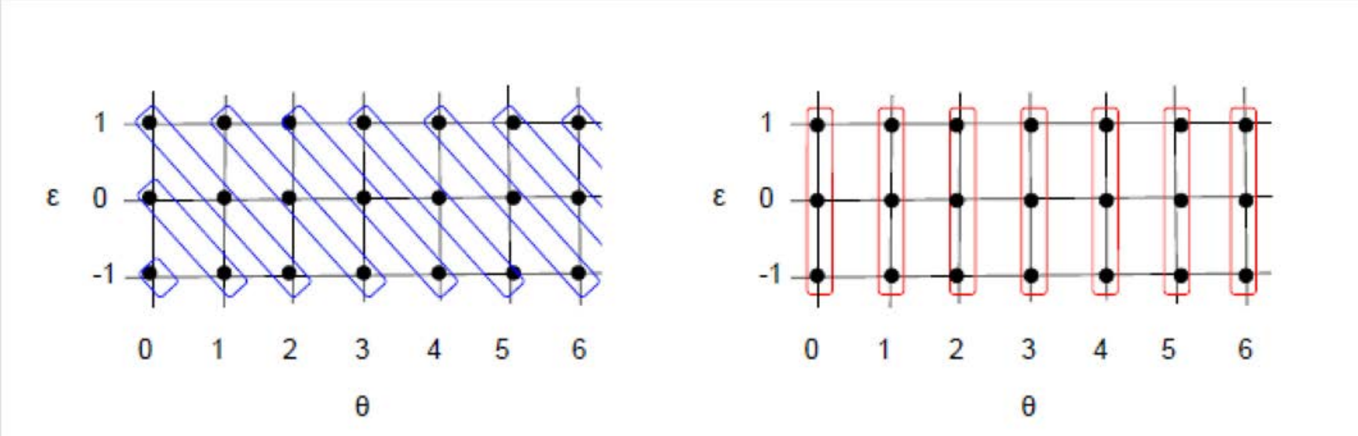

Figure 8.1: Information sets of Ann (on the left) and Bob (on the right)

she finds possible. In that sense, one can call it a *possibility set*, calling I a *possibility function*. The terms information set and information function implies that they are derived from a knowledge function through (8.1).

The information sets of Ann and Bob are plotted in Figure 8.1. The figure on the left-hand side plots the information set $I_A(\omega)$ of Ann. For each $\omega = (\theta, \varepsilon)$, the information set $I_A(\omega)$ at ω is the set that contains ω. These information sets have a $-45°$ angle, reflecting the fact that she observes the sum $\theta + \varepsilon$. At state $\omega = (0, -1)$, her information set is singleton, and she knows the true state. Bob's information sets are plotted on the right-hand side. His information sets are vertical lines, reflecting the fact that he knows θ and does not have any information about ε. Observe that (1) at each state ω, the true state is contained in the information set $I(\omega)$, and (2) the information sets form a partition for each individual. In particular, if ω' is not ruled out at ω, then the information sets at states ω and ω' are identical. These are general properties imposed by the fact that information function is derived from a knowledge function by (8.1):

I1 $\omega \in I(\omega)$;

I2 $I(\omega) \neq I(\omega') \Rightarrow I(\omega) \cap I(\omega') = \varnothing$.

Exercise 8.1 *Prove that a function $I : \Omega \to 2^{\Omega} \setminus \{\varnothing\}$ satisfied properties I1 and I2 if and only if I is derived from a knowledge function κ through (8.1).*

The property I1 is called *the Truth Axiom*, in this formulation of knowledge. It states that the individual does not rule out the true state, i.e., the individual is not delusional (although the term "non-delusion assumption" will be used for a slightly different assumption). The second property, I2, is the standard property of information sets in game theory: two distinct information set does not intersect each other. If it did, one could not have a coherent story about information sets without knowledge of the true state. I2 allows one to focus information sets without referring to the underlying true state. Of course, under bounded rationality, one could easily have intersecting distinct information set, when one would need to represent the knowledge through information function.

More generally, any function $I : \Omega \to 2^{\Omega} \setminus \{\varnothing\}$ with properties I1 and I2 is called an *information function*. Clearly, an information function is a specific case of a knowledge function, although every knowledge function yields an information function by (8.1). Throughout the course, we will mostly work with an information function, but knowledge function is also very useful, especially in concrete applications.

8.1.3 Information Partition

As discussed above, under I2, one can simply model the information of an individual by describing the information sets. Of course, under I1 and I2, the information sets form a partition

$$\mathcal{I} = \{I(\omega) \,|\omega \in \Omega\} \in \Pi, \tag{8.2}$$

where Π denotes the set of all partitions of Ω. The partition $\mathcal{I}$ is called the *information partition*. The information partitions in the ongoing example are plotted in Figure 8.1; the information partition of Ann consists of -45 degree lines, while the information partition of Bob consists of vertical lines. A partition seems to be the simplest way to represent the knowledge (when it is feasible). Clearly, every partition $\mathcal{I} \in \Pi$ is associated with an information function I where $I(\omega)$ is uniquely defined by

$$\omega \in I(\omega) \in \mathcal{I}, \tag{8.3}$$

i.e., the unique cell that contains ω. The information function I is defined by (8.3) when one starts with an information partition $\mathcal{I} \in \Pi$.

8.1.4 Knowledge Operator

While partition model is useful in representing the information, as it has been done in game theory classes at all levels, translating the information about the players' knowledge in those models to plain English is somewhat difficult. Knowledge operator comes in handy for such a task. The *knowledge operator* K on the set of events is defined by

$$K(E) = \{\omega | I(\omega) \subseteq E\}. \tag{8.4}$$

Note that when $I(\omega) \subseteq E$, the individual knows that event E holds at state ω. This is because the event E occurs no matter what possible state $\omega' \in I(\omega)$ is. Conversely, if $I(\omega) \nsubseteq E$, she does not rule out the possibility that the true state is in $I(\omega) \backslash E$ and the event E does not occur. Therefore, $K(E)$ is the states of the world at which the individual knows that event E occurs. One says that the individual *knows* event E at state ω if $\omega \in K(E)$. The event $K(E)$ is often simply denoted by KE.

The knowledge operator has several properties:

K1 $KE \subseteq E$;

K2 $E \subseteq F \Rightarrow KE \subseteq KF$;

K3 $KE \subseteq KKE$

K4 $\sim KE \subseteq K \sim KE$.

Exercise 8.2 *Show that an operator $K : 2^{\Omega} \to 2^{\Omega}$ satisfies the properties K1-K4 whenever K is defined by (8.4) from an information function I (with properties I1 and I2).*

Here, K1 is implied by $\omega \in I(\omega)$ and states that if the individual knows E then E is true (or E occurs). This property is referred to as the Truth Axiom. Philosophically, this is what distinguishes knowledge from belief; knowledge implies truth, while one may be wrong no matter how strongly she believes in something. Property K2 is known as *logical omniscience*. It states if an event E logically implies an event F (i.e., $E \subseteq F$), then whenever she knows E she also knows F. In other words, she can figure out all of the logical implications of her information. For example, knowing the basic axioms of mathematics, she can tell you any digit of the number π. Property K3 is known as the axiom of *positive introspection*. It states that if an individual knows that event E

occurs, then she knows that she knows that event E occurs. Property K4 is known as the axiom of *negative introspection*. It states that if an individual does not know that event E occurs, then she knows that she does not know that event E occurs. The axioms K1-K4 is the defining axiom of the knowledge operator in the partition model.

Definition 8.1 *A mapping* $K : 2^{\Omega} \to 2^{\Omega}$ *is called a* knowledge operator *if it satisfies the properties K1-K4.*

There are many other useful properties of a knowledge operator, as stated in the following exercise.

Exercise 8.3 *Using the properties K1-K4, prove the following.*

1. $K(E \cap F) = KE \cap KF$;

2. $KE = KKE$;

3. $\sim KE = K \sim KE$;

4. $K\Omega = \Omega$;

Property 1 is a useful case of logical omniscience: knowing events E and F is equivalent to knowing E and knowing F. Interestingly, property $K\Omega = \Omega$ states that the individual knows the state space, or the model. That is, if a feature is true throughout the state space, then the individual knows it at every state. For example, if we were to put Ann and Bob in the same model, then Ann would know that Bob knows θ because Bob knows θ at every state.

One can represent a knowledge operator—with properties K1-K4—via an information function or information partition. One can accomplish this by defining

$$I(\omega) = \sim K(\sim \{\omega\}). \tag{8.5}$$

The following exercise asks you to prove that one can indeed obtain an information partition in this way.

Exercise 8.4 *For any knowledge operator* K*, show that the mapping* I *defined by (8.5) satisfies the following properties:*

1. $\omega \in I(\omega)$;

2. $\omega' \in I(\omega) \Rightarrow I(\omega) = I(\omega')$;

3. $\omega' \notin I(\omega) \Rightarrow I(\omega) \cap I(\omega') = \varnothing$..

Example 8.3 *In the ongoing example, consider the event*

$$E_1 = \{(\theta, \varepsilon) | \theta \geq 1\}$$

that the price of bread is at least one dollar. In Figure 8.2, this is the shaded rectangular area. Observe that Ann would know that the price of bread is at least 1 dollar if her signal is at least 2:

$$K_A E_1 = \{(\theta, \varepsilon) | \theta + \varepsilon \geq 2\} \equiv E_{A,2}$$

where K_A is the knowledge operator for Ann, which is derived from information function I_A. This is the second shaded area contained in E_1 in Figure 8.2. This event is obtained by combining all of her information sets that are contained in E_1. Observe that in the triangular area on the left-bottom corner of event E_1, Ann does not know that the price is at least zero although it is indeed the case. On the other hand, knowing the price of a bread, Bob would know that the price of a bread is at least one dollar if and only if it is indeed at least one dollar:

$$K_B E_1 = \{(\theta, \varepsilon) | \theta \geq 1\} = E_1$$

where K_B is the knowledge operator for Ann, which is derived from information function I_B.This can be viewed vividly in Figure 8.2: he has no information set that spans both E_1 and $\sim E_1$, so that an information set is either in or out.

The above formulations of knowledge are quite standard: knowledge function is the standard formulation of information in economic applications, while information sets (and associated information function and information partition) are the standard devices for modeling players' information in Game Theory. On the other hand, although they are less common in Economics and Game Theory, knowledge operators are the standard formulation of knowledge in Epistemology, and they are easily translated to plain English by substituting "one knows" for K, "if" for subset, "and" for intersection and "or" for

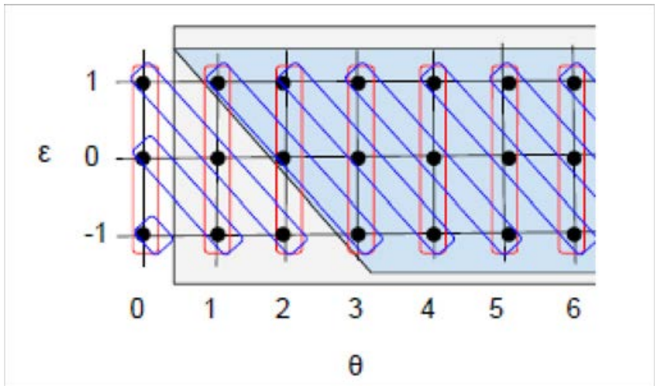

Figure 8.2: The events that the price of bread is at least 1 dollar (shaded rectangular area) and that Ann knows this (the second shaded area inside above rectangle)

union. For example, one can translate $K(E \cap F) \subseteq K(E) \cap K(F)$ as "if one knows E and F, then she knows E and she knows F".

I will next present two more representations of knowledge. These representations are not as standard or as natural as the previous representations. Nonetheless, they are quite useful in analyses, especially the analyses of multi-person information structures.

8.1.5 Knowledge Field

Given a partition $\mathcal{I} \in \Pi$, define

$$\mathcal{K} = \{\cup_{\omega \in \Omega'} I(\omega) \,|\Omega' \subseteq \Omega\}$$

as the set of events that can be written as a disjoint union of information sets (or the cells in the partition). The set $\mathcal{K}$ is called the *knowledge field*. A knowledge field represents the set of events that the individual knows whenever the event occurs. Indeed, one can easily check that

1. $E \in \mathcal{K} \iff E = KF$ for some event F;
2. $E \in \mathcal{K} \iff KE = E$.

These statements provide alternative equivalent definitions for a knowledge field. The first statement states that $\mathcal{K}$ is the set of events corresponding to knowing some other event. The second statement, which implies the first one, states that $\mathcal{K}$ is the set of fixed points of the knowledge operator. It is the set of knowable events, in that the individual knows those events whenever they occur. Formally, such events are called *self-evident* for the individual considered. A knowledge field has the following properties:

Closure under complement $E \in \mathcal{K} \iff \sim E \in \mathcal{K}$

Closure under union for any subset $\mathcal{K}' \subseteq \mathcal{K}$, $\bigcup_{E \in \mathcal{K}'} E \in \mathcal{K}$

Closure under intersection for any subset $\mathcal{K}' \subseteq \mathcal{K}$, $\bigcap_{E \in \mathcal{K}'} E \in \mathcal{K}$

Self-evidence of state space $\Omega \in \mathcal{K}$.

It is a straightforward exercise to derive these properties from the definition based on the information partition. In terms of knowledge operator, the first property follows from negative introspection and self evidence. By negative introspection, the individual knows that a self-evident event does not hold whenever it does not hold. Hence, the complement of a self-evident event is also a self-evident event. The first two statements imply the last two.

Any subset $\mathcal{K}$ of power set with closure under complement and union is called a *knowledge field*. One can obtain the other equivalent formulations of knowledge as follows. One can obtain the knowledge operator K from a knowledge field $\mathcal{K}$ by setting KE as the largest event $F \in \mathcal{K}$ with $F \subseteq E$. One can also obtain the information function I from $\mathcal{K}$ by setting $I(\omega) = \bigcap_{\omega \in E \in \mathcal{K}} E$.

8.1.6 Information Graph

Define the *information graph* G with the set of nodes Ω by

$$G_{\omega,\omega'} = 1 \iff \omega' \in I(\omega).$$

That is, the states ω and ω' are linked if they belong to the same information set. Hence, the information sets correspond to the components of the information graph (i.e., the maximally connected subgraphs). Note that the property I2 of information function implies that the graph G is undirected.

In our ongoing example, the information graphs are plotted in Figure 8.3. Observe that each graph consists of completely connected components. That is, if there is a path

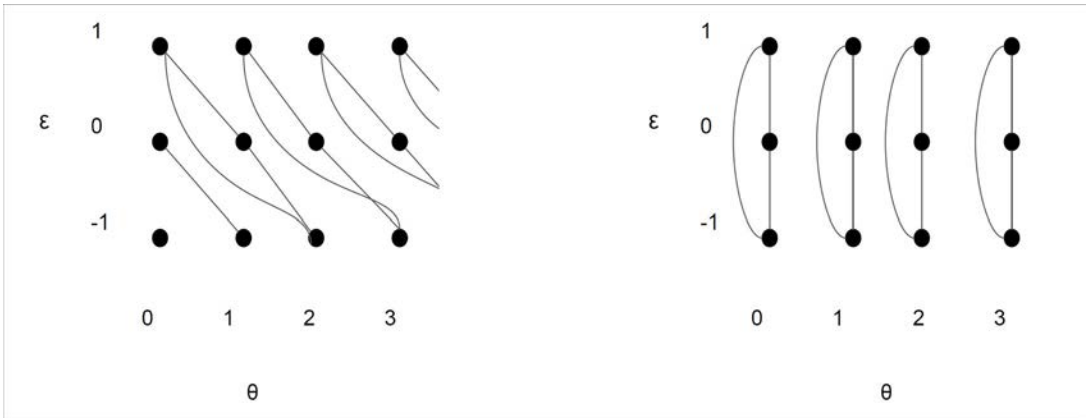

Figure 8.3: The information graphs for Ann (left) and Bob (right) in our ongoing example

$\omega_0, \omega_1, \ldots, \omega_m$ with $G_{\omega_k,\omega_{k+1}} = 1$ between any two states ω_0 and ω_m, then there is a direct link between ω_0 and ω_m: $G_{\omega_0,\omega_m} = 1$. Moreover, each state is connected to itself, i.e., $G_{\omega,\omega} = 1$. These two properties hold more generally for any information graph that is derived from an information function as above. Any undirected graph with these properties is called *information graph.* Given any information graph G, one can define an information function I by $\omega' \in I(\omega) \iff G_{\omega,\omega'} = 1$.

Note that the above formulations of knowledge (i.e., knowledge function, information function, information partition, knowledge operator, knowledge field, and information graph) are equivalent, in that one can start from any of these formulations and derive the other formulations using the relevant formulas in the text above, and the properties described above will be satisfies so long as the original formulation satisfies the properties for that formulation.

8.1.7 Beliefs and Certainty

Throughout the course, I will assume that each player has a belief about the events that she is uncertain about. These beliefs are represented by the *belief function* p that maps each state ω to a probability distribution

$$p_\omega \in \Delta(I(\omega))$$

on the information set $I(\omega)$, where $\Delta(X)$ denotes the set of probability distributions on X. The belief p_ω is extended to entire state space as usual by setting

$$p_\omega(E) = p_\omega(E \cap I(\omega))$$

for every event E. Here, $p_\omega \in \Delta(I(\omega))$ incorporates the assumption that the individual's beliefs respect her information; she does not believe in things that she knows to be false. It is also assumed that she knows her beliefs. That is,

$$\omega' \in I(\omega) \Rightarrow p_{\omega'} = p_\omega.$$

The individual is said to be *certain about event E at state ω* if she assigns probability 1 on event E at state ω. That is, $p_\omega(E) = 1$. The certainty operator C on events E is defined by

$$C(E) = \{\omega | p_\omega(E) = 1\}.$$

This is the event that the individual assigns probability 1 on event E.

Observe that the individual may be certain that the true state is not the case, i.e.,

$$p_\omega(\sim \{\omega\}) = 1.$$

In that case, she is called *delusional*. Indeed, this case arises if and only if the individual assigns zero probability on some state in her information set. In particular, if $p_\omega(\omega') = 0$ for some $\omega' \in I(\omega)$, then she will be delusional at state ω':

$$p_{\omega'}(\sim \{\omega'\}) = p_\omega(\sim \{\omega'\}) = 1.$$

Accordingly, the individual is said to be *non-delusional* if her belief has full support on all information sets: $p_\omega(\omega') > 0$ for all ω, ω' with $\omega' \in I(\omega)$.

This is the main difference between knowledge and certainty: if one knows something, then it is true (the Truth Axiom), but one may be certain about things that are false. The next exercise asks you to check that certainty exhibits all the other properties of knowledge.

Exercise 8.5 *Show that the operator C satisfies the following properties:*

C1 $KE \subseteq CE$

C2 $E \subseteq F \Rightarrow CE \subseteq CF$;

C3 $CE = KCE \subseteq CCE$

C4 $\sim CE = K \sim CE \subseteq C \sim CE$.

The first property above states that knowledge implies certainty: if one knows something, the she will be certain of it. When an individual is delusional the truth axiom K1 does not hold, and we may have $CE \nsubseteq E$. The properties C2-C4 are versions of the knowledge axioms K2-K4 for certainty, taking into account that K1 does not hold. Property C2 is the logical omniscience axiom K2 for the operator C, stating that if one is certain of an event, then she will be certain of all its implications. Property C3 is has two parts. First part states that if one is certain of an event, then she will know that she is certain, as one knows her own beliefs. The second part, which is implied by C1 and the first part, is positive introspection for the operator C: if one is certain of something, then she will be certain of the fact that she is certain. We do not have equality because K1 does not hold for certainty. Likewise, C4 establishes negative introspection for certainty in two parts: if one is not certain of an event, then she will know that she is not certain, and she will be certain of the fact that she is not certain.

As in the case of certainty, other weaker forms of knowledge are obtained by dropping or weakening some of the knowledge axioms K1-K4.

8.2 Informativeness and Lattice of Information Partitions

This section defines a partial order on information partition according to their informativeness. Observing that the set of information partitions forms a complete lattice under this order, it computes the join and meet operators for the lattice. These operators are highly important for interactive epistemology as they define distributed and common knowledge in any interactive epistemic model.

Definition 8.2 *For any information partitions $\mathcal{I}$ and $\mathcal{I}'$ with information functions I and I', the information partition $\mathcal{I}$ is said to be* finer than $\mathcal{I}'$, *written as $\mathcal{I} \geq \mathcal{I}'$ if*

$$I'(\omega) = \bigcup_{\omega' \in I'(\omega)} I(\omega')$$

for all ω. The information partition $\mathcal{I}'$ is said to be coarser than $\mathcal{I}$ *if* $\mathcal{I} \geq \mathcal{I}'$.

That is, the cells in $\mathcal{I}'$ can be written as disjoint unions of the cells in $\mathcal{I}$. This is the case when information partition $\mathcal{I}$ has more information than $\mathcal{I}'$, in that one obtains $\mathcal{I}$ from $\mathcal{I}'$ by breaking some of the information sets into smaller sets. Consequently, an individual with information partition $\mathcal{I}$ knows more than an individual with information partition $\mathcal{I}'$, as the next result formalizes.

Proposition 8.1 *Let $\mathcal{I}$ be an information partition associated with information function I, knowledge operator K, and knowledge field $\mathcal{K}$. Let also $\mathcal{I}'$ be an information partition associated with information function I', knowledge operator K', and knowledge field $\mathcal{K}'$. Then, the following are equivalent:*

1. $\mathcal{I} \geq \mathcal{I}'$.

2. $I(\omega) \subseteq I'(\omega)$ *for all* $\omega \in \Omega$.

3. $K'E \subseteq KE$ *for all events* E.

4. $\mathcal{K}' \subseteq \mathcal{K}$.

Exercise 8.6 *Prove Proposition 8.1. Translating the statements 2-4 to plain English, briefly discuss the equivalence. Finally, suppose Alex observes the value of a knowledge function κ_1, while Beatrice observes the value of the knowledge function κ_1 and the value of another knowledge function κ_2. Show that Beatrice knows more than Alex in the sense of Proposition 8.1.*

One can show that, for any fixed state space Ω, the set of all information partitions forms a lattice with join operator $\vee$ and meet operator $\wedge$, defined as follows.

For any two information partitions $\mathcal{I}$ and $\mathcal{I}'$, their join $\mathcal{I} \vee \mathcal{I}'$ is defined as the coarsest information structure that is finer than both $\mathcal{I}$ and $\mathcal{I}'$. That is,

1. $\mathcal{I} \vee \mathcal{I}' \geq \mathcal{I}$ and $\mathcal{I} \vee \mathcal{I}' \geq \mathcal{I}'$, and

2. if $\mathcal{I}'' \geq \mathcal{I}$ and $\mathcal{I}'' \geq \mathcal{I}'$, then $\mathcal{I}'' \geq \mathcal{I} \vee \mathcal{I}'$.

Here, the first condition states that $\mathcal{I} \vee \mathcal{I}'$ is more informative than both $\mathcal{I}$ and $\mathcal{I}'$. The second condition states that it is least informative information partition with that property. Hence, the join $\mathcal{I} \vee \mathcal{I}'$ is obtained by combining the information distributed among the information structures $\mathcal{I}$ and $\mathcal{I}'$ and nothing more. Indeed, the join $\mathcal{I} \vee \mathcal{I}'$ can be defined through the information function

$$(I \vee I')(\omega) = I(\omega) \cap I'(\omega),$$

obtained by combining the information contained in information functions I and I'. Observe that if an individual observed the value of the two functions I and I' (as her knowledge function), her information function would be $I \vee I'$. In this sense, $\mathcal{I} \vee \mathcal{I}'$ formally represents the distributed information among $\mathcal{I}$ and $\mathcal{I}'$.

Example 8.4 *In our ongoing example, the join $\mathcal{I}_A \vee \mathcal{I}_B$ of information partitions of Ann and Bob is the full information partition:*

$$\mathcal{I}_A \vee \mathcal{I}_B = \{\{\omega\} \mid \omega \in \Omega\}.$$

Indeed, for any $\omega = (\theta, \varepsilon)$, the intersection of the information sets $I_A(\omega) = \{(\theta', \varepsilon') \mid \theta + \varepsilon = \theta' + \varepsilon'\}$ and $I_B(\omega) = \{\theta\} \times \{-1, 0, 1\}$ is $\{\omega\}$. Using Bob's information, one obtains the price, and combining this with Ann's information, one finds the noise in Ann's observation.

Now, we turn to meet operator $\wedge$. The meet operator for information partitions is essential for Game Theory. For any two information partitions $\mathcal{I}$ and $\mathcal{I}'$, their meet $\mathcal{I} \wedge \mathcal{I}'$ is defined as the coarsest information structure that is finer than both $\mathcal{I}$ and $\mathcal{I}'$, and it is often called the *common coarsening* of $\mathcal{I}$ and $\mathcal{I}'$. That is,

1. $\mathcal{I} \geq \mathcal{I} \wedge \mathcal{I}'$ and $\mathcal{I}' \geq \mathcal{I} \wedge \mathcal{I}'$, and
2. if $\mathcal{I} \geq \mathcal{I}''$ and $\mathcal{I}' \geq \mathcal{I}''$, then $\mathcal{I} \wedge \mathcal{I}' \geq \mathcal{I}''$.

The first condition states that both $\mathcal{I}$ and $\mathcal{I}'$ are more informative than $\mathcal{I} \wedge \mathcal{I}'$. That is, the information contained in $\mathcal{I} \wedge \mathcal{I}'$ is also contained in both $\mathcal{I}$ and $\mathcal{I}'$. Thus, $\mathcal{I} \wedge \mathcal{I}'$ contains only the knowledge that is common to both information partitions. The second condition states that it is the most informative information partition with that property. Hence, $\mathcal{I} \wedge \mathcal{I}'$ represents the knowledge that is common in both $\mathcal{I}$ and $\mathcal{I}'$. Indeed, the information partition $\mathcal{I} \wedge \mathcal{I}'$ will formally represent the common knowledge with respect to $\mathcal{I}$ and $\mathcal{I}'$.

The conditions 1 and 2 that define $\wedge$ take a useful form in terms of information functions; it immediately follows from the definition of $\geq$ and the above properties.

Proposition 8.2 *Let $I \wedge I'$ be the information function associated with $\mathcal{I} \wedge \mathcal{I}'$. Then, for each $\omega \in \Omega$, $(I \wedge I')(\omega)$ is the smallest set $\tilde{I}$ (in the sense of set inclusion) with*

$$\omega \in \tilde{I} = \bigcup_{\omega' \in \tilde{I}} I(\omega') = \bigcup_{\omega' \in \tilde{I}} I'(\omega').$$

That is the information sets under the common coarsening can be written as disjoint union of information sets for each information sets, and they are the smallest sets that can be written that way.

Unfortunately, computation of the common coarsening is not as straightforward as taking intersection or union of information sets. The next result provides an algorithm to compute common coarsening.

Proposition 8.3 *Let $I \wedge I'$ be the information function associated with $\mathcal{I} \wedge \mathcal{I}'$. For any $\omega, \omega' \in \Omega$, $\omega' \in (I \wedge I')(\omega)$ if and only if there exists a sequence $\omega_0 = \omega$, $\omega_1, \ldots, \omega_n = \omega'$ such that $\omega_{k+1} \in I(\omega_k) \cup I'(\omega_k)$ for each $k = 0, \ldots, n-1$.*

The proposition gives an algorithm to compute the information set $(I \wedge I')(\omega)$ under common coarsening for each ω: take all ω' that is not ruled out at ω by either of the information partitions I and I'. This is the set $I_1 = I(\omega) \cup I'(\omega)$. Then, for each $\omega' \in I_1$, pick all the states ω'' that is not ruled out at ω' by either of the information partitions I and I'. This is the set $I_2 = \bigcup_{\omega' \in I_1} I(\omega') \cup I'(\omega')$. Continuing in this fashion construct an increasing sequence of sets I_k, $k \geq 0$, by setting $I_k = \bigcup_{\omega' \in I_{k-1}} I(\omega') \cup I'(\omega')$. The set $(I \wedge I')(\omega)$ is the limit $\bigcup_{k=0}^{\infty} I_k$.

Example 8.5 *Consider $\omega = (0, -1)$ in our ongoing example; this is the state at the bottom left corner in Figure 8.1. Observe that $I_A(\omega) = \{\omega\}$ while $I_B(\omega) = \{0\} \times \{-1, 0, 2\}$. Hence, the set*

$$I_1 = I_A(\omega) \cup I_B(\omega) = \{0\} \times \{-1, 0, 1\}$$

is contained in $(I_A \wedge I_B)(\omega)$. *Then, the set*

$$I_2 = \bigcup_{\omega' \in I_1} I_A(\omega') \cup I_B(\omega') = I_1 \cup \bigcup_{\omega' \in I_1} I_A(\omega') = \{(\theta, \varepsilon) \mid \theta + \varepsilon \leq 1\}$$

is contained in $(I_A \wedge I_B)(\omega)$. *This is the largest triangle at the lower corner of the state space. Next one obtains the rectangular event*

$$I_3 = I_2 \cup \bigcup_{\omega' \in I_2} I_B(\omega') = \{0, 1, 2\} \times \{-1, 0, 1\}.$$

Continuing in this fashion one obtains a larger rectangle in every two rounds by including two more values of the price. Thus,

$$(I_A \wedge I_B)(\omega) = \bigcup_k \{0, 1, \ldots, 2k\} \times \{-1, 0, 1\} = \Omega.$$

The common coarsening of Ann and Bob's information partitions is the trivial partition

$$\mathcal{I}_A \wedge \mathcal{I}_B = \{\Omega\},$$

void of any information beyond the state space.

As it turns out, the set of all information partitions is a *complete* lattice, in which every family of information partitions have greatest lower bound and the least upper bound. This is formally stated next.

Theorem 8.1 *For any state space* Ω, *the set of all information partitions of* Ω *is a complete lattice under the order* $\geq$. *Every family* $\mathcal{I}_\alpha$, $\alpha \in A$, *of information partitions associated with information functions* I_α *and knowledge fields* $\mathcal{K}_\alpha$ *has the least upper bound* $\bigvee_{\alpha \in A} \mathcal{I}_\alpha$ *and the greatest lower bound* $\bigwedge_{\alpha \in A} \mathcal{I}_\alpha$ *(called the common coarsening of the family* $\mathcal{I}_\alpha$, $\alpha \in A$*) where*

1. $\bigvee_{\alpha \in A} \mathcal{I}_\alpha$ *is associated with information function* $\bigvee_{\alpha \in A} I_\alpha$ *with*

$$\left(\bigvee_{\alpha \in A} I_\alpha\right)(\omega) = \bigcap_{\alpha \in A} I_\alpha(\omega)$$

2. $\bigwedge_{\alpha \in A} \mathcal{I}_\alpha$ *is associated with the knowledge field*

$$\bigcap_{\alpha \in A} \mathcal{K}_\alpha.$$

Proof. It suffices to show the Properties 1 and 2 in the proposition for an arbitrary family of information partitions. Take any family as in the statement of the theorem.

Proof of Property 1: First observe that $\bigvee_{\alpha \in A} I_\alpha$ is indeed an information function, i.e., it satisfies the properties I1 and I2 (by the virtue of each I_α satisfying these properties). Observe also that, for each $\hat{\alpha} \in A$,

$$\left(\bigvee_{\alpha \in A} I_\alpha \right)(\omega) = \bigcap_{\alpha \in A} I_\alpha(\omega) \subseteq I_{\hat{\alpha}}(\omega),$$

showing by Proposition 8.1 that $\bigvee_{\alpha \in A} \mathcal{I}_\alpha \geq \mathcal{I}_{\hat{\alpha}}$, i.e., $\bigvee_{\alpha \in A} \mathcal{I}_\alpha$ is an upper bound. Moreover, take any upper bound $\mathcal{I}^*$ associated with information function I^*, where $I^*(\omega) \subseteq I_\alpha(\omega)$ for each α and ω (by Proposition 8.1). Then, for each ω,

$$I^*(\omega) \subseteq \bigcap_{\alpha \in A} I_\alpha(\omega) = \left(\bigvee_{\alpha \in A} I_\alpha \right)(\omega),$$

showing that $\mathcal{I}^* \geq \bigvee_{\alpha \in A} \mathcal{I}_\alpha$.

Proof of Property 2: First observe that $\bigcap_{\alpha \in A} \mathcal{K}_\alpha$ is indeed a knowledge field, i.e., it is closed under negation and arbitrary unions (by virtue of the fact that each $\mathcal{K}_\alpha$ is closed under these operations). Let $\bigwedge_{\alpha \in A} \mathcal{I}_\alpha$ be the information partition associated with $\bigcap_{\alpha \in A} \mathcal{K}_\alpha$. To show that $\bigwedge_{\alpha \in A} \mathcal{I}_\alpha$ is a lower bound, take any $\hat{\alpha} \in A$. Observe that

$$\bigcap_{\alpha \in A} \mathcal{K}_\alpha \subseteq \mathcal{K}_{\hat{\alpha}}.$$

Hence, by Proposition 8.1, $\mathcal{I}_{\hat{\alpha}} \geq \bigwedge_{\alpha \in A} \mathcal{I}_\alpha$. Moreover, it is the finest such lower bound. Indeed, take lower bound $\mathcal{I}^*$ associated with knowledge field $\mathcal{K}^*$, where $\mathcal{I}_\alpha \geq \mathcal{I}^*$ for each α. Then, by Proposition 8.1, $\mathcal{K}^* \subseteq \mathcal{K}_\alpha$ for each α. Hence,

$$\mathcal{K}^* \subseteq \bigcap_{\alpha \in A} \mathcal{K}_\alpha,$$

showing that $\bigwedge_{\alpha \in A} \mathcal{I}_\alpha \geq \mathcal{I}^*$. ■

Theorem 8.1 provides a characterization of the common coarsening $\bigwedge_{\alpha \in A} \mathcal{I}_\alpha$ of an arbitrary family of information partitions. It shows that it is the information partition that corresponds to the intersection of all knowledge fields in the family. That is, it corresponds to the set of events that are knowable under every information partition. In that sense, it is the information partition that captures the information that is common to all information partitions in the family. Unfortunately, this does not necessarily tell us how to compute the common coarsening directly from information partitions. The next result shows that one can do that as in the computation of meet when the family is finite.

Proposition 8.4 *Consider any family* $\mathcal{I}_1, \ldots, \mathcal{I}_n$ *of information partitions, associated with information functions* $I_1, \ldots, I_n$*. Let* $\bigwedge_{i=1}^n I_i$ *be the information partition associated with* $\bigwedge_{i=1}^n \mathcal{I}_i$*. For any* $\omega, \omega' \in \Omega$*,* $\omega' \in \left(\bigwedge_{i=1}^n I_i\right)(\omega)$ *if and only if there exists a sequence* $\omega_0 = \omega$*,* ω_1*,. . . ,*$\omega_n = \omega'$ *such that* $\omega_{k+1} \in I_1(\omega_k) \cup \cdots \cup I_n(\omega_k)$ *for each* $k = 0, \ldots, n-1$*.*

8.3 Interactive Epistemology

This section builds on the single-person knowledge model to develop epistemology for multiple players who interact with each other. It presents several formulations of common knowledge as well as some other basic notions.

8.3.1 Epistemic Models

When multiple players interact with each other, one needs to specify not only players' information and beliefs about the underlying fundamentals but also their beliefs and information about each other. This is accomplished as follows.

Definition 8.3 *A* model *(or interchangeably an* information structure*) is a tuple* (N, Ω, I, p) *consisting of*

- *a set of players* $N = \{1, 2, \ldots, n\}$*;*
- *a state space* Ω*;*
- *a collection of information functions* I_i *for each player* i*;*
- *a collection of* $p_{i,\omega} \in \Delta(I_i(\omega))$ *beliefs over possible states in* $I_i(\omega)$ *for each player* i*, such that*
$$\omega' \in I(\omega) \Rightarrow p_{i,\omega'} = p_{i,\omega}.$$

A multi-person model extends a single-person model by simply describing information function I_i and state-dependent belief $p_{i,\omega}$ of each player i. Observe that the belief $p_{i,\omega}$ is constant over $I_i(\omega)$, so that player i knows her belief at state ω. For each player i, her information partition $\mathcal{I}_i$, her knowledge operator K_i, her knowledge field $\mathcal{K}_i$, and her information graph G_i are derived from her information function as in the previous section.

Although a state of the world is an abstract object, the mappings I_i and p_i give meaning about the players' knowledge and beliefs. As such, a state also describes players' knowledge and beliefs about the states, their knowledge and beliefs about the other players' knowledge and beliefs an so on. When the states also describe some underlying fundamentals, they will also describe each player's beliefs and knowledge about these fundamentals, each player's beliefs and knowledge about the other players' beliefs and knowledge about the fundamentals, and so on, as illustrated in our ongoing example below. Recall that only one of these states is the true state of the world. The other states and the rest of the model are introduced as hypothetical constructs to the players' beliefs and knowledge at the true state.

Example 8.6 *When Ann and Bob interact, the situation described in our ongoing example can be modeled as follows. The set of players is* $N = \{1, 2\}$*, where Player 1 is Ann and Player 2 is Bob. The state space is* $\Omega = \mathbb{N} \times \{-1, 0, 1\}$*, where each sate* ω *is a pair* (θ, ε) *of a price for a bread and a noise level in Ann's observation. At each* $\omega = (\theta, \varepsilon)$*, the information functions are given by*

$$\begin{aligned} I_1(\theta, \varepsilon) &= \{(\theta', \varepsilon') \,|\theta' + \varepsilon' = \theta + \varepsilon\} \\ I_2(\theta, \varepsilon) &= \{\theta\} \times \{-1, 0, 1\}. \end{aligned}$$

The information functions I_1 *and* I_2 *yield information partitions* $\mathcal{I}_1$ *and* $\mathcal{I}_2$*, respectively. The information partitions are plotted in Figure . The beliefs of the players were not specified previously. For the sake of completeness, assume that* $p_{i,\omega}$ *is the uniform distribution on* $I_i(\omega)$ *for each* i *and* ω*.*

In this example, one can deduce a hierarchy of knowledge and belief of players at each states as follows. At any state $\omega = (\theta, \varepsilon)$ with $t_1 \equiv \theta + \varepsilon \geq 1$, Ann does not know the price of bread; she knows that the price of bread is either $t_1 - 1$, or t_1, or

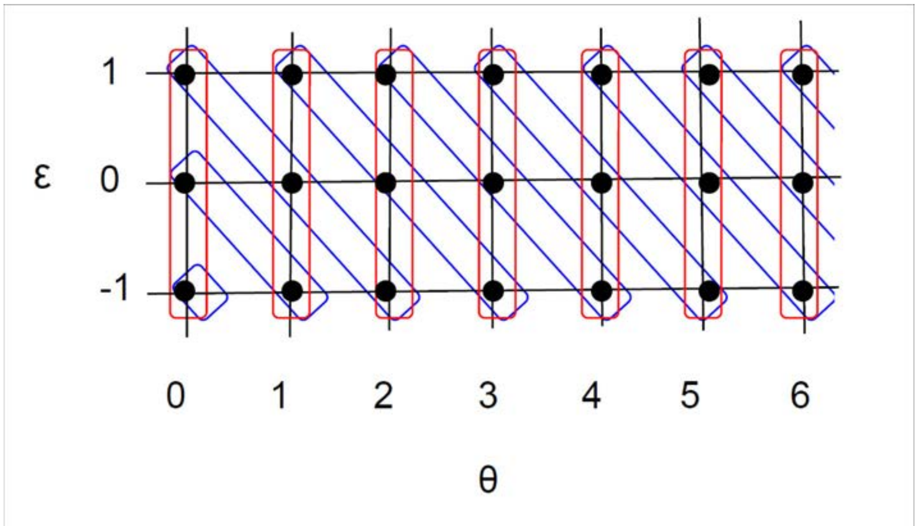

Figure 8.4: The information partitions for Ann and Bob in Example 8.6.

$t_1 + 1$; she assigns probability 1/3 on each case. When $t_1 \equiv \theta + \varepsilon = 0$, she knows that the price of bread is either 0 or 1, and she assigns probability 1/2 on each case. When $t_1 \equiv \theta + \varepsilon = -1$, she knows that the price of bread is 0. On the other hand, at each state $\omega = (\theta, \varepsilon)$, Bob does know that the price of bread is θ, but he does not know what Ann thinks. He knows that there are three possibilities: $t_1 = \theta - 1$, $t_1 = \theta$, and $t_1 = \theta + 1$. He assigns probability 1/3 on each case, and knows that in each of these cases will be determined by t_1 as above.

This leads to a hierarchy of knowledge and beliefs at any state as follows. Take $\omega = (1, 0)$, so that the price of bread is 1 and Ann happens to observe without noise (although she does not know it). Ann knows that the price of bread is either 0 or 1 or 2, and she believes that each has probability 1/3. Bob knows that the price of bread is 1. The above knowledge and beliefs are called the *first-order knowledge* and *first-order belief*, respectively. Ann knows that in each of the three cases above, Bob knows the price of bread, assigning probability 1/3 on each case: Bob knows that the price of bread is 0; Bob knows that the price of bread is 1; and Bob knows that the price of bread is 0. Bob does not know Ann's belief/information. He also think that there are three possible cases, each with probability 1/3: ($t_1 = 0$) Ann assigns equal probabilities on prices 0 and 1; ($t_1 = 1$) Ann assigns equal probabilities on prices 0, 1, and 2; and ($t_1 = 2$) Ann assigns equal probabilities on prices 1, 2 and 3. He also knows that Alice will know that

Bob knows the price of bread in each case. These describe the *second-order* knowledge and beliefs. In principle, one can describe the third-, the fourth-, ... order knowledge and belief iteratively in this manner. (Note that it gets exponentially complex as one consider higher orders.)

Note that in the above hierarchy of knowledge and beliefs, something will always be true. For example, Bob will know the price, Ann will know that Bob knows the price (although she does not necessarily know what that price is), Bob will know that Ann knows that Bob knows the price etc. These are the features that are true at every state of the world in this example. Such features will be common knowledge.

8.3.2 Mutual Knowledge

We now turn to the concept of mutual knowledge, which corresponds to the statement "everybody knows". As in the example above, we will also define higher-order mutual knowledge.

Fix any model $M = (N, \Omega, I, p)$. Recall that an event E is any subset of Ω. On these events, define the operator K_N by

$$K_N E = \cap_{i \in N} K_i E.$$

This is the event that everybody in group N knows that event E holds. Mutual knowledge differs from knowledge. The next example asks you to check how the mutual knowledge differs from knowledge in various formulations of knowlegde.

Exercise 8.7 *Check which properties K1-K4 of knowledge operator is satisfied by the mutual knowledge operator K_N. Determine the "mutual knowledge information set" $I_N(\omega)$ that corresponds to the mutual knowledge operator K_N, in that*

$$\omega \in K_N(E) \iff I_N(\omega) \subseteq E.$$

Which of the properties I1 and I2 are satisfied by I_N? Can one have an information partition representation of mutual knowledge? Briefly discuss your results.

Now, for any order $m = 0, 1, 2, \ldots$, define operator K_N^m by $K_N^0 E = E$ and

$$K_N^m E = K_N\left(K_N^{m-1} E\right)$$

for $m \geq 1$. The operator K_N^m is called *mth-order mutual knowledge*. It corresponds to the event that everybody in group N knows everybody in group N knows... that event E holds, where one repeats everybody in group N knows m times. (One can define mutual knowledge within a subset of players for any subset N' of players using the operator $K_{N'}E = \cap_{i \in N'} K_i E$.)

Example 8.7 *In our ongoing Example 8.6, consider the event*

$$E_1 = \{(\theta, \varepsilon) : \theta \geq 1\}$$

that the price of bread is at least one dollar. Since Bob knows the price of bread, he knows this whenever it is true:

$$K_2 E_1 = E_1.$$

Ann does not know the price, but she will know that the price is at least 1 if and only if her signal is at least 2:;

$$K_1 E_1 = \{(\theta, \varepsilon) \,|\theta + \varepsilon \geq 2\} \equiv E_{A,2}.$$

This is the largest knowable event contained in E_1 for Ann; it is the union of all her information sets that are contained in E_1. The events E_1 and $E_{A,1}$ are plotted in Figure 8.2. Everybody knows that the price is at least 1 if the state is in

$$K_N(E_1) = K_1 E_1 \cap K_2 E_1 = E_{A,2}.$$

Thus, the (first-order) mutual knowledge of E_1 corresponds to the event

$$K_N^1(E_1) = K_N(E_1) = E_{A,2}.$$

There will be some event (in the lower triangle in the left bottom corner of E_1) in which the price will be at least one but this fact is not mutual knowledge.

To compute the second-order mutual knowledge, observe that Ann knows whether they have first-order mutual knowledge:

$$K_1 K_N^1 E_1 = K_1 E_{A,2} = E_{A,2}.$$

This is because $K_N^1 E_1$ is the event that she knows that the price is at least 1. Bob does not necessarily know whether they have first-order mutual knowledge. For this he must know that $\theta + \varepsilon \geq 2$, and this is the case if and only if $\theta \geq 3$. Hence,

$$K_2 K_N^1 E_1 = K_2 E_{A,2} = \{(\theta, \varepsilon) : \theta \geq 3\} \equiv E_3.$$

Intuitively, in Figure 8.2, the largest rectangle that fits in the event $E_{A,1}$ is E_3. There is second-order mutual knowledge of that the price is at least 1 if and only if the state is in

$$K_N^2 E_1 = E_{A,2} \cap E_3 = E_3.$$

That is, the price of bread is at least 3. Following in this way, for any $m > 2$, one can compute that

$$\begin{aligned} K_N^m(E_1) &= \{(\theta,\varepsilon)\,|\theta+\varepsilon \geq m+1\} \equiv E_{A,m+1} \qquad (\text{if } m \text{ is odd}) \\ K_N^m(E_1) &= \{(\theta,\varepsilon)\,|\theta \geq m+1\} \equiv E_{m+1} \qquad (\text{if } m \text{ is even}). \end{aligned}$$

In particular, for any even m, there is mth-order mutual knowledge of the fact that the price is at least 1 if and only if the price is at least $m+1$. For any odd price m, if the price is m, there will be $(m-1)$th-order mutual knowledge of the fact that the price is not zero, but there will not be mth-order mutual knowledge of this fact.

Exercise 8.8 *For each $m \geq 1$, determine the "mth-order mutual knowledge information set" $I_N(\omega)$ that corresponds to the mth-order mutual knowledge operator K_N, in that*

$$\omega \in K_N^m(E) \iff I_N^m(\omega) \subseteq E.$$

Which of the properties I1 and I2 are satisfied by I_N? Can one have an information partition representation of mutual knowledge? Briefly discuss your results.

Exercise 8.9 *In Exercise 8.6, for state $\omega = (0,0)$, construct the sets $I_N^m(\omega)$ for $m \geq 1$.*

For economic modeling it turns out that what is important are the states in which there remains mutual knowledge at all orders. This is called *common knowledge*. The above example illustrates that one cannot have common knowledge of price being in a given non-trivial interval regardless of how large the interval is and how far the boundaries of the interval are from the true state. One can only have the common knowledge of the trivial fact that the price is non-negative. In any given model some facts will remain mutual knowledge at any order. In this example, as it is often the case, these are the features of the model that are true at all states, such as the fact that Bob knows the price and the fact that Ann's beliefs will be determined by $t_1 = \theta + \varepsilon$ as described above. The next section formally introduces common knowledge.

8.3.3 Common Knowledge

In essence, informally, common knowledge events are the features of a model that are true throughout the model (or a "submodel" where a state lies). There are several equivalent formal definitions of common knowledge. This section introduces these definitions.

First, one can define common knowledge as the limit of mutual knowledge for arbitrarily high orders. In particular, common knowledge operator CK is defined on events by setting

$$CK(E) = \bigcap_{m=0}^{\infty} K_N^m(E).$$

That is, event E is common knowledge at a state ω if and only if there is mth-order mutual knowledge of E at ω for each order m. As it is hinted in the previous section, there cannot be common knowledge of any non-trivial event in our ongoing example.

Example 8.8 *In Example 8.7, consider the event E_1 that the price of bread is at least one. As shown in that example for any even m,*

$$K_N^m(E_1) = \{(\theta, \varepsilon) | \theta \geq m+1\} \equiv E_{m+1}.$$

Then, since the sets $K_N^m(E_1)$ are decreasing, the common knowledge of E_1 can be written as the intersection of the sets $K_N^m(E_1)$ for even numbers m:

$$CK(E_1) = \bigcap_{m=0}^{\infty} K_N^{2m}(E_1) = \bigcap_{m=0}^{\infty} E_{2m+1} = \varnothing.$$

That is, there is no state at which it is common knowledge that the price is not zero.

In this example, one can show that there is no non-trivial event that is common knowledge at any state. The next exercise asks you to prove it (this will be obvious by the analysis later, but I ask you to prove it by brut force.)

Exercise 8.10 *In Example 8.6, show that, for any $E \neq \Omega$,*

$$CKE = \varnothing.$$

Show also that

$$CK\Omega = \Omega.$$

The following theorem states several properties of common knowledge which in particular it verifies that CK is a particular knowledge operator (i.e. it satisfies properties 1, 2, and 3 below).

Theorem 8.2 *The common knowledge operator CK is a knowledge operator, i.e., it satisfies the following properties:*

1. $CKE \subseteq E$;
2. $E \subseteq F \implies CKE \subseteq CKF$;
3. $CKE = CKCKE$;
4. $\sim CKE = CK \sim CKE$.

That is, if one used the information that is common knowledge in a society as what is known by that society, it would look like a knowledge operator for a unitary agent. In particular, it would satisfy the truth axiom; the unitary agent would have logical omniscience, positive introspection and negative introspection. The latter two properties are important. They state that there cannot be ambiguity or private information about whether something is common knowledge. It will be common knowledge whether it is common knowledge or not. This result has many interesting corollaries. I will present two of them next.

Corollary 8.1 $CK\Omega = \Omega$.

That is the state is always common knowledge. That is, the features that are true throughout the model are common knowledge. There is a sense in which the converse is also true. If CKE is non-empty for a proper subset of E, one could define a "submodel" by considering CKE as the state space, and the submodel can be analyzed in isolation. The second corollary is highly important and useful:

Corollary 8.2 *For each $i \in N$ and event E, $K_iCKE = CKE$.*

This of course immediately follows from Property 3 in Theorem 8.2; it is often used as a lemma to prove the theorem. It states that if common knowledge of an event is a self-evident event for every player. If something is common knowledge, it wouldn't be news for any player that it is common knowledge. The events identified in this corollary are highly important, and I will use this corollary to obtain alternative, useful definitions for common knowledge next. (While the inductive definition of common knowledge is intuitive, in many situations it is hard to work with.)

8.3.4 Public Events

Corollary 8.2 identifies an important property of common knowledge events: everybody knows that the event holds whenever it indeed holds. Such events are self-evident for everybody. Such events are called public, and they play a central role in interactive epistemology. In this section, I will provide a definition of common knowledge based on the concept of common knowledge.

Definition 8.4 *An event E is said to be* public *if*

$$K_i E = E \qquad (\text{for all } i \in N).$$

Observe that a public event is common knowledge whenever it occurs. Indeed, for any public event E, since $K_i E = E$ for all $i \in N$, $K_N E = \cap_{i \in N} K_i E = E$. Hence, using mathematical induction on m, one can easily show that

$$K_N^m (E) = E$$

for all m. Therefore,

$$CK (E) = E.$$

One can view this as a justification for the term public. Corollary 8.2 shows that the converse is also true: CKE is public event for any E. The following theorem gives an alternative definition of common knowledge.

Theorem 8.3 *(1) An event E is a public event if and only if $CK(E) = E$. (2) For any event E, the set CKE is the largest public event F (in the sense of set inclusion) such that $F \subseteq E$.*

The above result shows that computing common knowledge operator boils down to finding public events. Formally, it establishes that the set of public events is the knowledge field for common knowledge operator CK. I will next present a way to compute the set of public events, based on knowledge fields.

8.3.5 Knowledge Field for Common Knowledge Operator

Since the common knowledge operator CKE is a knowledge operator, it is associated with a knowledge field—as well as an information partition and information graph. Building on Theorem 8.3, I will show that the knowledge field for the common knowledge operator is simply the intersections of the individual knowledge fields, as that intersection consists of all public events.

Recall that a definition of knowledge field $\mathcal{K}_i$ for a player i is that it is the set of all self-evident events for player i:

$$\mathcal{K}_i = \{E | K_i E = E\}.$$

On the other hand, a public event is defined as an event that is self-evident for every player i:

$$K_i E = E \qquad \text{(for all } i \in N).$$

Therefore, an event E is public if and only if it is in the intersection of all individual knowledge fields, i.e.,

$$E \in \cap_{i=1}^{n} \mathcal{K}_i.$$

But Theorem 8.3 has established that the set of public events is the knowledge field for the common knowledge operator. Therefore, the knowledge field for the common knowledge operator is the intersection of the players' knowledge fields.

Theorem 8.4 *The set of all public events is*

$$\mathcal{K}_{CK} = \cap_{i=1}^{n} \mathcal{K}_i,$$

where $\mathcal{K}_i$ is the knowledge field of player i. Moreover, the common knowledge operator is the knowledge operator associated with knowledge field $\mathcal{K}_{CK}$.

Theorem 8.2 has established that the common knowledge operator is a knowledge operator. Theorem 8.4 establishes that the knowledge field associated by the common knowledge operator is the intersection of the individual knowledge fields. Intuitively, knowledge field of a player defines her knowledge as it is the set of events that are knowable by that player. The intersection of these fields defines knowledge that is common to all these players, as the set of events knowable by every player. In models with simple structure, one can be able to compute that intersection easily, leading to a useful method when the knowledge fields are easily accessible. The next exercise asks you to do that.

Exercise 8.11 *In Example 8.6, show that* $\mathcal{K}_1 \cap \mathcal{K}_2 = \{\Omega, \varnothing\}$. *Conclude that* $CKE = \varnothing$ *whenever* $E \neq \Omega$.

8.3.6 Information Partition for Common Knowledge

One can also use Theorem 8.4 to compute the information partition $\mathcal{I}_{CK}$ associated with common knowledge operator. After all, by Theorem 8.4, $\mathcal{I}_{CK}$ is the information partition associated with the common knowledge field $\mathcal{K}_{CK} = \cap_{i=1}^{N}\mathcal{K}_i$. Using the results in Section 8.2, I will next present direct computation of $\mathcal{I}_{CK}$ and the associated information function I_{CK}.

Theorem 8.4 establishes that $\mathcal{I}_{CK}$ is the information partition associated with the common knowledge field $\mathcal{K}_{CK} = \cap_{i=1}^{N}\mathcal{K}_i$. Theorem 8.1 in Section 8.2 further establishes that the information partition associated with $\mathcal{K}_{CK} = \cap_{i=1}^{N}\mathcal{K}_i$ is the common coarsening of the information partitions $\mathcal{I}_1, \ldots, \mathcal{I}_n$. That is, $\mathcal{I}_{CK}$ is the finest information partition that is coarser than each information partition $\mathcal{I}_i$. In other words, each $\mathcal{I}_i$ contains more information than $\mathcal{I}_{CK}$, and $\mathcal{I}_{CK}$ contains all the knowledge common in all information partitions. In that sense, $\mathcal{I}_{CK}$ represents the common knowledge in the model. The key properties of $\mathcal{I}_{CK}$ and associated information function I_{CK} is listed next (these are already obtained in Section 8.2).

Theorem 8.5 *The common knowledge operator* CK *is associated with the common knowledge information partition* $\mathcal{I}_{CK}$ *and common knowledge information partition* I_{CK}, *which satisfy the following properties.*

1. $\mathcal{I}_{CK}$ *is the common coarsening of the information partitions* $\mathcal{I}_1, \ldots, \mathcal{I}_n$, *so that* $\mathcal{I}_i \geq \mathcal{I}_{CK}$ *for each* i, *in that the cells in* $\mathcal{I}_{CK}$ *are disjoint unions of the cells in* $\mathcal{I}_i$, *and it is the finest such partition.*

2. *For each* $\omega \in \Omega$, $I_{CK}(\omega)$ *is the smallest set* $\tilde{I}$ *(in the sense of set inclusion) such that*
$$\omega \in \tilde{I} = \bigcup_{\omega' \in \tilde{I}} I_i(\omega') \qquad (\forall i \in N).$$

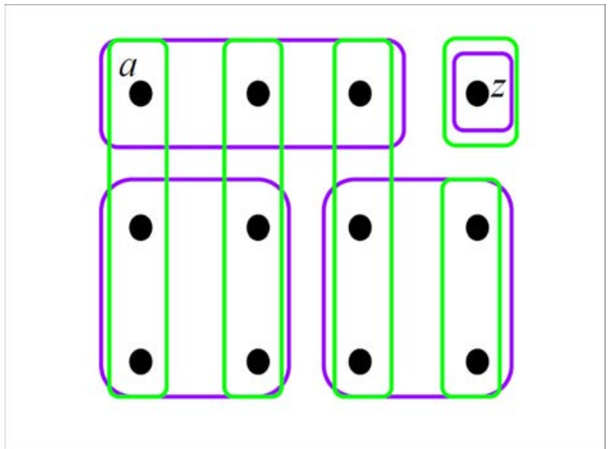

Figure 8.5: An Epistemic model with non-trivial common knowledge operator

3. *For each* $\omega, \omega' \in \Omega$, $\omega' \in I_{CK}(\omega)$ *if and only if there exists a sequence* $\omega_0, \ldots, \omega_m \in \Omega$ *such that* $\omega_0 = \omega, \omega_m = \omega'$, *and*

$$\omega_{k+1} = I_1(\omega_k) \cup \cdots \cup I_n(\omega_k) \qquad (\forall k = 0, \ldots m-1).$$

The first two properties state the fundamental—and highly useful—fact about the common knowledge: the information sets under common knowledge can be written as the disjoint union of the information sets of each individual player:

$$I_{CK}(\omega) = \bigcup_{\omega' \in I_{CK}(\omega)} I_i(\omega').$$

Moreover, this is the smallest set that can be decomposed into each player's information partitions (and contains the true state). The last property provides an algorithm to compute this set, Set $E_{0,\omega} = \{\omega\}$, and for each $k > 0$, define iteratively

$$E_{k,\omega} = \bigcup_{\omega' \in E_{k-1,\omega}} I_1(\omega') \cup \cdots \cup I_n(\omega').$$

The common knowledge information set is the limit

$$I_{CK}(\omega) = \bigcup_{k=0}^{\infty} E_{k,\omega}.$$

Example 8.9 *Consider the epistemic model depicted in Figure 8.5, where the purple horizontal information sets belong to Player 1, and the green vertical information sets belong to Player 2. First consider the state z. Observe that*

$$I_1(z) = I_2(z) = \{z\}.$$

Since the set $\{z\}$ can be written as (trivial) disjoint unions of the information sets of each player, we have

$$I_{CK}(z) = \{z\}.$$

The event $\{z\}$ is a public event, and it becomes common knowledge whenever the state of the world is z. Next consider the state $\omega = a$. The set $E_{1,a}$ in the above algorithm is the L-shaped shaded event on the right panel of Figure 8.6. This is simply the union of the information sets at $\omega = a$. In the second round, we included the information sets that intersect this event. We include the square in the lower-left corner for Player 1 and the first three rectangular vertical information sets for Player 2. The resulting set, $E_{2,a}$, is the rectangular shaded area in the middle panel. Finally, in the third round, we include the square in the lower-right corner for Player 1, obtaining the L-shaped shaded set on right panel, which is $E_{3,a}$. Observe that $E_{3,a}$ can be written as a disjoint union of information sets of each player. Consequently, the process of inclusion stops there, and

$$I_{CK}(a) = E_{3,a}.$$

This gives the information partition on the right panel as the common knowledge information partition, $\mathcal{I}_{CK}$, for this model.

8.3.7 Information Graph for Common Knowledge

In finite models, the graph G_{CK} associated with common knowledge operator is quite useful visually determining the common knowledge information partition, and it can be easily computed as follows. Consider the information graphs $G_1, \ldots, G_n$ associated with the information functions $I_1, \ldots, I_n$. Define

$$\tilde{G}_{CK} = \max\{G_1, \ldots, G_n\},$$

so that $\tilde{G}_{N,\omega,\omega'} = 1$ if and only if $G_{i,\omega,\omega'} = 1$ for some i. Note that $G_{i,\omega,\omega'} = 1$ if and only if i cannot rule out ω' at ω. By Theorem 8.5, $\omega' \in I_{CK}(\omega)$ if and only if there is a

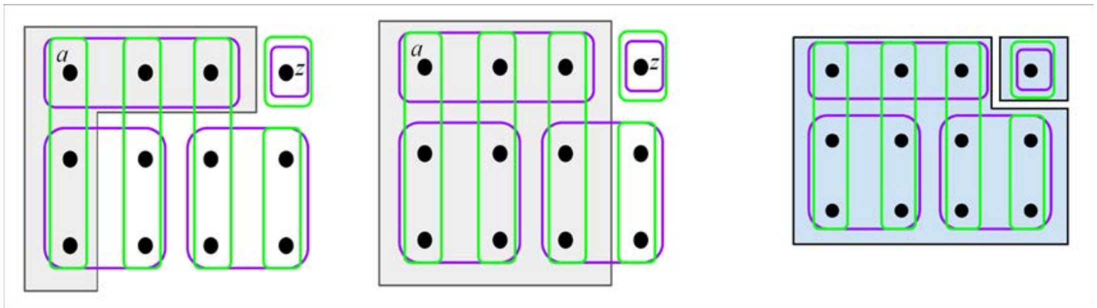

Figure 8.6: Computing common knowledge information partition in Figure 8.5.

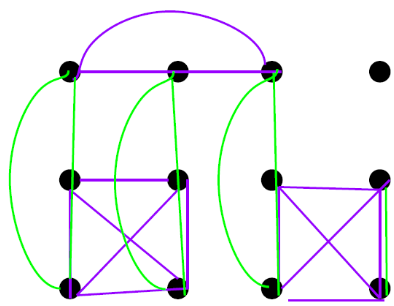

Figure 8.7: Information graphs for the model in Figure 8.5.

path that connects ω to ω' in G_{CK}. The graph $\tilde{G}_{CK}$ is not an information graph in that there some connected states may not have direct link. The information graph G_{CK} for the common knowledge is defined as the sparsest information graph larger than $\tilde{G}_{CK}$: $G_{N,\omega,\omega'} = 1$ if and only if there exists a path that connects ω to ω' in $\tilde{G}_{CK}$.

This result is quite useful in visually determining common knowledge event. Plot the information graphs of all players in one network. The information sets for common knowledge operator is simply the connected components of the resulting network.

Example 8.10 *Information graphs of the model in Figure 8.5 are plotted in Figure 8.7. In this figure, the purple edges correspond to the graph of Player 1, and the green edges correspond to the graph of Player 2. To compute the graph $\tilde{G}_{CK}$, we simply ignore the colors, and consider the whole thing as one graph. Now, clearly, the graph $\tilde{G}$ has two components: a singleton component, which corresponds to $I_{CK}(z) = \{z\}$, and the rest* *of the state space, which corresponds to* $I_{CK}(a) = \Omega \setminus \{z\}$.

8.4 Common Certainty

Common knowledge is a strong assumption: it essentially reflects the features of the model that are true throughout the model. An important weakening of common knowledge is obtained by weakening knowledge to certainty. Recall that certainty fails one important property of knowledge: the Truth Axiom, K1. To account for this failure, in definition of common certainty, one does not require that event E necessarily hold when there is common certainty of E.

Along the same lines with common knowledge, define the *mutual certainty* operator C_N by

$$C_N E = \cap_{i \in N} C_i E.$$

This is the event that everybody in group N is certain that event E holds. Now, for any order $m = 2, \ldots$, define operator C_N^m by

$$C_N^m E = C_N \left(C_N^{m-1} E \right)$$

where $C_N^1 = C_N$. The operator C_N^m is called *mth-order mutual certainty.* It corresponds to the event that everybody in group N is certain that everybody in group N is certain that ... that event E holds, where one repeats everybody in group N is certain that m times.

Common certainty of event is defined as having mutual certainty at all orders. One says that there is common certainty of event E at state ω if ω is in the event

$$CC(E) = \bigcap_{m=1}^{\infty} C_N^m(E).$$

Observe that the index starts at $m = 1$. Hence, there can be common certainty of false event, i.e., one may have $CC(E) \nsubseteq E$.

Exercise 8.12 *Which properties C1-C4 of certainty operator are satisfied by common certainty CC? (Replace K with CK when necessary.)*

Chapter 9

Common-Prior Assumption

In real life players often hold different beliefs. Such belief disagreements often have large impact on strategic outcomes. For example, litigants go through costly litigation process instead of settling their disputes; countries go to war, and many profitable partnerships do not come to existence. There are two main causes of such belief disagreements:

1. Informational Differences: the players may have different information.

2. Differing priors: the players may have different priors to start with.

Plausibly, in most situations there is a little both of each differences when there is a belief disagreement. From a modeling perspective, informational differences are important when the players attribute the belief differences to their private information, asking "What do they know that I do not know?" Differing priors are relevant when they attribute the belief disagreement to imperfections in other players' models, such as lack of information on the part of the other players, or other players' putting too much emphasis on some factors and putting too little emphasis on some others, according to our player. For example, in forming their beliefs, the players may draw on certain historical facts that they deem relevant but they may disagree on the relevance of the historical facts to the problem at hand. In these situations, players may simply recognize the differences and move on, or if there is uncertainty about the nature of the disagreement they would rather ask "What is it that she does not understand?" (Note that this does not mean that they think that the other players are irrational. They may trust each others' rationality and intelligence while not agreeing with their priors, which may reflect the parties cultural background and upbringing.)

Harsanyi (1967) argued that all belief differences should be traced back to informational differences, ruling out differing priors as a cause of belief disagreements. This view is known as the *Harsanyi Doctrine*, and formally corresponds to the *Common Prior Assumption*. This allowed Harsanyi to model games of incomplete information as complete information games with chance moves, as in a poker game.[1] This focus on incomplete information led to major advances in economic theory, especially in information economics and mechanism design.

This view has been challenged both empirically and theoretically, and eventually led to modern game theoretical models that incorporate both incomplete information and differing priors. The first empirical challenge came in the form of a mathematical theorem: Agreeing to Disagree Theorem. Aumann (1974) showed that if the players' beliefs are common knowledge, then they must be identical—and there must be no belief disagreement. This suggests that the common-prior assumption is violated in many real-world situations in which the parties disagree and the nature of their disagreement is known. An application of the idea behind the agreeing to disagree theorem led to a twin mathematical theorem challenging the common-prior assumption further: No-Trade Theorem (Milgrom and Stokey 19xx). This theorem establishes that if there were no incentive to trade, the rational players cannot have an incentive to trade when they obtain private information. In other words, adverse-selection will only reduce the scope of trade. Applying this to financial markets, one can conclude that risk sharing is the only motive for trade under the common prior assumption. The amount of trade that cannot be explained by risk-sharing must come from violations of the common prior assumptions. Empirical and experimental literature further showed the existence of systematic biases in perception of uncertain outcomes, such as optimism, overconfidence and self-serving biases, biases that can be modeled via differing priors.

This Chapter presents the above theorems. There is also a sense in which the common-prior assumption per se does not have any significant bite, and the empirical content of the common-prior assumption comes rather from the common knowledge

[1] As seen in previous chapters, the common-prior assumption is not needed to define the Bayesian games, as one can allow differing priors on the type space.

assumptions. In particular, Lipman (2003) shows that given any belief hierarchy h_i and arbitrary k, one can find a type t_i in some type space with common-prior assumption such that the first k orders of beliefs under t_i are as described by the hierarchy h_i. Hence, one cannot rule out the possibility that the common knowledge assumptions rather than the common-prior assumption per se are violated when there is disagreement or excessive trade. Later in the course, we will present Lipman's result along with Samet's characterization of the common-prior assumption via convergence of the higher-order expectations.

9.1 Agreeing to Disagree

Imagine two players who use the same decision rule, so that they would make the same decision if they had the same information. In general, when they have different information, they can naturally choose different actions. Can they choose different actions if their decisions are common knowledge (although they have possibly different information)? It turns out that the answer is No if the common decision rule satisfies the sure-thing principle, the main axiom of the expected utility maximization. Under the common-prior assumption, the rule that assigns conditional probability of an event under various pieces of information satisfies the sure-thing principle, and hence the players' beliefs cannot be common knowledge and different at the same time. This section presents this result formally.

For a set of players N, fix an information structure $(\Omega, (I_i)_{i\in N}, (\pi_{i,\omega})_{i\in N,\omega\in\Omega})$ where Ω is a finite state space, I_i is an information partition of Ω for each $i \in N$, and $\pi_{i,\omega}$ is a probability distribution on $I_i(\omega)$ for each i and ω. Recall that the common-knowledge partition I_{CK} is the finest partition of Ω that is weakly coarser than each partition I_i, so that

$$I_{CK}(\omega) = \bigcup_{\omega' \in I_{CK}(\omega)} I_i(\omega') \tag{9.1}$$

for each player i and each ω. A *decision rule* is a mapping

$$d : 2^{\Omega} \backslash \{\varnothing\} \to Y$$

that maps each non-empty subset E of the state space to a decision $d(E) \in Y$ where Y is a set of decisions. That is, if one's information set is E, then she is to choose decision

$d(E)$. A decision rule is said to satisfy the *sure-thing principle* if for any disjoint events E and F,

$$d(E) = d(F) = y \Rightarrow d(E \cup F) = y.$$

That is, if one is to choose the same action when she learns E or when she learns F then she should take the same action if she only learns that E or F occur. This is the main axiom in Savage's characterization of the expected utility maximization. For each player i, let

$$d_i : \Omega \to Y$$

be the decision she makes under the above information structure if she uses a decision rule d, where

$$d_i(\omega) = d(I_i(\omega)).$$

The next results states that if all players use the same decision rule, then they make the same decision whenever their decisions are common knowledge.

Theorem 9.1 (Agreement Theorem) *Assume that $d_i(\omega) = d(I_i(\omega))$ for each ω and i for some decision rule d that satisfies the sure-thing principle. If it is common knowledge that $d_i = y$ and $d_j = y'$ at some ω for some $i, j \in N$ and $y, y' \in Y$, then $y = y'$.*

Proof. Since it is common knowledge that $d_i = y$ at ω, $d_i(\omega') = y$ for each $\omega' \in I_{CK}(\omega)$, i.e.,

$$d(I_i(\omega')) = d_i(\omega') = y \qquad \forall \omega' \in I_{CK}(\omega).$$

Since d satisfies the sure-thing principle, by (9.1), this implies that

$$d(I_{CK}(\omega)) = y.$$

Likewise, since it is common knowledge that $d_j = y'$ at ω, $d(I_{CK}(\omega)) = y'$. Therefore, $y = y'$. ■

A typical "decision rule" is the conditional probability of a given event E^* under a prior belief Pr:

$$d^*(E) = \Pr(E^*|E).$$

By the Bayes' rule, this decision rule satisfies the sure-thing principle. To see this, suppose $\Pr(E^*|E) = \Pr(E^*|F) = q$ for some disjoint events E and F with positive probabilities. Then,

$$d^*(E \cup F) = \Pr(E^*|E \cup F) = \Pr(E^*|E)\Pr(E|E \cup F) + \Pr(E^*|F)\Pr(F|E \cup F) = q.$$

Applying the Agreement Theorem to this example, one obtains the Agreeing to Disagree Theorem.

Corollary 9.1 (Agreeing to Disagree Theorem) *Let* $(\Omega, (I_i)_{i\in N}, (\pi_{i,\omega})_{i\in N,\omega\in\Omega})$ *be an information structure with a common prior* P *where* $P(\omega) > 0$ *for every* $\omega \in \Omega$*, and fix some event* E *and players* $i, j \in N$*. If it is common knowledge at some* ω *that* $\pi_i(E) = p$ *and* $\pi_j(E) = q$*, then* $p = q$*.*

That is, under the common-prior assumption, if the players' beliefs are common knowledge, then their beliefs must be identical, ruling out belief disagreement when the beliefs are common knowledge. This does not mean however that belief disagreement cannot be common knowledge, as established by the next example.

Example 9.1 *Take* $\Omega = \{0, 1\}$ *with prior probabilities* $P(0) = P(1) = 1/2$*; take* $I_1 = \{\Omega\}$ *and* $I_2 = \{\{0, 1\}\}$*. Consider event* $E = \{0\}$*. Clearly,* $\Pr(E|I_1(\omega)) = 1/2$ *while* $\Pr(E|I_2(\omega)) \in \{0, 1\}$ *so that it is common knowledge that there is belief disagreement.*

I will nex presents stronger versions of Agreeing to Disagree Theorem. The first generalization is another corollary to the Agreement Theorem. Agreeing to Disagree Theorem above established that, under the Common-Prior Assumption, the players cannot agree to disagree about probability of a given event. More generally, the players cannot agree to disagree on the expectation of a random variable. Before stating this result formally it is useful to introduce a couple of notation that will be used in the sequel. For any model $(\Omega, (I_i)_{i\in N}, (\pi_{i,\omega})_{i\in N,\omega\in\Omega})$, a random variable X is any function $X : \Omega \to \mathbb{R}^m$ for some m; one can use more general spaces as the range as long as conditional expectations are well-defined.[2] For each $i \in N$ and $\omega \in \Omega$, define $X_i(\omega)$ as the conditional expectation of X according to player i at ω, i.e., the expectation of X under $\pi_{i,\omega}$.

Corollary 9.2 (Agreeing to Disagree Theorem for Random Variables) *Let* $(\Omega, (I_i)_{i\in N}, (\pi_{i,\omega})_{i\in$ *be an epistemic model with a common prior* P *where* $P(\omega) > 0$ *for every* $\omega \in \Omega$*, and fix some random variable* X *and players* $i, j \in N$*. If it is common knowledge at some* ω *that the conditional expectation according to* i *is* $X_i = x$ *and the conditional expectation according to* j *is* $X_j = y$*, then* $x = y$*.*

[2] We assume discrete σ-algebra, so that there is no measurability restriction on random variables.

Proof. Define the decision rule d by

$$d(E) = \mathbb{E}[X|E]$$

where $\mathbb{E}[X|E]$ denotes the conditional expectation of X with respect to event E under common prior P. Observe that d satisfies the Sure-Thing Principle, which is known as the Tower Property or the Law of Iterated Expectation in this context. Defining

$$d_i(\omega) \equiv \mathbb{E}[X|I_i(\omega)] = X_i(\omega),$$

one concludes from the Agreement Theorem that common knowledge of $d_i = X_i = x$ and $d_j = X_j = y$ implies that $x = y$. ■

Observe that Agreeing to Disagree Theorem for events is a special case of the last corollary because one can take X as the characteristic function of event E^*.

In real-world applications it is hard to verify common knowledge of each player's belief about a random variable. However, in many sitations, some aggragate statistics of their beliefs may be public information (e.g., in the form of opinion polls or prices). By taking a giant leap of faith, one may assume that such publicly available aggregate statistics is common knowledge, and ask if Agreeing to Disagree Theorem applies to such aggregate statistics. The next result states that this is indeed the case.

Theorem 9.2 *Let $(\Omega, (I_i)_{i\in N}, (\pi_{i,\omega})_{i\in N,\omega\in\Omega})$ be an epistemic model with a common prior P where $P(I_i(\omega)) > 0$ for every $i \in N$ and $\omega \in \Omega$. Let $X : \Omega \to \mathbb{R}$ be a random variable and $f : \mathbb{R}^n \to \mathbb{R}$ be a mapping with*

$$f(x_1, \ldots, x_n) = f_1(x_1) + \cdots + f_n(x) \tag{9.2}$$

for some strictly increasing functions $f_1, \ldots, f_n$. If $\omega^* \in CK(f(X_1, \ldots, X_n) = x)$, then

$$X_1(\omega^*) = \cdots = X_n(\omega^*) = \mathbb{E}[X|I_{CK}(\omega^*)],$$

where $X_i(\omega)$ denotes the expectation of X under $\pi_{i,\omega}$ and $\mathbb{E}[X|F]$ denotes the conditional expectation of X with respect to F under P.

Proof. We will focus on the set $I_{CK}(\omega^*)$ at which it is common knowledge that $f(X_1, \ldots, X_n) = x$. For each $\omega \in I_{CK}(\omega^*)$, write $X_0(\omega) = \mathbb{E}[X|I_{CK}(\omega^*)]$. By definition,

$$\mathbb{E}[X - X_0|I_{CK}(\omega^*)] = 0.$$

Multiplying both sides of this equality by x, one obtains

$$\begin{aligned} 0 &= \mathbb{E}\left[x\left(X-X_0\right)|I_{CK}\left(\omega^*\right)\right] \\ &= \mathbb{E}\left[f\left(X_1,\ldots,X_n\right)\left(X-X_0\right)|I_{CK}\left(\omega^*\right)\right] \\ &= \sum_{i\in N}\mathbb{E}\left[f_i\left(X_i\right)\left(X-X_0\right)|I_{CK}\left(\omega^*\right)\right] \\ &= \sum_{i\in N}\mathbb{E}\left[f_i\left(X_i\right)\left(X_i-X_0\right)|I_{CK}\left(\omega^*\right)\right] \end{aligned}$$

where the second equality is by substitution of $f\left(X_1\left(\omega\right),\ldots,X_n\left(\omega\right)\right)=x$ at each $\omega\in I_{CK}\left(\omega^*\right)$, the next equality is by (9.2), and the last equality is by the law of iterated expectation—as $X_i\left(\omega\right)=\mathbb{E}\left[X|I_i\left(\omega\right)\right]$ and $I_i\left(\omega\right)\subseteq I_{CK}\left(\omega^*\right)$ for each $\omega\in I_{CK}\left(\omega^*\right)$. Using $f\left(X_0,\ldots,X_0\right)$ instead of x in the above derivation, one can also obtain

$$0=\sum_{i\in N}\mathbb{E}\left[f_i\left(X_0\right)\left(X_i-X_0\right)|I_{CK}\left(\omega^*\right)\right].$$

Combining the last two displayed equalities, one further obtains

$$\sum_{i\in N}\mathbb{E}\left[\left(f_i\left(X_i\right)-f_i\left(X_0\right)\right)\left(X_i-X_0\right)|I_{CK}\left(\omega^*\right)\right]=0.$$

Now, since f_i is increasing, $\left(f_i\left(X_i\left(\omega\right)\right)-f_i\left(X_0\left(\omega\right)\right)\right)\left(X_i\left(\omega\right)-X_0\left(\omega\right)\right)\geq 0$ for each $\omega\in I_{CK}\left(\omega^*\right)$ and $i\in N$. Thus, if the last equality were strict for some ω and i, then the displayed sum of expectation would be strictly positive (because $P\left(I_i\left(\omega\right)|I_{CK}\left(\omega^*\right)\right)>0$ by hypothesis). Therefore, the displayed equation implies $\left(f_i\left(X_i\left(\omega\right)\right)-f_i\left(X_0\left(\omega\right)\right)\right)\left(X_i\left(\omega\right)-X_0\left(\omega\right)\right)=0$ for all $\omega\in I_{CK}\left(\omega^*\right)$ and $i\in N$. Since f_i is strictly increasing, this further implies that

$$X_i\left(\omega\right)=X_0\left(\omega\right)$$

for each $\omega\in I_{CK}\left(\omega^*\right)$ and $i\in N$. ■

Theorem 9.2 generalizes the Agreeing to Disagree Theorem to monotone aggregate statistics. A monotone aggregate statistic f is a separable and strictly increasing function of players' expectations, such as their weighted averages. The theorem states that if the value of such an aggregate statistic is common knowledge, then each player's conditional expectation is equal to the conditional expectation under that information:

$$X_i=E\left[X|CK\left(f\left(X_1,\ldots,X_n\right)=x\right)\right]$$

on $CK\left(f\left(X_1,\ldots,X_n\right)=x\right)$.

It is tempting to conclude that there will not be any belief disagreement when there are aggregate statistics in the form of large public opinion polls. This is not warranted. The aggregate statistics reflects the expectations before the aggregate statistics is revealed, and the players may update their beliefs in response to the aggregate statistics. Geanokoplos and Polimarchakis (198x) studies such an opinion dynamics, and show that the beliefs converge eventually.

9.2 No Trade Theorem

A version of the agreement theorem establishes that there will be no information based trade under the common-prior assumption. There are many versions of this No-Trade Theorem. This section presents two of them. The first version establishes that risk-neutral players cannot have a strict incentive to trade in any equilibrium under the common-prior assumption. The second one implies that if the initial allocation is Pareto-efficient there will be no Pareto-improving allocation in the interim stage. There will be only two players for clarity.

It is useful to start with a corollary to Agreement Theorem. Consider two risk neutral players $i\in N=\{1,2\}$. Consider any epistemic model (N,Ω,I,P) with a common prior P on Ω. A *trade* (or a bet) is a random variable

$$X:\Omega\to\mathbb{R}$$

with the understanding that Player 2 will pay $X(\omega)$ to Player 1 at state ω. Clearly, at the ex-ante stage—before the players get their private information, there is no individually rational trade with at least one player has strict preference towards trade. Indeed, if $\mathbb{E}[X]>0$, then Player 2 rejects such a trade; Player 1 rejects it if Player 2 wants it (i.e., $\mathbb{E}[X]>0$). One may think that perhaps there may be trade in the interim stage when players get their information. The next result establishes that there will not be strictly individually rational trade in the interim stage either.

Corollary 9.3 (No Trade Theorem under Risk Neutrality) *Consider any epistemic model* (N,Ω,I,P) *with a common prior* P *such that* $P\left(I_i(\omega)\right)>0$ *for each* $i\in N$ *and* $\omega\in\Omega$. *Then,*

$$CK\left(X_1\geq 0, X_2<0\right)=\varnothing$$

where $X_i(\omega) = \mathbb{E}[X|I_i(\omega)]$ *denotes the player i's interim expectation of transfer at* ω*. If* $\omega^* \in CK(X_1 \geq 0, X_2 \leq 0)$*, then* $X_1(\omega^*) = X_2(\omega^*) = 0$*.*

Proof. Let $Y = \{\text{buy, sell}\}$ be the set of decision, and the decision rule be

$$d(E) = \begin{cases} \text{buy} & \text{if } \mathbb{E}[X|E] \geq 0 \\ \text{sell} & \text{if } \mathbb{E}[X|E] < 0 \end{cases}$$

where buy and sell mean buying and selling an asset that pays $X(\omega)$ at ω, respectively; the ties are broken in favor of buy. There is trade if Player 1 buys and Player 2 sells. Observe that

$$CK(X_1 \geq 0, X_2 < 0) = CK(d_1 = \text{ buy}, d_2 = \text{ sell}).$$

But under common prior, the decision rule d satisfies the Sure-Thing Principle, and the Agreement Theorem concludes that the above set must be empty. The second statement is not strictly a corollary to the agreement theorem, but as in the proof of the Agreement Theorem, one can show that $\mathbb{E}[X|I_{CK}(\omega^*)] = 0$ and therefore $X_1(\omega) = X_2(\omega) = 0$ on $I_{CK}(\omega^*)$. ■

The previous result can be viewed as a No-Trade Theorem for risk neutral players. It states that it cannot be common knowledge that two risk neutral players are trading rationally and one of them strictly prefers trade to no trade. This statement is an immediate corollary to the Agreement Theorem because the players' decisions are common knowledge and the decisions are not equal; they are diagonally opposite to each other.

9.2.1 Examples

This sub-section presents simple examples to illustrate some of the main insights for the theoretical results in the remainder of this section. In all of the examples there is a divisible asset that pays $\theta \in \{0, 1\}$. The first example illustrates that there is no gain from trade when the players are risk neutral and have a common prior.

Example 9.2 *There are two players: Alice and Bob. Alice owns the asset, and both players assign probability 1/2 for each value of* θ*. Assume that both players are risk-neutral. If the price of the asset is* $p = 1/2$*, then the players are indifferent towards any amount of trade, and there is a continuum of Walrasian equilibria, but there is no gain from trade. At any other price, either both will want to have all of the asset or none of it, and there will not be any trade.*

When players have private information, it may appear that there is gain from trade, in that a party may strictly value an asset more than the other does. Nonetheless, under the common-prior assumption, such belief differences are all attributable to differences in information, and such trades will not be realized (assuming it remains common knowledge that the players are rational and risk neutral), as illustrated next.

Example 9.3 *Now, imagine that Alice privately observes a binary signal* $t \in \{0, 1\}$ *such that*

$$\Pr(t = \theta|\theta) = 3/4$$

for each θ*. Upon observing* t*, Alice updates her belief and assigns probability*

$$\pi_A(t) = \begin{cases} 3/4 & \text{if } t = 1 \\ 1/4 & \text{if } t = 0 \end{cases}$$

on $\theta = 1$*. Now at a price* $p \in (1/4, 1/2)$*, Alice strictly prefers to sell the asset, which improves her expected payoff by* $p - 1/4$*. At that price, without any adverse-selection, Bob would also want to buy the asset, as that trade would improve his expected payoff by* $1/2 - p$*. However, such gain from trade is illusory. It is due to the belief differences generated by Alice's private information. Knowing that Alice has observed* t *privately, Bob will not buy the asset at that price. Indeed, Bob knows that Alice would not sell the asset at price* p *if she had observed* $t = 1$*. He thus knows that if he can buy the asset at price* p*, then Alice, the sole owner of the asset, is willing to sell it at that price, and she must have observed* $t = 0$*. Taking this information into account, he also lowers his expect payoff to* $1/4$ *and rejects the trade.*

The above phenomenon is an instance of winners' curse or buyers' remors. If Bob did not think through carefully and bought the asset, then—sooner or later—would learn that Alice had adverse information and regret that he bought the asset.

When the players are risk averse, the players have an incentive to trade in order to share the risk—as shown next.

Example 9.4 *Suppose now that both players are risk-averse with constant absolute risk aversion factor* $\alpha > 0$*. There is a unique optimal allocation of risky asset, in which each player owns half of the asset, while any amount of fixed monetary transfer is consistent with optimal risk sharing. Since Walrasian equilibrium shares the risk optimally, this* *leads to a unique Wlarasian equilibrium. Indeed, for Alice to demand half of the asset, the price must be*

$$p(\pi) = \frac{1}{1 + \frac{1-\pi}{\pi} e^{\alpha/2}}$$

where π *is the probability of* $\theta = 1$. *Since* $\pi = 1/2$, *the uniqe equilibrium price is*

$$p = \frac{1}{1 + e^{\alpha/2}},$$

and each player demands half of the asset.

Once the player realize the equilibrium trade so that the allocation is Pareto-optimal, then there cannot be any further trade in equilibrium even if some of the players receive private information and would like to trade further under their private information—as illustrated next.

Example 9.5 *Now, imagine that Alice privately observes a binary signal* $t \in \{0, 1\}$ *above. If Alice observes* $t = 1$, *she values the asset more and would like to buy some of her asset back if the price remains* $p = 1/\left(1 + e^{\alpha/2}\right)$. *If she observed* $t = 0$, *she would like to sell more without a price change. But, of course, Bob would not want to trade any further even he were naive. Now imagine that, observing* $t = 1$, *Alice is willing to buy back at higher prices and the price goes up a little bit. Then, knowing that Alice would want to buy back at that price only if she observed* $t = 1$, *he would not sell even if this higher price. Indeed, he would like to buy more, and there would be excess demand. In a (rational expectations) equilibrium, the price goes up if Alice observes* $t = 1$ *and goes down if* $t = 0$, *and the equilibrium price fully reveals Alice's information. The equilibrium price is*

$$p_t = p(\pi_A(t)) = \frac{1}{1 + \frac{1-\pi_A(t)}{\pi_A(t)} e^{\alpha/2}} = \begin{cases} \frac{1}{1+\frac{1}{3}e^{\alpha/2}} & \text{if } t = 1 \\ \frac{1}{1+3e^{\alpha/2}} & \text{if } t = 0. \end{cases}$$

At the equilibrium price, each player demands half of the asset, and there is no trade from the ex-ante equilibrium allocation.

Of course, when the players have heterogenous priors, they have an incentive to trade in order to bet against each other. The next example illustrates that risk-neutral players would like to have infinite amount of bets—without any private information.

Example 9.6 *Now introduce Art and Beth. They do not own the asset, but they have heterogenous priors about the asset's dividend. In particular, Art assigns* $\pi_A = 2/3$ *on* $\theta = 1$*, while Beth assigns probability* $\pi_B = 1/3$ *on* $\theta = 1$*. Consider a future contract on the value of asset that pays* θ*, and it is traded at price* p*. Multiple copies of the contract can be traded, so that one can demand any amount of contracts, where negative amounts indicate that one sells the contract. Assume that both players are risk neutral. Then, at any price* $p \in (1/3, 2/3)$*, Art's payoff from buying* x *units of contract is* $x(\pi_A - p) = x(2/3 - p)$*, linearly increasing in* x*. Beth's payoff from selling* x *units of contract is* $x(p - \pi_B) = x(p - 1/3)$*, again linearly increasing in* x*. Therefore, they can obtain arbitrarily high payoffs by trading the contracts at that price.*

When the belief differences are solely due to heterogenous priors, they create gains from trade—in the form of bets—as illustrated by above example. Now, seeing that Beth is willing to sell a contract at price $p \in (1/3, 2/3)$, Art concludes that Beth undervalues the asset—attributing the belief difference to "Beth's bias—and keeps his belief as is. He thus sees this as an opportunity to get rich and buys as many contracts as he can.

However, the scope of speculative trade is limited by both private information and risk-aversion. The first example, illustrates this for private information.

Example 9.7 *Now, imagine that Art privately observes the binary signal* $t \in \{0, 1\}$ *above, such that* $\Pr(t = \theta|\theta) = 3/4$ *for each* θ*. Upon observing* t*, Art updates his belief and assigns probability*

$$\pi_A(t) = \begin{cases} 6/7 & \text{if } t = 1 \\ 2/5 & \text{if } t = 0 \end{cases}$$

on $\theta = 1$*. Bet would have assigned probability*

$$\pi_B(t) = \begin{cases} 3/5 & \text{if } t = 1 \\ 1/7 & \text{if } t = 0 \end{cases}$$

on $\theta = 1$ *if she observed* t*. If Art observes* $t = 0$*, he would not buy any contract at any price* $p > 2/5$*. Hence, if Bet sees that Art is willing to trade at a price* $p > 2/5$*, she concludes that Art must have observed* $t = 1$*, and updates her belief to* $\pi_B(t) = 3/5$*. She would not sell it unless* $p \geq 3/5$*. Thus, there is no trade at prices* $p \in (2/5, 3/5)$*.*

The next example illustrates that risk-aversion limits the scope of speculative trade.

Example 9.8 *Suppose that both Art and Beth are risk-averse with constant absolute risk-aversion $\alpha > 0$. At any price p, each player i demands*

$$x_i = \frac{1}{\alpha}\left[\ln\left(\frac{1-p}{p}\right) - \ln\left(\frac{1-\pi_i}{\pi_i}\right)\right] \tag{9.3}$$

units of contract, where π_i is the probability she assigns on $\theta = 1$. In equilibrium, the markets clear, so that $x_B = -x_A$, and the unique equilibrium price is

$$p(\pi_A, \pi_B) = \frac{1}{1 + \sqrt{\frac{1-\pi_A}{\pi_A}\frac{1-\pi_B}{\pi_B}}}.$$

Since $\pi_A = 2/3$ and $\pi_B = 1/3$, the equilibrium price is

$$p(\pi_A, \pi_B) = 1/2.$$

The amount of trade is

$$x_A = \frac{1}{\alpha}\ln(2).$$

The equilibrium trade is decreasing in risk-aversion and goes to zero as the players become extremely risk-averse.

Now imagine that (risk-averse) Art and Beth traded $x_A = \frac{1}{\alpha}\ln(2)$ contracts, reaching to a Pareto-efficient allocation. It turns out that the players cannot have any further trade if they received private information as above.

Example 9.9 *Imagine that, at equilibrium allocation above, Art observes the signal t above. Then, in the rational expectation equilibrium, the price changes reflect Art's private information, but there will not be trade. In particular, if Art observes $t = 1$, then the price becomes*

$$p(\pi_A(1), \pi_B(1)) = \frac{1}{1 + \sqrt{\frac{1-\pi_A(1)}{\pi_A(1)}\frac{1-\pi_B(1)}{\pi_B(1)}}} = \frac{1}{1 + \sqrt{\frac{1-6/7}{6/7}\frac{1-3/5}{3/5}}} = \frac{3}{4}.$$

At this price, given her updated beliefs, by (9.3), Art demands

$$x_A(1) = \frac{1}{\alpha}\ln(2),$$

his initial equilibrium allocation. Likewise, Beth also demands her initial equilibrium allocation, and there will not be trade. Similarly, when Art observes $t = 0$, then the price becomes

$$p(\pi_A(0), \pi_B(0)) = \frac{1}{4},$$

and there is no trade.

The lack of further trade here is due to the fact that the players' beliefs are *concordant*, in that the players agree on the conditional distributions of the signals, i.e., both assign probability $p(t = \theta|\theta) = 3/4$. Hence, players do not have incentive to bet on the value of t when they can bet on θ. This is a key insight for Milgrom-Stokey No Trade Theorem. If they have differing priors on that conditional distribution as well, they would have an incentive to trade—before the information arrives—but that opens another can of worms in the form of incentive compatibility, as t is observed privately.

9.2.2 No Trade Theorem under Risk Neutrality

Building on Corollary 9.3, this section presents a version of this No-Trade Theorem for risk neutral players.

Consider two risk neutral players $i \in N = \{1, 2\}$ with assets whose dividends depend on some state $\theta \in \Theta$; there are finitely many states. There is a trading mechanism that results in a state-dependent trade based on players' actions. In particular, each player i has a set A_i of actions with a special element $\bar{a}_i$ that ensures no trade. Trading rule is a function $\tau : A \times \Theta \to \mathbb{R}$ that transfers $\tau(a_1, a_2, \theta)$ from player 2 to player 1 at state ω if players 1 and 2 take actions a_1 and a_2, respectively. Moreover, $\tau(a_1, a_2, \theta) = 0$ if $a_i = \bar{a}_i$ for some i. Under a fixed type space (Θ, T, p), this yields a Bayesian game $\mathcal{B}$.

As illustrated in the previous section, in general, when players have heterogeneous priors, they will have a strict preference to trade, by betting against each other. Thus, without a common prior, there can be trade, and when players do not have any private information, they would like to have the maximum amount of trade possible. Without any restriction on trade, this leads to infinite bets that result in infinite improvement in expected payoffs. Under the common-prior assumption, this channel for trade is closed, and there will be no trade between risk neutral players:

Theorem 9.3 (No Trade Theorem—Equilibrium) *Assume that the game $\mathcal{B}$ above has a common prior that puts positive probability on each state. Under any correlated equilibrium $((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ for $\mathcal{B}$, each player i is indifferent between no trade and her equilibrium action at every ω with positive probability:*

$$E_{i,\omega}(\tau(\boldsymbol{a}, \boldsymbol{\theta})) = 0.$$

Proof. The proof is identical to the proof of the Agreement Theorem. Now, for player 1, since she can obtain 0 by playing $\bar{a}_1$,

$$E_{1,\omega}(\tau(\boldsymbol{a},\boldsymbol{\theta})) \geq 0 \qquad \forall \omega \in \Omega.$$

Suppose that there is a state $\hat{\omega}$ at which her expected transfer is strictly positive:

$$E_{1,\hat{\omega}}(\tau(\boldsymbol{a},\boldsymbol{\theta})) > 0.$$

Then, at the ex ante stage, the expected transfer is strictly positive:

$$E(\tau(\boldsymbol{a},\boldsymbol{\theta})) = E(E_{1,\omega}(\tau(\boldsymbol{a},\boldsymbol{\theta}))) > 0.$$

Likewise, by rationality of player 2, ex-ante expected transfer is non-positive:

$$E(\tau(\boldsymbol{a},\boldsymbol{\theta})) = E(E_{2,\omega}(\tau(\boldsymbol{a},\boldsymbol{\theta}))) \leq 0,$$

leading to a clear contradiction. ■

That is, under the common prior assumption, there will be no Pareto improving trade provided that it is common knowledge that the players are rational. It is tempting to conjecture that the same result obtains for rationalizability. This is not true, however, because the common-prior assumption is required for both underlying uncertainty about the parameters and the strategic uncertainty. This is illustrated in the following example.

Example 9.10 *There is no uncertainty about the assets, i.e.,* $|\Theta| = 1$, *and the trading mechanism is as follows:*

	Head	*Tail*	$\bar{a}_2$
Head	*1*	*-1*	*0*
Tail	*-1*	*1*	*0*
$\bar{a}_1$	*0*	*0*	*0*

As in the matching-penny game with outside options, each player has three actions, Head, Tail, and $\bar{a}_i$*, where the last action ensures no trade. If both players choose from Head and Tail, the transfer from player 2 to player 1 is positive if they choose the same action and negative otherwise. Clearly, all three actions are rationalizable, involving trade. In particular, there can be a model with common knowledge of rationality in which the players trade and obtain strictly positive expected payoff everywhere. To see this,* *consider $\Omega = \{Head, Tail\}^2$ where players 1 and 2 know the first and second components, respectively. Player 1 assigns probability 1 to (Head, Head) when her type is Head and probability 1 to (Tail, Tail) when her type is Tail. On the other hand, Player 2 assigns probability 1 to (Tail, Head) when her type is Head and probability 1 to (Head, Tail) when her type is Tail. Each player's action is equal to her or her type. Clearly, at each state, the expected payoff of each player is 1 and thus the player is rational.*

9.2.3 No-Trade Theorem under Ex-ante Optimality

Thgis sectionpresentgs another version of the No-Trade Theorem, stated in terms of Pareto optimality for general risk preferences. This version of the No Trade Theorem does not require the common prior assumption per se, and I will state it without assuming a common prior. In general, the no trade in this result is attributed to the common-prior assumption as there would be a Pareto-improving trade, based on mutual bets, at ex-ante stage without a common prior.

Consider n players $i \in N = \{1, \ldots, n\}$, endowed with assets $\bar{x}_i : \Omega \to X$ for some finite state space Ω and consumption space X. Each player i has von-Neumann and Morgenstern utility function $u_i : X \to \mathbb{R}$ and has a prior belief $P_i \in \Delta(\Omega)$ with support Ω. Each player i also has information partition I_i of Ω. The set of feasible allocations is the set of functions $(x_1, x_2) : \Omega \to X^2$ such that

$$x_1(\omega) + \cdots + x_n(\omega) = \bar{x}_1(\omega) + \cdots + \bar{x}_n(\omega)$$

at each ω. An allocation x is said to be *ex-ante Pareto optimal* if there does not exist a feasible allocation y with $E_i(u_i(y_i)) \geq E_i(u_i(x_i))$ for each i where at least one of these inequalities are strict. The No-Trade Theorem in this setup is as follows.

Theorem 9.4 (No-Trade Theorem—Pareto Optimality) *Assume that the initial allocation $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ is ex-ante Pareto optimal. If it is common knowledge at some ω_0 that each player's interim expected payoff from some feasible allocation x is at least as high as $\bar{x}$ (i.e., $E_{i,\omega}(u_i(x_i)) \geq E_{i,\omega}(u_i(\bar{x}_i))$ for all $\omega \in I_{CK}(\omega_0)$), then it is common knowledge at ω_0 that the players are indifferent between x and the initial allocation $\bar{x}$ (i.e., $E_{i,\omega}(u_i(x_i)) = E_{i,\omega}(u_i(\bar{x}_i))$ for all $\omega \in I_{CK}(\omega_0)$).*

Proof. Suppose that $E_{i,\omega_1}(u_i(x_i)) > E_{i,\omega_1}(u_i(\bar{x}_i))$ for some $\omega_1 \in I_{CK}(\omega_0)$ for some player i. Consider the allocation y where

$$y(\omega) = \begin{cases} x(\omega) & \text{if } \omega \in I_{CK}(\omega_0) \\ \bar{x}(\omega) & \text{otherwise.} \end{cases}$$

Clearly, y is feasible, and by construction, $E_i(u_i(y_i)) > E_i(u_i(\bar{x}_i))$ and $E_j(u_j(y_j)) \geq E_j(u_j(\bar{x}_j))$ for $j \neq i$, showing that the initial allocation is not ex-ante Pareto optimal. ■

This result shows that there cannot be trade based on information; if there were a Pareto-improving trade based on information, they could realize such a gain at the ex-ante stage via a state-dependent trade contract. In this result it is crucial that there are no cross-state restrictions on trade agreements. If there were cross-state restrictions on trades (e.g. if they could only trade the physical assets), then the players may need to wait for information to realize some Pareto-improving trades, as illustrated next.

Example 9.11 *The players are Alice and Bob. They are risk neutral. Alice owns an asset that pays $\theta \in \{0,1\}$. At the interim stage both players publicly observe a signal t. The players' beliefs about (θ, t) are as follows (Alice's belief is on the left, and Bob's belief is on the right):*

$t\backslash\theta$	1	0
1	1/3	1/6
0	1/6	1/3

$t\backslash\theta$	1	0
1	1/6	1/3
0	1/3	1/6

Ex-ante, each player assign probability 1/2 on $\theta = 1$. Hence, they have no incentive to trade the asset. However, they do disagree about how the asset return is related to signal. Alice thinks that they are positively correlated while Bob thinks that they are negatively correlated. Hence, if they could bet on the (θ, t), they would write contracts that would transfer money from Bob to Alice when $\theta = t$ and from Alice to Bob when $\theta \neq t$. Unfortunately, they could not bet. Hence, they wait for the interim stage to trade the asset instead. Indeed, at the interim stage, observing t, Alice assigns probability $2/3$ on $\theta = t$ while Bob assigns probability $1/3$ on $\theta = t$. If they observe $t = 1$, Alice is more optimistic about the asset return and keeps the asset. If they observe $t = 0$, then Bob is more optimistic about the asset return and they would be willing to trade the asset at some price $p \in (1/3, 2/3)$.

In this example, the players would not have an incentive to trade at the ex-ante stage if they could write contracts on θ (but not on (θ, t)), as they agree on the marginal distribution of θ. In that sense, the initial allocation is Pareto-efficient with respect to θ-contingent trades, although it is not Pareto efficient more generally. This is closely related to the fact that they disagree on the statistical relationship between θ and t. This is formally established next.

Formally, let

$$\Omega = \Theta \times T$$

where Θ is a set of payoff-relevant parameters θ, while T is a set of payoff-irrelevant signals t, which may be informative about θ. The initial allocation $\bar{x}$ depends only on θ. Assume also that the space X of possible consumoptions is a convex set.

Definition 9.1 *A* trade *is any $\tau = x - \bar{x}$ where $x : \Omega \to X^n$ a feasible allocation, i.e., $\sum_{i \in N} \tau_i = 0$. A trade τ is said to be θ-contingent if $\tau(\theta, t) = \tau(\theta, t')$ for any θ, t, t'. An allocation x is said to be Pareto-efficient with respect to θ-contінget trade if there does not exists any θ-contingent trade τ such that $x + \tau$ Pareto-dominates x.*

Clearly, every Pareto-efficient allocation is Pareto-efficent with respect to θ-contingent trades. The previous example illustrates that the converse is not true. In that example, the initial allocation, where $\bar{x}(\theta, t) = (\theta, 0)$, is Pareto-efficient with respect to θ-contingent trades but not Pareto efficient, as the allocation y with

$$y(\theta, t) = \begin{cases} (2\theta, -\theta) & \text{if } \theta = t \\ (0, \theta) & \text{otherwise} \end{cases}$$

Pareto-dominates $\bar{x}$. This is because betting on the outcome of t in addition to θ provides additional beneficial trading opportunities. In contrast, in the examples in Section 9.2.1, betting on the outcome of t did not provide additional beneficial trading opportunities. The contrast arises from the fact that the players agreed on the conditional distributions of signals in Section 9.2.1 but not in the previous example.

Definition 9.2 *Beliefs are said to be* concordant *if*

$$P_1(t|\theta) = \cdots = P_n(t|\theta) \qquad (\forall \theta, t).$$

That is, the players agree on the statistical relationship between the payoff relevant parameters and the signals. Any belief difference between the players is attributed to their belief differences about the payoff releval parameters. For example, in Examples 9.6-9.7, although Art and Beth hade differing priors, the belief differences were all attributed to the belief difference about θ, as we had $\Pr(t = \theta|t) = 3/4$ for both players. Their beliefs were concordant. In contrast, in Example 9.11, the players disagreed on the conditional distribution of the signals:

$$\begin{aligned} P_A(t = \theta|\theta) &= 2/3 \\ P_A(t = \theta|\theta) &= 1/3 \end{aligned}$$

for each $\theta, t \in \{0, 1\}$. The beliefs were not concordant, leading to gains from trades that condition on (θ, t). The next result shows that Pareto-efficiency is equivalent to Pareto-efficiency with respect to θ-contingent trades when the beliefs are concordant.

Fact 9.1 *Assume that the beliefs are concordant, and the players are (weakly) risk-averse. Initial allocation $\bar{x}$ is Pareto efficient if and only if it is Pareto-efficient with respect to θ-contingent trades.*

Proof. Assume that initial allocation $\bar{x}$ is not Pareto-efficient, so that there exists a feasible allocation y that Pareto-dominates x. Now, let $\tau = y - \bar{x}$, and define θ-contingent trade τ^* by

$$\tau^*(\theta) = \mathbb{E}_1[\tau|\theta] \qquad (\forall\theta).$$

Since X is convex, τ^* is a feasible trade. Since the beliefs are concordant, $\tau^*(\theta) = \mathbb{E}_i[\tau|\omega]$ for all i and θ. Observe that, for each player i, since player i is risk-averse and $\bar{x}$,

$$\mathbb{E}_i[u_i(\bar{x} + \tau^*)|\theta] = u_i(\bar{x}(\theta) + \tau^*(\theta)) \geq \mathbb{E}_i[u_i(\bar{x} + \tau)|\theta],$$

where the equality is by the fact that $\bar{x} + \tau^*$ is θ-contingent, and the inequality is by Jensen's inequality and by definition of τ^*. Thus,

$$\mathbb{E}_i[u_i(\bar{x} + \tau^*)] \geq \mathbb{E}_i[u_i(\bar{x} + \tau)]$$

for each player i, showing that θ-contingent trade τ^* Pareto dominates $\bar{x}$.[3] ■

[3]Observe that the inequalities would be strict if players were strictly risk-averse and τ were not θ-contingent with positive probability.

This leads to the following No-Trade Theorem for (weakly) risk-averse players with concordant beliefs.

Corollary 9.4 *Assume that the beliefs are concordant; the players are (weakly) risk-averse; and the initial allocation $\bar{x}$ is Pareto-efficient with respect to θ-contingent trades. If it is common knowledge at some ω_0 that each player's interim expected payoff from some feasible allocation x is at least as high as $\bar{x}$ (i.e., $E_{i,\omega}(u_i(x_i)) \geq E_{i,\omega}(u_i(\bar{x}_i))$ for all $\omega \in I_{CK}(\omega_0)$), then it is common knowledge at ω_0 that the players are indifferent between x and the initial allocation $\bar{x}$ (i.e., $E_{i,\omega}(u_i(x_i)) = E_{i,\omega}(u_i(\bar{x}_i))$ for all $\omega \in I_{CK}(\omega_0)$). If, in addition, the players are strictly risk averse, then it is common knowledge at ω_0 that $x = \bar{x}$.*

This is a version of the celebrated Milgrom-Stokey No-Trade Theorem, in which the players do not have any private information initially. It considers risk-averse players with concordant beliefs. It states that if players do not have any incentive to make θ-contingent trades at the ex-ante stage, then it cannot be common kowledge that they all willingly trade on receiving private information and some of them enjoy strict gains from trade. In particular, if the players are strictly risk averse, then it cannot be common knowledge that all players rationally agree on a trade. Milgrom and Stokey further allow players to have private information initially as follows.

Theorem 9.5 (Milgrom-Stokey No-Trade Theorem) *Assume that the beliefs are concordant and the players are (weakly) risk-averse. Imagine that and the initial allocation $\bar{x}$ is Pareto-efficient with respect to θ-contingent trades. If it is common knowledge at some ω_0 that each player's interim expected payoff from some feasible allocation x is at least as high as $\bar{x}$ (i.e., $E_{i,\omega}(u_i(x_i)) \geq E_{i,\omega}(u_i(\bar{x}_i))$ for all $\omega \in I_{CK}(\omega_0)$), then it is common knowledge at ω_0 that the players are indifferent between x and the initial allocation $\bar{x}$ (i.e., $E_{i,\omega}(u_i(x_i)) = E_{i,\omega}(u_i(\bar{x}_i))$ for all $\omega \in I_{CK}(\omega_0)$). If, in addition, the players are strictly risk averse, then it is common knowledge at ω_0 that $x = \bar{x}$.*

9.2.4 No-Trade Theorem under Incentive Compatibility

9.3 Exercises

Exercise 9.1 *Assuming a common prior on a finite state space with full support, show that if it is common knowledge that a player assigns weakly higher probability to an event than another player, then it is common knowledge that they both assign the same probability to the event.*

Chapter 10

Epistemic Foundations of Solution Concepts

10.1 Normal-Form Games

10.1.1 Rationalizability

Many common knowledge assumptions are embedded in the description of a game $G = (N, S, u)$: it is common knowledge that the set of players is N; it is common knowledge that each player i can play any strategy in set S_i (an assumption which contains many other assumptions on the information structure), and it is common knowledge that each player i chooses a strategy s_i that maximizes expected value of u_i under her belief about the other players strategies (i.e., player i is rational). This section establishes that rationalizability captures precisely the strategic implications of all these assumptions. In particular, it characterizes the strategies that are consistent with common knowledge of rationality.

It is useful to recall the definition of rationalizability, which can also be defined as the result of iterated elimination of strictly dominated strategies.

Definition 10.1 (Rationalizability) *Set $S^0 = S$, and for every $m > 0$, set*

$$S_i^m = B_i\left(\Delta\left(S_{-i}^{m-1}\right)\right)$$

as the set of all strategies that are best responses to beliefs that put positive probability only on the strategies $s_{-i} \in S_{-i}^{m-1}$. For any player i, a strategy is said to be rationalizable *if* $s_i \in S_i^\infty$ *where*

$$S_i^\infty = \bigcap_{m \geq 0} S_i^m.$$

Note that, for any m, a strategy s_i is in S_i^m if and only if it is rationally played by i in a situation in which (1) i is rational, (2) i knows that every player is rational, (3) i knows that everybody knows that every body is rational, and ... (m) i know that every body knows that ... everybody knows that everybody is rational. That is, s_i is a best response to a belief μ_{-i}^1 such that every s_j^1 in the support of μ_{-i}^1 is a best response to some belief μ_{-j}^2 such that every every s_k^2 in the support of μ_{-j}^2 is a best response to some belief μ_{-k}^3 ... up to order m. It is in that sense S^m is the set of strategy profiles that are consistent with mth-order mutual knowledge of rationality. Rationalizability corresponds to the limit of the iterative elimination of strictly-dominated strategies.

The proofs in this section will use the fixed-point definition for rationalizability. Recall that fixed-point definition uses the following property.

Definition 10.2 *A set* $Z = Z_1 \times \cdots \times Z_n \subseteq S$ *is said to* have best-response property *(or to be* closed-under rational behavior*) if for each* $i \in N$,

$$Z_i \subseteq B_i\left(\Delta\left(Z_{-i}\right)\right);$$

i.e., every $z_i \in Z_i$ *is a best response to some belief* $\mu \in \Delta\left(Z_{-i}\right)$ *that puts zero probability outside of* Z_{-i}.

Theorem 2.3 establishes that S^∞ is the largest set with best-response property, providing a fixed-point definition for rationalizability. The theorem is replicated below:

Theorem 10.1 *If* Z *has best-response property, then* $Z \subseteq S^\infty$*. Moreover, under Assumption 1.1,* S^∞ *has the best response property.*

I will now formalize the idea that rationalizability captures the implications of common knowledge of rationality precisely. I have so far considered an abstract information structure for players N. In order to give a strategic meaning to the states, we also need to describe what players play at each state by introducing a strategy profile $\boldsymbol{s} : \Omega \to S$.

Definition 10.3 *A strategy profile* $\boldsymbol{s} : \Omega \to S$ *with respect to* $(\Omega, (I_i)_{i\in N}, (p_{i,\omega})_{i\in N,\omega\in\Omega})$ *is said to be* adapted *if* $\boldsymbol{s}_i(\omega) = \boldsymbol{s}_i(\omega')$ *whenever* $I_i(\omega) = I_i(\omega')$.

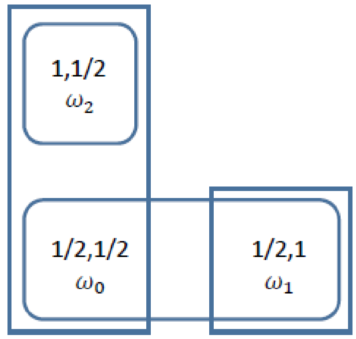

Figure 10.1: An information structure

The last condition on the strategy profile ensures that each player knows what she is playing. The possibility that $\boldsymbol{s}_i(\omega) \neq \boldsymbol{s}_i(\omega')$ for some $I_i(\omega) = I_i(\omega')$ would contradict the fact $\boldsymbol{s}_i(\omega)$ is what player i plays at state ω and that he cannot distinguish the states ω and ω' when $I_i(\omega) = I_i(\omega')$.

Definition 10.4 *An* epistemic model for a game $G = (N, S, u)$ *is a pair* $M = (\Omega, (I_i)_{i\in N}, (p_{i,\omega})_{i\in N,\omega\in\Omega}, \boldsymbol{s})$ *of an information structure and an adapted strategy profile* $\boldsymbol{s}$ *with respect to the information structure.*

The states ω in the information structure are abstract objects that describe the players' information and beliefs about those states. The mapping $\boldsymbol{s} : \Omega \to S$ gives a strategic meaning for those states and relates it to the game G. Now, at each state ω, each player i plays strategy $s_i(\omega)$. At any state ω, each player i has also a belief

$$p_{i,\omega} \circ \boldsymbol{s}_{-i}^{-1}$$

about the other players' strategies; this is the belief induced by the belief $p_{i,\omega}$ about the states and the mapping $\boldsymbol{s}_{-i}$. Consequently, each state also describe a belief and information hierarchy about the players' strategy. Therefore, the model describes what players play and what they think about the other players and so on when they make their decision, implicitly describing their rationale.

Example 10.1 *Consider the game*

$$\begin{array}{c|c|c|} & a & b \\ \hline a & 5,1 & 0,0 \\ \hline b & 4,4 & 1,5 \\ \hline \end{array} \tag{10.1}$$

and the information structure in Figure 10.1. Define $\boldsymbol{s}$ *by*

$$\boldsymbol{s}(\omega_0) = (b,a), \ \boldsymbol{s}(\omega_1) = (b,b), \ \boldsymbol{s}(\omega_2) = (a,a).$$

Note that the strategy of each player remains constant over his information sets (e.g. $\boldsymbol{s}_1(\omega_0) = \boldsymbol{s}_1(\omega_1) = b$*). Hence,* $\boldsymbol{s}$ *is adapted. In the figure, the information sets of players 1 and 2 are depicted by rounded and regular rectangles, respectively. For example, at* ω_0*, player 1 finds* ω_1 *possible and rules out* ω_2*. Similarly, player 2 rules out* ω_1*. At each state* ω*, the probabilities* $p_{1,\omega}(\omega)$ *and* $p_{2,\omega}(\omega)$ *that players 1 and 2 assign to the true state* ω *are depicted in the figure, in the given order. For example, at state* ω_1*, player 1 knows that the state is* ω_1 *and assigns probability 1 on* ω_1*, while player 2 assigns probability 1/2 on* ω_1 *and probability 1/2 on* ω_0*. In this epistemic model, the players have beliefs about the other players' strategies as well. For example, at* ω_0*, Player 1 plays b believing that Player 2 plays a with probability 1/2 and b with probability 1/2. She also believes that in the case Player 2 plays a, he believes that Player 1 plays a with probability 1/2 and b with probability 1/2, and so on.*

The ideas of rationality and common knowledge of it can be formalized as follows.

Definition 10.5 *For any epistemic model* $M = (\Omega, (I_i)_{i\in N}, (p_{i,\omega})_{i\in N,\omega\in\Omega}, \boldsymbol{s})$ *for a game* $G = (N, S, u)$*, a player i is said to be rational at a state* $\omega \in \Omega$ *if*

$$\boldsymbol{s}_i(\omega) \in B_i\left(p_{i,\omega} \circ \boldsymbol{s}_{-i}^{-1}\right),$$

where $B_i(\mu)$ *finds all of the best responses for player i to belief* μ *on* S_{-i}*.*

That is, $\boldsymbol{s}_i(\omega)$ is a best response to $\boldsymbol{s}_{-i}$ under player i's belief at ω. When Ω is finite, the condition can be written more transparently as

$$\boldsymbol{s}_i(\omega) \in \arg\max_{s_i \in S_i} \sum_{\omega' \in I_i(\omega)} u_i(s_i, \boldsymbol{s}_{-i}(\omega'))p_{i,\omega}(\omega').$$

Example 10.2 (Example 10.1, continued) *In the previous example, every player is rational at every state. For example, at state ω_0, Player 1 assigns equal probabilities on states ω_0 and ω_1, assigning equal probabilities on $\boldsymbol{s}_2(\omega_0) = a$ and $\boldsymbol{s}_2(\omega_1) = b$. Thus, she is indifferent between a and b:*

$$\begin{aligned} E\left[u_1\left(a, \boldsymbol{s}_2\right)|I_1\left(\omega_0\right)\right] &= \frac{1}{2}\cdot 5 + \frac{1}{2}\cdot 0 = 5/2 \\ E\left[u_1\left(b, \boldsymbol{s}_2\right)|I_1\left(\omega_0\right)\right] &= \frac{1}{2}\cdot 4 + \frac{1}{2}\cdot 1 = 5/2. \end{aligned}$$

Therefore, she is rational at ω_0 (and ω_1). At ω_2, she assigns probability 1 on ω_2, and thereby on $\boldsymbol{s}_2(\omega_2) = a$. She is clearly rational, as her expected payoff from a is 5 and her expected payoff from b is 4.

Let's write

$$R_i = \{\omega | \text{player } i \text{ is rational at } \omega\}$$

for the event that corresponds to the rationality of player i. It is common knowledge that player i is rational at ω if and only if event R_i is common knowledge at ω. We say that it is common knowledge in model M that player i is rational if $R_i = \Omega$. In the example above it is common knowledge that everybody is rational.

Definition 10.6 *For any $i \in N$, a strategy $s_i \in S_i$ is said to be* consistent with common knowledge of rationality *if there exists a model $M = (\Omega, (I_j)_{j\in N}, (p_{j,\omega})_{j\in N, \omega\in\Omega}, \boldsymbol{s})$ with state ω^* at which it is common knowledge that all players are rational and $\boldsymbol{s}_i(\omega^*) = s_i$.*

Since one can construct a model with state space $I_{CK}(\omega^*)$, this is equivalent to saying that there exists a model M such that $\boldsymbol{s}_j(\omega')$ is a best response to $\boldsymbol{s}_{-j}$ at each $\omega' \in \Omega$ for each player $j \in N$. For example, in game (10.1), each strategy is consistent with common knowledge of rationality because each strategy is played at some state ω in the epistemic model above, where rationality is common knowledge.

The next result states that rationalizability is equivalent to common knowledge of rationality in the sense that S_i^∞ is precisely the set of strategies that are consistent with common knowledge of rationality.

Theorem 10.2 *For any player $i \in N$ and any $s_i \in S_i$, s_i is consistent with common knowledge of rationality if and only if s_i is rationalizable (i.e. $s_i \in S_i^\infty$).*

Proof. ($\Longrightarrow$) First, take any s_i that is consistent with common knowledge of rationality. Then, there exists a model $M = (\Omega, (I_j)_{j\in N}, (p_{j,\omega})_{j\in N,\omega\in\Omega}, \boldsymbol{s})$ with a state $\omega^* \in \Omega$ such that $\boldsymbol{s}_i(\omega^*) = s_i$ and for each j and ω,

$$\boldsymbol{s}_j(\omega) \in \arg\max_{s_j \in S_j} B_j\left(p_{j,\omega} \circ \boldsymbol{s}_{-j}^{-1}\right). \tag{10.2}$$

Define Z by setting $Z_j = \boldsymbol{s}_j(\Omega)$ for each $j \in N$. By Theorem 10.1, in order to show that $s_i \in S_i^\infty$, it suffices to show that $s_i \in Z_i$ and Z has best response property. First part is immediate, as $s_i = \boldsymbol{s}_i(\omega^*) \in s_i(\Omega) = Z_i$. To see the second part, for each $z_j \in Z_j$, noting that $z_j = s_j(\omega)$ for some $\omega \in \Omega$, define belief $\mu_{-j,\omega}$ on $Z_{-j} = \boldsymbol{s}_j(\Omega)$ by setting

$$\mu_{-j,\omega} = p_{j,\omega} \circ \boldsymbol{s}_{-j}^{-1}. \tag{10.3}$$

(By definition $\mu_{-j,\omega}$ is a probability distribution on Z_{-j}.) Then, by (10.2),

$$z_j = \boldsymbol{s}_j(\omega) \in \arg\max_{s_j\in S_j} B_j\left(p_{j,\omega} \circ \boldsymbol{s}_{-j}^{-1}\right) = \arg\max_{s_j\in S_j} B_j\left(\mu_{-j,\omega}\right),$$

showing that Z has best-response property. Therefore, $s_i \in S_i^\infty$.

($\Longleftarrow$) Conversely, since S^∞ has best-response property, for every $s_j \in S_j^\infty$, there exists a probability distribution μ_{-j,s_j} on S_{-j}^∞ against which s_j is a best response. Define model

$$M^* = (S^\infty, (I_j, p_{j,s})_{j\in N, s\in S^\infty}, \boldsymbol{s}) \tag{10.4}$$

with

$$\begin{aligned} I_j(s) &= \{s_j\} \times S_{-j}^\infty \qquad \forall s \in S^\infty \\ p_{j,s}(s') &= \mu_{-j,s_j}\left(s'_{-j}\right) \qquad \forall s' \in I_j(s) \\ \boldsymbol{s}(s) &= s \qquad \forall s \in S^\infty \end{aligned}$$

In model M^* it is common knowledge that each player j is rational. Indeed, for each $s \in S^\infty$,

$$\begin{aligned} \boldsymbol{s}_j(s) &= s_j \in \arg\max_{s'_j\in S_j} \sum_{s_{-j}\in S_{-j}^\infty} u_j\left(s'_j, s_{-j}\right)\mu_{-j,s}\left(s'_{-j}\right) \\ &= \arg\max_{s'_j\in S_j} \sum_{s'\in I_j(s)} u_j\left(s'_j, s_{-j}\right) p_{j,s}(s'), \end{aligned}$$

where the equalities are by the definition of M^*, and the inclusion is by the definition of $\mu_{-j,s}$. Of course, for every $s_i \in S_i^\infty$, there exists $s = (s_i, s_{-i}) \in S^\infty$ such that $s_i(s) = s_i$, showing that s_i is consistent with common knowledge of rationality. ■

Theorem 10.2 establishes that the rationalizable strategies are precisely the ones that are consistent with common knowledge of rationality. In particular, there is a model M^* in which it is common knowledge that players are rational and every rationalizable strategy is played at some state. Hence, all rationalizable strategies are rationalized by the same model. In this model, the states are indeed rationalizable strategy profile and the players have belief about purely what strategy each player plays and what they believe about the other players' strategies and so on. As such, model M^* can be viewed as a theory of how people behave and how they rationalize their behavior. In this model, players' can hold somewhat arbitrary beliefs about the other players' strategies. In the next section, we will study Nash equilibrium, which assumes that players know the other players' strategies.

10.1.2 Correlated Equilibrium

This section establishes that correlated equilibrium characterizes the strategies that are consistent with common knowledge of rationality and a common prior on epistemic states. It is useful to recall the relevant definitions.

Definition 10.7 *An information structure* $(\Omega, (I_j)_{j\in N}, (p_{j,\omega})_{j\in N,\omega\in\Omega})$ *is said to* admit a common prior $P \in \Delta(\Omega)$ *if*

$$p_{i,\omega} = P(\cdot|I_i(\omega)) \qquad \forall i, \omega. \tag{CPA}$$

In that case, the information structure is denoted as $(\Omega, (I_j)_{j\in N}, P)$*. A* common-prior model *is any model* $M = (\Omega, (I_j)_{j\in N}, P, \boldsymbol{s})$ *in which information structure admits a common prior.*

Here, $P(\cdot|I_i(\omega))$ is a conditional probability distribution on Ω given information set $I_i(\omega)$. Since it puts probability 1 on $I_i(\omega)$ it is also viewed as a probability distribution on $I_i(\omega)$. Recall that when Ω is finite and $P(I_i(\omega)) > 0$,

$$p_{i,\omega}(\omega') = P(\omega'|I_i(\omega)) = P(\{\omega'\})/P(I_i(\omega)) \qquad (\forall \omega' \in I_i(\omega)). \tag{10.5}$$

Example 10.3 *The information structure in Figure 10.1 admits a common prior P where $P(\omega_0) = P(\omega_1) = P(\omega_2) = 1/3$. For example, $p_{1,\omega_0}(\omega_0) = P(\omega_0 | \{\omega_0, \omega_1\}) = (1/3)/(1/3 + 1/3) = 1/2$.*

One can define corretaled equilibrium directly in terms of common knowledge of rationality.

Definition 10.8 (Correlated Equilibrium) *A* correlated equilibrium *is any common-prior model $M = (\Omega, (I_j)_{j\in N}, P, \boldsymbol{s})$ in which it is common knowledge that every player is rational, i.e., $R_i = \Omega$ for each i.*

Therefore, common prior assumption is simply conjunction of common knowledge of rationality and the common prior assumption. Notice however that the common prior assumption here is made hypothetical epistemic states that are deviced to describe the players' reasoning. One may find such an assumption unwarranted.

10.1.3 Nash Equilibrium

Aumann and Brandenburger (1995) provides an epistemic foundation for Nash equilibrium. Their characterization also provides a revealing definition for Nash equilibrium in terms of conjectures—especially in two-player games. This section presents their characterization (focusing on knowledge instead of certainty for clarity).

Consider a game $G = (N, S, u)$ and an epistemic model $M = (N, \Omega, I, p, \mathbf{s})$ for game G. A *conjecture* ψ_i of player i is any probability distribution on S_{-i}, representing her belief about the other players' strategies. The conjecture of player i at any state ω is

$$p_{i,\omega} \circ \boldsymbol{s}_{-i}^{-1}.$$

For every conjecture $\psi_i \in \Delta(S_{-i})$, define the event

$$[\psi_i] = \left\{\omega | p_{i,\omega} \circ \boldsymbol{s}_{-i}^{-1}\right\}$$

that player i holds conjecture i. Observe that in two-player games a player's conjecture is a mixed strategy for the other player. The next theorem presents a characterization of Nash equilibrium in terms of conjectures for two-player games.

Theorem 10.3 *For any finite two-player game $G = (N, S, u)$ with $N = \{1, 2\}$, a (mixed) strategy profile (ψ_2, ψ_1) is a Nash equilibrium of G if and only if there exists an epistemic model M with a state ω_0 at which it is mutually known that each player i holds conjecture ψ_i and that the players are rational:*

$$\omega_0 \in K_N\left([\psi_1] \cap [\psi_2] \cap R\right).$$

Proof. First assume that $\omega_0 \in K_N\left([\psi_1] \cap [\psi_2] \cap R\right)$ for some model M and some ω_0. To prove that (ψ_2, ψ_1) is a Nash equilibrium, it suffices to show that for each i, j, and s_i,

$$\psi_j\left(s_i\right) > 0 \Longrightarrow s_i \in B_i\left(\psi_i\right);$$

recall that ψ_i and ψ_j the mixed strategy of player j and i, respectively. To establish the displayed implication, take any s_i with $\psi_j\left(s_i\right) > 0$. Then, there exists $\omega \in I_j\left(\omega_0\right)$ such that

$$s_i = \mathbf{s}_i\left(\omega\right).$$

Since $\omega_0 \in K_N\left([\psi_1] \cap [\psi_2] \cap R\right)$, $I_j\left(\omega_0\right) \subseteq [\psi_1] \cap [\psi_2] \cap R$. Hence,

$$\omega \in [\psi_1] \cap [\psi_2] \cap R.$$

Thus,

$$s_i = \mathbf{s}_i\left(\omega\right) \in B_i\left(p_{i,\omega} \circ \boldsymbol{s}_{-i}^{-1}\right) = B_i\left(\psi_i\right),$$

where the inclusion is by $\omega \in R$, and the last equality is by $\omega \in [\psi_i]$, so that $p_{i,\omega} \circ \boldsymbol{s}_{-i}^{-1} = \psi_i$. This shows that the conjectures (ψ_2, ψ_1) form a Nash equilibrium.

To prove the converse, take any mixed staretgy Nash equilibrium (ψ_2, ψ_1). Let $\psi_2 \otimes \psi_1$ be the product of the conjectures, and let

$$\mathit{Support}\left(\psi_2 \otimes \psi_1\right) = \left\{s \in S | \psi_2 \otimes \psi_1\left(s\right) > 0\right\}$$

be its support. Define an epistemic model $M = (N, \Omega, I, P, \mathbf{s})$ with common prior P by setting

$$\begin{aligned} \Omega &= \mathit{Support}\left(\psi_2 \otimes \psi_1\right) \\ I_i\left(s\right) &= \{s_i\} \times S_{-i} \\ P &= \psi_2 \otimes \psi_1 \\ \boldsymbol{s}\left(s\right) &= s. \end{aligned}$$

Observe that $[\psi_1] \cap [\psi_2] \cap R = \Omega$. ■

When there are more than two players, one needs more than mutual knowledge of conjectures and rationality. Indeed, the epistemic foundation for Nash equilibrium requires common-prior assumptiona ant the common knowledge of conjectures:

Theorem 10.4 *For any finite game $G = (N, S, u)$ with $N = \{1, \ldots, n\}$ with $n > 0$, assume that*

$$\omega_0 \in CK\left([\psi_1] \cap \cdots \cap [\psi_n]\right) \cap K_N(R)$$

for some common prior model $M = (N, \Omega, I, P, \mathbf{s})$ with $P(\omega_0) > 0$. Then, for each distinct $i, j, j' \in N$,

$$marg_{S_i}\psi_j = \; marg_{S_i}\psi_{j'} \equiv \sigma_i,$$

and the mixed strategy profile $(\sigma_1, \ldots, \sigma_n)$ is a Nash equilibrium of G.

Here the equality of marginals is direct implication of the Agreement Theorem. Since the conjectures $\mathrm{marg}_{S_i}\psi_j$ and $\mathrm{marg}_{S_i}\psi_{j'}$ of players j and j' about S_i are common knowledge and we have a common prior, then they must be the same. Note that this inequality is a necessary condition for defining a Nash equilibrium, and one needs the common knowledge of these conjectures to apply the Agreement Theorem. When there are more than two players, Nash equilibrium requires further that the players' equilibrium conjectures are in product form. Aumann and Brandenburger further shows that, under the common-prior assumption, the common knowledge of conjectures implies that the conjectures are of the product form. Then mutual knowledge of rationality implies that the conjectures put positive probability only on strategies s_i that are best responses to ψ_i.

10.2 Bayesian Games

10.2.1 Interim Correlated Rationalizability

As discussed in Section 2.1 (see Remark 1.1), the fact that two players choose their actions independently does not mean that a third player's belief about their actions will have a product form. In particular, just because all of player j's information about θ, which is the action of the nature, is summarized by t_j does not mean the belief of i about the state θ and the action of j does not have any correlation once one conditions on t_j. Once again i might find it possible that the factors that affect the payoffs may also affect how other players will behave given their beliefs (regarding the payoffs). This leads to the following notion of rationalizability, called *interim correlated rationalizability*.

Iterated Elimination of Strictly Dominated Actions Consider a Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$. For each $i \in N$ and $t_i \in T_i$, set $S_i^0[t_i] = A_i$, and define sets $S_i^k[t_i]$ for $k > 0$ iteratively, by letting $a_i \in S_i^k[t_i]$ if and only if

$$a_i \in B_i(\pi) \equiv \arg\max_{a_i'} \int u_i(\theta, a_i', a_{-i}) d\pi(\theta, t_{-i}, a_{-i})$$

for some $\pi \in \Delta(\Theta \times T_{-i} \times A_{-i})$ such that

$$\text{marg}_{\Theta \times T_{-i}} \pi = p_i(\cdot|t_i) \text{ and } \pi\left(a_{-i} \in S_{-i}^{k-1}[t_{-i}]\right) = 1.$$

That is, a_i is a best response to a belief of t_i that puts positive probability only on the actions that survive the elimination in round $k-1$. Write $S_{-i}^{k-1}[t_{-i}] = \prod_{j \neq i} S_j^{k-1}[t_j]$ and $S^k[t] = \prod_{i \in N} S_i^k[t_i]$.

Definition 10.9 *The set of all* interim correlated rationalizable *(ICR) actions for player i with type t_i is*

$$S_i^\infty[t_i] = \bigcap_{k=0}^{\infty} S_i^k[t_i].$$

Since interim correlated rationalizability allows more beliefs, interim correlated rationalizability is a weaker concept than interim independent rationalizability, i.e., if an action is interim independent rationalizable for a type, then it is also interim correlated rationalizable for that type. When all types have positive probability, ex-ante rationalizability is stronger than both of these concepts because it imposes not only independence but also the assumption that a player's conjecture about the other actions is independent of his type. Since all of the equilibrium concepts are refinements of ex-ante rationalizability, interim correlated rationalizability emerges as the weakest solution concept we have seen so far, i.e., all of them are refinements of interim correlated rationalizability.

Exercise 10.1 *Consider a Bayesian game $\mathcal{B} = (N, A, \Theta, T, u, p)$ in which each type t_i has positive ex-ante probability of $p_i(t_i)$.*

1. *Show that if a player i plays a strategy s_i with positive probability in a Bayesian Nash equilibrium, then s_i is ex-ante rationalizable.*

2. *For any ex-ante rationalizable strategy $s_i : T_i \to A_i$ and for any t_i show that $s_i(t_i)$ is interim independent rationalizable for t_i.*

Exercise 10.2 *Show that if a_i is interim independent rationalizable for some type t_i, then a_i is interim correlated rationalizable for t_i.*

Fixed-Point Definition of ICR I will next present a fixed point definition of ICR. A solution concept $\Sigma : T \rightrightarrows A$ is said to have the *best-response property* (or closed under rational behavior) if for every $t_i \in T_i$ and $a_i \in \Sigma(t_i)$, there exists $\mu^{a_i,t_i} \in \Delta(\Theta \times T_{-i} \times A_{-i})$ such that

$$a_i \in BR_i\left(\mu^{a_i,t_i}\right), \tag{10.6}$$

$$p_i(\cdot|t_i) = \text{marg}_{\Theta\times T_{-i}}\mu^{a_i,t_i}, \tag{10.7}$$

$$\mu^{a_i,t_i}(a_{-i} \in \Sigma_{-i}(t_{-i})) = 1. \tag{10.8}$$

As in the case of complete information games, the next result establishes that ICR is the largest solution concept with best-response property.

Theorem 10.5 *If $\Sigma : T \rightrightarrows A$ has best-response property, then $\Sigma \subseteq S^\infty$. Moreover, under Assumption 3.1, S^∞ has the best response property.*

I will next show that ICR characterizes the strategies that are consistent with common knowledge of rationality in any given Bayesian game.

Fix a finite Bayesian game $\mathcal{B} = (N, A, \Theta, u, T, p)$. Recall that an information (or belief) structure is a list $(\Omega, (I_i)_{i\in N}, (\pi_{i,\omega})_{i\in N,\omega\in\Omega})$ where

- Ω is a (finite) state space,

- I_i is a partition of Ω for each $i \in N$, called information partition of i,

- $\pi_{i,\omega}$ is a probability distribution on $I_i(\omega)$, which is the cell of I_i that contains ω, representing belief of i.

An *epistemic model* for $\mathcal{B}$ is a list $M = ((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ where

$$\begin{aligned} \boldsymbol{\theta} &: \Omega \to \Theta \\ \boldsymbol{t} &: \Omega \to T \\ \boldsymbol{a} &: \Omega \to A \end{aligned}$$

such that $\boldsymbol{t}$ and $\boldsymbol{a}$ are adapted and[1]

$$\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{t}_{-i})^{-1} = p_i(\cdot|\boldsymbol{t}_i(\omega)) \qquad \forall \omega, i. \tag{10.9}$$

Here, $\boldsymbol{\theta}(\omega)$ and $\boldsymbol{t}(\omega)$ determine the payoff parameter θ and the type profile t at state ω. The mapping $\boldsymbol{a}$ gives strategic meaning to the state determining what action profile is played at a given state. The condition that $\boldsymbol{t}$ and $\boldsymbol{a}$ are adapted (i.e. that $\boldsymbol{t}_i$ and $\boldsymbol{a}_i$ are constant over information sets of i for each i) ensures that each player i knows his own type $(\boldsymbol{t}_i(\omega))$ and his own action $(\boldsymbol{a}_i(\omega))$ at any state ω.

At any state in epistemic model, each player i has beliefs regarding the parameter value θ, the other players' types t_{-i} and actions a_{-i}. (In fact, he has an infinite hierarchy of beliefs regarding these variables, allowing him to reason about not only what the other players are doing but also why they are behaving that way.) In particular, at a state ω, his belief regarding the payoff parameter θ and the other players' types t_{-i} is $\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{t}_{-i})^{-1}$. His beliefs regarding (θ, t_{-i}) are however determined already as $p_i(\cdot|t_i)$ by the type space in game $\mathcal{B}$. The equation (10.9) states that the belief according to the epistemic model is identical to the belief of his type $\boldsymbol{t}_i(\omega)$ at ω according to the game. That is, the beliefs in the epistemic model are consistent with the beliefs in the underlying game.

Rationality and common knowledge of rationality are defined as before.

[1] Note that for any probability distribution P on X and any measurable function $f : X \to Y$, $P \circ f^{-1}$ is the probability distribution on Y induced by P and f. It is defined by setting

$$P \circ f^{-1}(Y') = P(\{x \in X | f(x) \in Y'\})$$

for every event $Y' \subset Y$. Thus, for finite spaces, the formula in (10.9) can be spelled out as

$$\sum_{\substack{\omega' \in I_i(\omega) \\ \boldsymbol{\theta}(\omega)=\theta, \boldsymbol{t}_{-i}(\omega)=t_{-i}}} \pi_{i,\omega}(\omega') = p_i(\theta, t_{-i}|\boldsymbol{t}_i(\omega)).$$

Definition 10.10 *A player i is said to be* rational *at ω if*

$$\boldsymbol{a}_i(\omega) \in B_i\left(\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{a}_{-i})^{-1}\right),$$

i.e., his action $\boldsymbol{a}_i(\omega)$ is a best response to his belief $\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{a}_{-i})^{-1}$ regarding (θ, a_{-i}). It is said that rationality is common knowledge in $((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ *if and only if everybody is rational throughout Ω:*

$$\boldsymbol{a}_i(\omega) \in B_i\left(\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{a}_{-i})^{-1}\right) \qquad \forall i \in N, \omega \in \Omega.$$

Here, I define common knowledge as being true throughout the state space (see Remark **??**). The next definition formalizes the strategic implications of common knowledge of rationality as follows.

Definition 10.11 *An action a_i is said to be* consistent with common knowledge of rationality for t_i *if and only if there exists a model $M = ((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ and $\omega \in \Omega$ such that $\boldsymbol{t}_i(\omega) = t_i$; $\boldsymbol{a}_i(\omega) = a_i$ and rationality is common knowledge in M.*

The next result establishes that ICR characterizes the strategic implications of common knowledge of rationality.

Theorem 10.6 *For any Bayesian game $\mathcal{B} = (N, A, \Theta, u, T, p)$ and any (i, t_i^*, a_i^*), action a_i^* is consistent with common knowledge of rationality for type t_i^* if and only if $a_i^* \in S_i^\infty[t_i^*]$.*

Proof. First assume that a_i^* is consistent with common knowledge of rationality for t_i^*, i.e., there exists a model $M = ((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ with $\omega^* \in \Omega$ such that $a_i^* = \boldsymbol{a}_i(\omega^*)$, $t_i^* = \boldsymbol{t}_i(\omega^*)$, and rationality is common knowledge in M. Define the solution concept

$$\Sigma : t \mapsto \mathbf{a}\left(\boldsymbol{t}^{-1}(t)\right).$$

Since rationality is common knowledge in M, one can easily show that Σ has best-response property. Therefore,

$$a_i^* \in \mathbf{a}_i\left(\boldsymbol{t}_i^{-1}(t_i^*)\right) \subseteq S_i^\infty[t_i^*].$$

Here, $a_i^* \in \mathbf{a}_i\left(\boldsymbol{t}_i^{-1}(t_i^*)\right)$ by definition of ω^*, and the inclusion is by Theorem 3.3.

Conversely, since S^∞ has best response property, one can define the following epistemic model with common knowledge of rationality:

$$\begin{aligned}
\Omega &= \{(\theta,t,a)\,|a\in S^\infty[t]\}\\
I_i(\theta,t,a) &= \{(\theta',t',a')\,|a'_i=a_i,t'_i=t_i\}\\
\pi_{i,(\theta,t,a)} &= \mu^{a_i,t_i},\\
\boldsymbol{\theta}(\theta,t,a) &= \theta\\
\boldsymbol{t}(\theta,t,a) &= t\\
\boldsymbol{a}(\theta,t,a) &= a.
\end{aligned}$$

Here, μ^{a_i,t_i} is as in the definition of best-response property, satisfying (10.6-10.8) for $\Sigma = S^\infty$. By (10.7) and (10.8), $((\Omega,I,\pi),\boldsymbol{\theta},\boldsymbol{t},\boldsymbol{a})$ is a well-defined epistemic model for $\mathcal{B}$. By (10.6), rationality is common knowledge in $((\Omega,I,\pi),\boldsymbol{\theta},\boldsymbol{t},\boldsymbol{a})$. Clearly, for each $a_i\in S_i^\infty[t_i]$, $\boldsymbol{a}_i(\theta,t,a)=a_i$ and $\boldsymbol{t}_i(\theta,t,a)=t_i$ for some $(\theta,t,a)\in\Omega$. ■

==============================

Aumann (1976) introduced the formulations of information structure, knowledge, and common knowledge. This is a canonical model of interactive epistemology. It is known as partition model, which implies what is known as the Truth Axiom: what you know must be true. Aumann (1987) introduced the solution concept of correlated equilibrium and showed that it captures the idea of common knowledge of rationality under common-prior assumption using this formulation. Aumann's template has been used to study the epistemic foundations of other solution concepts, most notably by Brandenburger and Dekel (1987) and Tan and Werlang (1988), who formally show that rationalizability characterizes the strategies that are consistent with common knowledge of rationality. (The arguments of Bernheim (1984) and Pearce (1984) were less formal.)

Chapter 11

Incomplete Information and Higher-Order Beliefs

In Bayesian games, players' private information is modeled by type spaces. This formulation has been quite useful, and Bayesian games play a central role in modern economics. In these models, some of the features affect the equilibrium outcomes disproportionately vis a vis their impact on players' beliefs, often for a good reason. For example, in the coordination game studied in Section 6.2, the prior mean of the fundamentals has a disproportionately large impact on equilibrium strategies because players coordinate their actions using the public information. This leads to a challenging problem in the modeling stage, when a researcher selects a model to represent an actual situation. In the actual situation, players do not get clear signals. They have some beliefs about the fundamentals and other players' beliefs, and those beliefs are vaguely articulated and incomplete at best. Then, one may not be able to identify the above features from those beliefs, reducing the scope of predictions that she can make. For example, in the coordination game above, one may not be able to rule out a wide range of prior beliefs, and may not be able to say which action the players will choose despite the fact that the game is dominance solvable.

The identification problem above leads to many related challenging questions:

- How sensitive are the solutions to misspecification of model?
- Which predictions are robust to misspecification of model?

- What do we need to know about hierarchies to verify a prediction?

This section is devoted to analyses of hierarchies of beliefs and their strategic impact towards answering those questions. I will also explore many other important questions: can we model all hierarchies by type spaces? Can we approximate them using tractable/small type spaces? Can we approximate them using common prior models? What are the implications of the common prior assumption?

The outline of the chapter is as follows. The first section is devoted to important examples. The first example provides an explicit formula for higher order expectations in the normal model for coordination game above, illustrating the challenge of making prediction under a dominance solvable model discussed above. The next example explores the role of higher-order expectations in linear Cournot oligopoly. The impact of the higher-order beliefs on equilibrium strategies is exponentially decreasing in the duopoly case and *exponentially increasing* when there are more firms. This will be closely related to the fact that Cournot duopoly is dominance solvable while Cournot oligopoly is not. These are special cases of games with a linear best response function, for which I present a general formula. The last example is the celebrated e-mail game example of Rubinstein (1989), which inspired most of the work explored in this and the previous chapter.

After formally introducing the universal type space, the next section explores the continuity properties of the solution concepts with respect to the belief hierarchies. In particular it presents a Structure Theorem for interim correlated rationalizability, which implies that one cannot obtain any sharper robust prediction by refining (interim correlated) rationalizability. (The set of robust predictions is the set of predictions from rationalizability alone no matter what refinement one uses. In particular, robust predictions of equilibrium coincide with the predictions of rationalizability.) That section also explores the important concept of common p belief and applies it to a variety of topics related to robustness, such as the strategic topology, ex-ante robustness and foundations of risk dominance.

The final section will explore the role of common-prior assumption and present some well known results, such as the Agreeing to Disagree Theorem and No Trade Theorem.

11.1 Examples

11.1.1 Higher-order Expectations with Normal Distributions

Consider the investment game

	a	b
a	θ, θ	$\theta - 1, 0$
b	$0, \theta - 1$	$0, 0$

in Section 6.2 where the return θ from investment and the idiosyncratic noise terms in players signals have standard Normal distributions:

$$\theta = y + \tau\eta \text{ and } x_i = \theta + \sigma\varepsilon_i \text{ where } (\eta, \varepsilon_1, \varepsilon_2) \overset{iid}{\sim} N(0, 1). \tag{11.1}$$

In the model, parameters y, $\sigma > 0$ and $\tau > 0$ are known, while θ is not. Assuming $\lambda = \frac{\sigma}{\tau^2}\frac{1}{\sqrt{\alpha+1}} < \sqrt{2\pi}$ where $\alpha = \tau^2/(\sigma^2 + \tau^2)$, recall that the game is dominance solvable. According to the unique solution, each player invests (i.e. plays a) if her signal is above a threshold $\hat{x}$ where

$$\Phi(\lambda(E[\theta|\hat{x}] - y)) = E[\theta|\hat{x}]. \tag{11.2}$$

In terms of interim beliefs the cutoff depends only on the prior mean y and the sensitivity parameter λ, which is a function of the variances σ^2 and τ^2. Recall also that the cutoff is a decreasing function of y. When y is very low, $E[\theta|\hat{x}]$ is nearly 1 so that the players do not invest unless investing is nearly dominant (due to their initial pessimism), and when y is very large, $E[\theta|\hat{x}]$ is nearly 1 and the players invest when there is a possibility of gain from investment.

Now imagine a researcher who wants to apply the above model to an actual situation. He does not have access to ex-ante model, which is often a hypothetical modeling device and does not know the above parameters. Instead, suppose she has access to a player's expectation about θ (namely her first-order expectation), her expectation about other players' expectations (namely her second-order expectation) and so on. He is willing to assume that the beliefs are approximately normal, and she wants to use choose parameters y, $\sigma > 0$ and $\tau > 0$ in order to model these beliefs. When can she apply this method successfully to predict which action the player will play?

Under (11.1), it is straightforward to compute players' higher-order expectations as a function of their type. For a player i with type x_i, her first-order expectation is the conditional expectation of θ given x_i:

$$E_i^1[\theta|x_i] \equiv E[\theta|x_i] = y + \alpha(x_i - y). \tag{11.3}$$

His second-order expectation the conditional expectation of the first-order expectation $E[\theta|x_j]$ of the other player given x_i:

$$\begin{aligned} E_i^2[\theta|x_i] &\equiv E\left[E_j^1[\theta|x_j]\,|x_i\right] \\ &= E[y + \alpha(x_j - y)\,|x_i], \text{ by (11.3)}, \\ &= y + \alpha(E[x_j|x_i] - y), \\ &= y + \alpha(E[\theta|x_i] - y), \text{ by } E[x_j|x_i] = E[\theta|x_i], \\ &= y + \alpha^2(x_i - y), \text{ by (11.3)}. \end{aligned}$$

Following in the same fashion, her kth-order expectation can be computed as

$$E_i^k[\theta|x_i] \equiv E\left[E_j^{k-1}[\theta|x_i]\,|x_i\right] = y + \alpha^k(x_i - y). \tag{11.4}$$

Note that the distance between the kth-order expectations and the prior mean reduced at the rate α:

$$E_i^k[\theta|x_i] - y = \alpha^k(x_i - y).$$

Thus, the higher-order expectations converge to the prior mean:

$$\lim_{k\to\infty} E_i^k[\theta|x_i] = y. \tag{11.5}$$

This is not a coincidence. For finite type spaces, Samet (1989) shows that, in a common-prior model, the higher-order expectations of any random variable converges to the ex-ante expectation of the random variable, and this is true only for common prior models.

Hence, if a researcher has access to entire hierarchy of expectations and the hierarchy is indeed induced by a model as in (11.1), she can compute the prior mean simply by taking the limit of the expectations; she can also compute α from the expectations alone. But in order to decide which action the player will take, she needs to decide whether $E_i^1[\theta|x_i]$ is above or below $E_i^1[\theta|\hat{x}]$. Although the above information is sufficient to deduce the action in some cases (e.g., $E_i^1[\theta|x_i] < 1/2$ and $y < 0$), she would need to know λ, for which she must know more about the beliefs.

Now imagine that the researcher has only a partial information about the player's belief hierarchy:

$$E_i^1[\theta|x_i] = \hat{\theta} \in (\epsilon, 1/2 - \epsilon) \text{ and } E_i^2[\theta|x_i], \ldots, E_i^k[\theta|x_i] \in \left(\hat{\theta} - \epsilon, \hat{\theta} + \epsilon\right). \tag{11.6}$$

That is, she knows the first-order expectation precisely, which happens to be some $\hat{\theta}$ at which b is risk dominant, and she also knows that the first k orders of expectations are all within ϵ neighborhood of $\hat{\theta}$. Since, one can write the kth-order expectation in terms of the first-order expectations as

$$E_i^k[\theta|x_i] - y = \alpha^{k-1}\left(E_i^1[\theta|x_i] - y\right), \tag{11.7}$$

(11.6) yields

$$\hat{\theta} - \frac{\epsilon}{1 - \alpha^{k-1}} < y < \hat{\theta} + \frac{\epsilon}{1 - \alpha^{k-1}}. \tag{11.8}$$

For small k and large α, the bounds for y are arbitrarily large. For a fixed α, as $k \to \infty$, the bounds for y approach ϵ. Hence, if the researcher knows α and k is sufficiently large, she can verify that $y < 1/2$, in which case she can conclude that $E[\theta|\hat{x}] > 1/2 > \hat{\theta}$, concluding that the player plays the risk-dominant action b.

What if the researcher does not know α? In that case, no matter how small ϵ and how large k are, the researcher cannot rule out any value of y. Indeed, fixing $\sigma = \tau^2$, so that $\lambda \in \left(1/\sqrt{2}, 1\right)$, as $\sigma \to 0$, α approaches 1, and the bound $\frac{\epsilon}{1-\alpha^{k-1}}$ gets arbitrarily large, allowing a wide range of possible y. The cutoff $E[\theta|\hat{x}]$ can take any value in $(0,1)$ for such (σ, τ, y), and thus the researcher cannot rule out either action, although she has a dominance solvable model.

11.1.2 Cournot Duopoly

Consider linear Cournot duopoly in with demand uncertainty. Two firms, 1 and 2, with zero marginal cost choose production levels, q_1 and q_2, and sell at price

$$P = \theta - (q_1 + q_2)$$

where $\theta \in \mathbb{R}$ is an unknown demand parameter. The players have private information about the demand, which is modeled by a type space. The specific features of the type space will not be important for the analysis, and we will instead work with higher-order expectations. As in the previous section, the kth-order expectation of any random variable X is

$$E_i^k[X] = \underbrace{E_i E_j E_i \cdots E_j}_{k \text{ times}}[X].$$

Assuming that the players can produce negative amounts as well as positive amounts, observe that the best response function is linear:

$$B_i(q_j) = E_i[\theta]/2 - E_i[q_j]/2.$$

Each firm produces half of the expected difference between θ and the other firms production.

Consider a Bayesian Nash equilibrium (q_1^*, q_2^*), where each q_i^* maps types to real line. We are interested in the sensitivity of equilibrium strategies to higher-order expectations. Since player i plays a best response, her equilibrium strategy is

$$q_i^* = \frac{1}{2}E_i^1[\theta] - \frac{1}{2}E_i^1\left[q_j^*\right] \tag{11.9}$$

where $E_i^1[\theta]$ is a function of her first order belief and $E_i^1\left[q_j^*\right]$ is a function of her higher-order beliefs. His first-order expectation and higher order beliefs contribute equally to her equilibrium strategy, as they each has coefficient $1/2$. Since q_j^* is a also a best response to q_i^*, by substituting $q_j^* = \frac{1}{2}E_i^1[\theta] - \frac{1}{2}E_j^1[q_i^*]$ in the previous equation, one can rewrite (11.9) as

$$q_i^* = \frac{1}{2}E_i^1[\theta] - \frac{1}{4}E_i^2[\theta] + \frac{1}{4}E_i^2[q_i^*].$$

Observe that the second-order expectation $E_i^2[\theta]$ affects the equilibrium strategy half as much as the first-order expectation $E_i^1[\theta]$ does, the remaining contribution $\frac{1}{4}E_i^2[q_i^*]$ coming from the third and higher-order beliefs. Substituting (11.9) in the previous equation, one can further clarify the role of higher-order expectations:

$$q_i^* = \frac{1}{2}E_i^1[\theta] - \frac{1}{4}E_i^2[\theta] + \frac{1}{8}E_i^3[\theta] - \frac{1}{8}E_i^3\left[q_j^*\right].$$

Following in this fashion, for any $k > 0$, one can write

$$q_i^* = \frac{1}{2}E_i^1[\theta] - \frac{1}{4}E_i^2[\theta] + \frac{1}{8}E_i^3[\theta] - \cdots - \left(-\frac{1}{2}\right)^k E_i^k[\theta] + \left(-\frac{1}{2}\right)^k E_i^k\left[q_j^*\right],$$

where $j = i$ when k is even. Note that the impact of higher-order expectations is decreasing exponentially, each contributing the equilibrium strategy half as much as the previous order does. The impact of beliefs higher than order k is $1/2^k$ in total. Assume that $\lim_k E_i^k\left[q_j^*\right]$ is bounded. For example, under common prior assumption $\lim_k E_i^k\left[q_j^*\right] = E\left[q_j^*\right]$, a real number as in the previous section with normal distributions. Then, the equilibrium strategy can be written in terms of higher-order expectations as

$$q_i^* = -\sum_{k=1}^{\infty} (-1/2)^k E_i^k [\theta].$$

That is, the game has a unique equilibrium, which is a linear function of higher-order expectations where the coefficient of the kth-order expectation is $1/2^k$.

Since the players are allowed to play any strategy and their best response functions are linear, any strategy is best response to a belief, making every strategy rationalizable (for every type). If one further assumes that the strategy space is compact (or that the firms cannot produce negative amounts), then the linear Cournot duopoly here becomes dominance solvable. As it turns out, this is closely related to the decreasing impact of higher-order beliefs, which will be clear in due course.

11.1.3 Cournot Oligopoly

Now imagine that there are $n \geq 3$ firms in the previous example. To investigate the sensitivity of equilibrium strategies to higher-order expectations in a Bayesian Nash equilibrium $(q_1^*, \ldots, q_n^*)$, write

$$E^k_{i,,j_1,\ldots,j_k}[X] = \underbrace{E_i E_{j_1} E_{j_2} \cdots E_{j_{k-1}}}_{k \text{ times}} [X]$$

for the i's expectation of the j_1's expectation of ... of the j_{k-1}'s expectation of X. Note that the best response function is now

$$B_i(q_j) = E_i[\theta]/2 - \sum_{j \neq i} E_i[q_j]/2.$$

Hence,

$$q_i^* = \frac{1}{2} E_i^1 [\theta] - \frac{1}{2} \sum_{j \neq i} E_i^1 \left[q_j^*\right]$$

Substituting the best response function iteratively, as in the case of duopoly, one can write the equilibrium strategy as

$$\begin{aligned} q_i^* &= \frac{1}{2}E_i^1[\theta] - \frac{1}{4}\sum_{j\neq i} E_{ij}^2[\theta] + \cdots - \left(-\frac{1}{2}\right)^k \sum_{j_{k-1}\neq\cdots j_1\neq i} E_{ij_1\ldots j_{k-1}}^k[\theta] \\ &\quad + \left(-\frac{1}{2}\right)^k \sum_{j_k\neq j_{k-1}\neq\cdots j_1\neq i} E_{i_0 j_1\ldots j_{k-1}}^k\left[q_{j_k}^*\right]. \end{aligned}$$

Note that there are $(n-1)^{k-1}$ sequences $i_0 i_1 \ldots i_{k-1}$ in the sum relating to the kth-order expectations. For example, for $n = 3$, the impact of the second-order expectation E_{ij}^2 is $1/4$ for each $j \neq i$. Since there are two rival firms, the total impact of the second-order expectations is $1/2$, the same as the total impact of the first-order expectation. Indeed, although the impact of each kth-order expectation $E_{i_0 j_1\ldots j_{k-1}}^k$ is $1/2^k$, since there are 2^{k-1} such sequences, the total impact of the kth-order expectations is $1/2$, for each k. That is, each order of beliefs contributes equal amount. Interestingly, when $n > 3$, the total impact of the kth-order expectations is

$$\left(\frac{n-1}{2}\right)^k,$$

exponentially increasing in k. This is closely related to the fact the Cournot duopoly is stable under best response dynamics, while the unique equilibrium in Cournot oligopoly is unstable for $n > 3$, and $n = 3$ is a knife-edge case. This is also closely related to the fact that Cournot oligopoly has a large set of rationalizable strategies even if the players cannot produce negative amounts.

11.1.4 Games with Linear Best Responses

The linear Cournot oligopoly is a special case of an important class of games in which the best response function is linear. Consider an n player Bayesian game (N, A, u, Θ, T, p) with one dimensional action spaces

$$A_1 = \cdots = A_n = \Theta = \mathbb{R}$$

and with linear best response function:

$$B_i(a_{-i}) = \alpha_i E_i[\theta] + \sum_{j\neq i} \beta_{ij} E_i[a_j].$$

As in the case of Cournot oligopoly, one can write any equilibrium strategy as a sum of higher-order expectations:

$$\begin{aligned} a_i^* &= \alpha_i E_i^1[\theta] + \sum_{j \neq i} \beta_{ij}\alpha_j E_i^2[\theta] + \cdots + \sum_{j_{k-1} \neq \cdots j_1 \neq i} \beta_{ij_1} \cdots \beta_{j_{k-2}j_{k-1}} \alpha_{j_{k-1}} E^k_{i_0 j_1 \ldots j_{k-1}}[\theta] \\ &\quad + \sum_{j_k \neq j_{k-1} \neq \cdots j_1 \neq i} \beta_{ij_1} \cdots \beta_{j_{k-2}j_{k-1}} \beta_{j_{k-1}} E^k_{ij_1 \ldots j_{k-1}}\left[a^*_{j_k}\right] \end{aligned}$$

If $\sum_{j \neq i} \left|\beta_{ij}\alpha_j/\alpha_i\right| < 1$ for all i, then the higher-order beliefs have exponentially decreasing impact. When the action spaces are bounded, the game becomes dominance solvable. With bounded A, the best response function is no longer linear, but the unique rationalizable strategy is approximately

$$a_i^* = \alpha_i E_i^1[\theta] + \sum_{k=2}^{\infty} \sum_{j_{k-1} \neq j_{k-2} \neq \cdots j_1 \neq i} \beta_{ij_1} \cdots \beta_{j_{k-2}j_{k-1}} \alpha_{j_{k-1}} E^k_{i_0 i_1 \ldots i_{k-1}}[\theta]$$

away from the boundaries. This is a useful formula that allows analyses of such games, abstracting away from the irrelevant features of the associated type space. Under the reverse inequality, i.e., $\sum_{j \neq i} \left|\beta_{ij}\alpha_j/\alpha_i\right| > 1$ for all i, these predictions are also reversed and the underlying equilibria are unstable.

Beauty Contest and Network Games A special member of the above class is the Beauty Contest game that models a general coordination game over networks where the utility function is

$$u_i = -(\theta - a_i)^2 - \sum_{j \neq i} \gamma_{ij}(a_i - a_j)^2$$

where $\left[\gamma_{ij}\right]$ is the payoff-interaction network. Here, the players face a trade off between matching the underlying state and matching the fellow players' actions where each player is weighed by γ_{ij}. This can be an actual network where a players' payoff depends only on her neighbors actions. The beauty contest game comes from the interpretation that the actions are declared estimates of a fundamental variable, such as the underlying value of stock, and the players do not want to be far off with respect to other players. Clearly, the best response function is linear:

$$B_i = \underbrace{\frac{1}{1 + \sum_{j \neq i} \gamma_{ij}}}_{\alpha_i} E_i[\theta] + \sum_{j \neq i} \underbrace{\frac{\gamma_{ij}}{1 + \sum_{j \neq i} \gamma_{ij}}}_{\beta_{ij}} E_i[a_j].$$

Once again, when $\sum_{j\neq i}\left|\beta_{ij}\alpha_j/\alpha_i\right| < 1$ for all i, the equilibrium strategies reflect the players' own information substantially and the equilibrium is stable. When $\sum_{j\neq i}\left|\beta_{ij}\alpha_j/\alpha_i\right| > 1$ for all i, the equilibrium strategies does not reflect the players' own information in any substantial way, and the equilibrium is unstable.

More generally, when Θ and action spaces are all compact metric spaces, one can define stability condition

$$d_i\left(B_i\left(\pi\right), B_i\left(\pi'\right)\right) \leq b d_{-i}\left(\pi, \pi'\right)$$

for some $b < 1$ where d_i is a metric on A_i and d_{-i} is a metric on $\Delta\left(A_{-i}\right)$ preserving the metric on A_{-i}. Then, the impact of higher-order beliefs is exponentially decreasing with rate b, and the game is dominance solvable whenever there is always a unique best response (e.g. payoff functions are strictly concave in own strategies and the action spaces are convex).

11.1.5 E-mail Game

Now, I will present a version of Rubinstein's (1989) e-mail game example that shows that mutual knowledge of a game at arbitrarily high orders may have quite different implications than common knowledge case. This example illustrates all of the structural properties of rationalizable solutions vis a vis knowledge assumptions.

Consider the following version of the investment game

	a	b
a	θ, θ	$\theta - 5, 0$
b	$0, \theta - 5$	$0, 0$

where θ can take only three values: $\theta \in \Theta = \{-2, 2, 6\}$. In this section, action a will be called Attack, while b will called No Attack; in this context, the game is often referred to as the coordinated attack problem, reflecting the story behind the type space below..

Write $\left\{t^{CK}(2)\right\}$ for the model in which it is common knowledge that $\theta = 2$. In this game there are two pure-strategy equilibria, one in which both players attack and obtain the payoff of 2, and one in which nobody attacks, each receiving zero. Pareto-dominant Nash equilibrium selects the former equilibrium.

Now imagine an incomplete information game in which the players may find it possible that $\theta = -2$. Ex ante, players assign probability 1/2 to each of the values -2

and 2. Player 1 observes the value of θ and automatically sends a message if $\theta = 2$. Each player automatically sends a "confirmation of receipt" message back whenever she receives one, and each message is lost with probability 1/2. When a message is lost, the process automatically stops, and each player is to take one of the actions a (Attack) and b (No Attack).

The information structure described in this story can be modeled by the type space $\tilde{T} = \{-1, 1, 3, 5, ...\} \times \{0, 2, 4, 6, ...\}$ with common prior p on $\Theta \times \tilde{T}$ where

$$\begin{aligned} p(\theta &= -2, t_1 = -1, t_2 = 0) = 1/2, \\ p(\theta &= 2, t_1 = 2m-1, t_2 = 2m-2) = 1/2^{2m}, \\ p(\theta &= 2, t_1 = 2m-1, t_2 = 2m) = 1/2^{2m+1} \end{aligned}$$

for each integer $m \geq 1$. Here, the type t_i is the total number of messages sent or received by player i (except for type $t_1 = -1$ who knows that $\theta = -2$).

For $k \geq 1$, type k knows that $\theta = 2$, knows that the other player knows $\theta = 2$, and so on through k orders. Hence, for high k, the beliefs of type k are similar to the common knowledge type $t^{CK}(2)$, in that both types know that other player knows that $\ldots$ $\theta = 2$ up to kth order.

Nevertheless, the rationalizable solution for type k is quite different from the type $t^{CK}(2)$. Indeed, the incomplete information game is dominance-solvable, and No Attack is the only rationalizable action for any given type. To see this, observe that type $t_1 = -1$ knows that $\theta = -2$, and hence her unique rationalizable action is No Attack. Type $t_2 = 0$ does not know θ but puts probability 2/3 on type $t_1 = -1$, thus believing that player 1 will play No Attack with at least probability 2/3, so that No Attack is the only best reply and hence the only rationalizable action. More interestingly, type $t_1 = 1$ knows that $\theta = 2$, but her unique rationalizable action is still No Attack. Although she knows that $\theta = 2$, she does not know that player 2 knows it. He assigns probability 2/3 to type 0, who does not know that $\theta = 2$, and probability 1/3 to type 2, who knows that $\theta = 2$. Since type 0 plays No Attack as her unique rationalizable action, under rationalizability, type 1 assigns at least probability 2/3 that player 2 plays No Attack. As a unique best reply, she plays No Attack. Applying this argument inductively for each type k, one concludes that the incomplete-information game is dominance-solvable, and the unique rationalizable action for all types is No Attack.

Now imagine a researcher who believes in Pareto-dominant equilibrium, in that she believes that the player ought to coordinate on an equilibrium that is better than all equilibria for all players if such an equilibrium exists. What can she conclude from her solution concept? In the complete information game, she would conclude that the players attack, while in the incomplete information game she would conclude that the players *do not* attack. What should she predict in a given situation? The answer seems to be clear. He could ask if it is common knowledge that $\theta = 2/3$. If the answer is Yes, she would predict Attack. If the answer is No, then she would figure out the beliefs of the type, and if it is as in one of the types in the incomplete information game, then she would predict No Attack. Unfortunately, however, it is an understatement that a researcher may not be able to access to the players' entire hierarchy of beliefs. After all, the players' actual beliefs may not be as clearly articulated as those of the types in type spaces. In order to explore the implications of such an incomplete knowledge of beliefs, suppose that the researcher can learn only finitely many orders of beliefs. For example, she can learn whether the player knows that $\theta = 2$, whether she knows that the other player knows that $\theta = 2$, and so on up to a finite order, but she cannot learn whether it is common knowledge that $\theta = 2$. In that case, the researcher cannot verify that the players will attack even if she observes that the players know that $\theta = 2$, that the players know that everybody knows that $\theta = 2$, and so on, no matter how many times we repeat the clause. If it is common knowledge that $\theta = 2$ as in the complete information game, then the researcher verifies that there is mutual knowledge of $\theta = 2$ at the order k that she can check, but she cannot verify whether the actual type is as in the complete information game or whether it is a type $k' > k$ in the incomplete information game, failing to verify that there will be an attack. What if the actual beliefs are as those of a type k in the incomplete information game? In that case, the researcher can find this out if she can learn the beliefs up to some order $m > k$. He could then verify that the actual type is k from $\tilde{T}$ and confirm that the player will not attack. (To see the latter confirmation, notice from the above analysis that if a type's first kth order beliefs are as in type k, then her only rationalizable action is b, as a is eliminated at round $k + 1$.)

Should the researcher then select the No Attack in the complete information game? It turns out that the same criticism applies if she selected No Attack. Indeed, if we replace $\theta = -2$ with $\theta = 6$, we obtain another model, for which Attack is the unique rationalizable action. We consider type space $\hat{T} = \{-1, 1, 3, 5, ...\} \times \{0, 2, 4, 6, ...\}$ and the common prior q on $\Theta \times \hat{T}$ where

$$\begin{aligned} q(\theta &= 6, t_1 = -1, t_2 = 0) = 1/2 \\ q(\theta &= 2, t_1 = 2m-1, t_2 = 2m-2) = 1/2^{2m} \\ q(\theta &= 2, t_1 = 2m-1, t_2 = 2m) = 1/2^{2m+1} \end{aligned}$$

for each integer $m \geq 1$. One can easily check that this game is dominance-solvable, and all types play Attack.

Note that for $k > 0$, type k knows that it is kth-order mutual knowledge that $\theta = 2$, but she does not know if the other player knows this, assigning probability $2/3$ to the type who only knows that it is $k-1$th-order mutual knowledge that $\theta = 2$. While the interim beliefs of the types with low k differ substantially from those of the common knowledge type, the beliefs of the types with sufficiently high k are indistinguishable from those of the common knowledge type according to the researcher above. But it is the behavior of those far away types that determines the behavior of the indistinguishable types; the unique behavior of $k = -1$, determines a unique behavior for $k = 0$, which in turn determines a unique behavior for $k = 1$, which in turn determines a unique behavior for $k = 2$ . . . up to arbitrarily high orders. A generalization of such a contagion argument leads to the Structure Theorem below, establishing a general structure of ICR and characterizing the robust predictions of any rationalizable solution concept to higher-order beliefs. In the meantime the next section introduces the model.

11.2 Model

Fix a finite set $N = \{1, ..., n\}$ of players and a finite set A of action profiles. Let $\Theta^* = \left([0,1]^A\right)^N$ be the space of all possible payoff functions. For any $\theta = (\theta_1, ..., \theta_n) \in \Theta^*$, the payoff of player i from any $a \in A$ is $u_i(\theta, a) = \theta_i(a)$. Note that u_i is continuous. Consider the Bayesian games with varying finite type spaces (Θ, T, π) with $\Theta \subset \Theta^*$. Recall that for each t_i in T_i, one can compute the first-order belief $h_i^1(t_i)$ about θ, the second-order belief $h_i^2(t_i)$ about $\left(\theta, h_{-i}^1\right)$, and so on, where the dependence of h_i on (Θ, T, π) is suppressed for simplicity. The type t_i and (Θ, T, π) are meant to model the infinite belief hierarchy

$$h_i(t_i) = \left(h_i^1(t_i), h_i^2(t_i), ...\right).$$

The universal type space, denoted by T^* throughout the chapter, consists of all belief hierarchies generated by all type spaces (Θ, T, π). These hierarchies incorporate a common knowledge assumption that the belief hierarchies are *coherent*: any two higher-order beliefs $h_i^k(t_i)$ and $h_i^{k'}(t_i)$ induce the same probability distribution on lower order beliefs $\left(\theta, h_{-i}^1(t_{-i}), \ldots, h_{-i}^l(t_{-i})\right)$ where $l < \min\{k, k'\}$. For example, both the first-order belief $h_i^1(t_i)$ and the marginal distribution $\mathrm{marg}_\Theta h_i^2(t_i)$ of the second-order belief on Θ are probability distributions on Θ; hence coherence requires that they must be equal: $h_i^1(t_i) = \mathrm{marg}_\Theta h_i^2(t_i)$. As it turns out, by Kolmogorov's Extension Theorem, any such hierarchy can be generated by a probability distribution $\pi_i^*(\cdot|h_i(t_i)) \in \Delta(\Theta \times T_{=i}^*)$, in that each kth-order belief $h_i^k(t_i)$ is the marginal distribution of $\pi_i^*(\cdot|h_i(t_i))$ on the space of lower-order beliefs $\left(\theta, h_{-i}^1, h_{-i}^2, ..., h_{-i}^{k-1}\right)$. Thus, (Θ^*, T^*, π^*) is a type space. Since the beliefs $\pi_i^*(\cdot|h_i(t_i))$ are identified from hierarchies $h_i(t_i)$, the beliefs $\pi_i^*(\cdot|h_i(t_i))$ will be omitted below. This shows that Harsanyi's idea of modeling incomplete information is without loss of generality, so long as one is willing to assume common knowledge of coherence.[1]

As in the examples above, assume that, in the modeling stage, the researcher can have information only on finite orders of beliefs $h_i^1(t_i), h_i^2(t_i), ..., h_i^k(t_i)$, where k can be arbitrarily high but finite and the information about these finite orders can be arbitrarily precise (without knowing $h_i^1(t_i), h_i^2(t_i), ..., h_i^k(t_i)$). If we consider the open sets generated the sets of hierarchies such a researcher can find possible, then we obtain the following (point-wise) convergence notion: For any sequence of types $t_i(m)$ and any type t_i, the sequence $t_i(m)$ converges to type t_i if kth-order beliefs under $t_i(m)$ converge to the kth-order belief under t_i for each k pointwise, i.e.,

$$t_i(m) \to t_i \iff h_i^k(t_i(m)) \to h_i^k(t_i) \qquad \forall k. \tag{11.10}$$

Here $h_i^k(t_i(m)) \to h_i^k(t_i)$ in the usual sense of convergence in distribution (i.e. for every bounded, continuous function f, $\int f dh_i^k(t_i(m)) \to \int f dh_i^k(t_i)$). The convergence here is pointwise in that we consider convergence of the kth-order beliefs for each k separately (as opposed to requiring to be uniform over all k). The topology on T^* associated with the convergence notion in (11.10) is called *the product topology*. Here, it is defined directly on the types in usual type spaces transparency.

[1]Technically, the set of available belief hierarchies also depend on the topology on Θ, $\Delta(\Theta)$, $\Delta(\Theta \times \Delta(\Theta))$, and so on.

For example, the sequence of types m from $\tilde{T}$ converges to $t^{CK}(2)$. To see this, observe that there is kth-order mutual knowledge of $\theta = 2$ according to any type $m > k$, and hence for any fixed k, the kth order beliefs become identical to those of $t^{CK}(2)$ as m goes to infinity. Similarly, the sequence of types m from $\hat{T}$ also converges to $t^{CK}(2)$.

11.3 Robustness to Incomplete Information

This section presents a general structure theorem for ICR: for any Bayesian game $\mathcal{B} = (N, A, \Theta, u, T, p)$, if $a_i \in S_i^{\infty}[t_i]$, then there exists another Bayesian game $\mathcal{B}'$ with a type t_i' such that t_i and t_i' have "similar" belief hierarchies, and yet a_i is the only ICR action for t_i', i.e., $S_i^{\infty}[t_i'] = \{a_i\}$. In particular, for each rationalizable action of a given type t, there is a sequence of types $t(m)$ that converges to t in the sense of (11.10) and the given action is the unique ICR action for types $t(m)$. For example, in the e-mail game, for the ICR action a of the complete information type $t^{CK}(2)$, the types m from $\tilde{T}$ converge to $t^{CK}(2)$ and a is the only ICR action for m. For the ICR action b, the types m from $\hat{T}$ play that role.

This implies that any refinement of rationalizability will be highly sensitive to higher-order beliefs. In particular, if a prediction is not valid under an ICR solution (but it is possibly valid for a refinement of ICR), then it will not be robust to perturbation of higher-order beliefs under any refinement of ICR.

Conversely, if a prediction is valid under all ICR solutions, then it will be robust to small perturbations of higher-order beliefs. This is because the ICR is upper-hemicontinuous with respect to belief hierarchies, so that the variations of the very high orders of beliefs do not expand the set of ICR actions. For example, in the e-mail game, the types k in type spaces $\tilde{T}$ and $\hat{T}$ converge to the common knowledge type $t^{CK}(2)$. For types k in $\tilde{T}$, the unique rationalizable action is a, and a remains to be an ICR action for the limiting type $t^{CK}(2)$. For types k in $\hat{T}$, the unique rationalizable action is b, and again b remains to be an ICR action for the limiting type $t^{CK}(2)$.

11.3.1 Upper-hemicontinuity of ICR

Theorem 11.1 *S^{∞} is upperhemicontinuous in t under the convergence notion in (11.10).*

Proof. Assuming that S^∞ is a function of hierarchies, the result follows from applying Theorem 3.4 to the Bayesian game $(N, \Theta^*, T^*, A, u, \pi^*)$, defined in the previous section, where the universal type space T^* is endowed with the product topology in (11.10). To check the conditions of Theorem 3.4, first note that, since A is finite, Θ^* is a compact metric space, as it is a subset of $\mathbb{R}^{|A \times N|}$. Second, it is well-known that in that case T^* is also a compact metric space. Finally, u_i is continuous, and $\pi_i^* (\cdot | h_i (t_i))$ is also known to be continuous in $h_i (t_i)$. Therefore, by Theorem 3.4, S^∞ is upper-hemicontinuous on $(N, \Theta^*, T^*, A, u, \pi^*)$. ■

Theorem 11.1 can be spelled out as follows. For any sequence $t_i(m)$ and any type t_i with $t_i(m) \to t_i$ as in (11.10), if $a_i \in S_i^\infty\ [t_i(m)]$ for all large m, then $a_i \in S_i^\infty[t_i]$. Note that since A is finite, a sequence $a(m)$ converges to a if and only if $a(m) = m$ for all large m. Hence, the last statement states that if $a_i(m) \to a_i$ for some $a_i(m) \in S_i^\infty[t_i(m)]$,then $a_i \in S_i^\infty[t_i]$. To appreciate the result, consider the following two implications.

Fact 11.1 *For any upperhemicontinuous solution concept F that maps each t to a subset $F[t] \subseteq A$,*

1. *F is invariant to the way hierarchies of beliefs are modeled, i.e., $F_i(t_i) = F_i$ for any two types t_i and t_i' with $h_i(t_i) = h_i(t_i')$;*
2. *F is locally constant when the solution is unique, i.e., if $F[t] = \{a\}$, then for any sequence $t(m) \to t$, $F[t(m)] = \{a\}$ for all large m.*

Exercise 11.1 *Prove these facts.*

The proof of Theorem 11.1 above assumed that S^∞ is invariant to the way hierarchies of beliefs are modeled, i.e., the set of ICR actions are identical for any two types with identical belief hierarchies. Theorem 11.1 generally holds without this assumption, and the above fact shows that Theorem 11.1 implies such an invariance as a corollary.

Many solution concepts, such as Bayesian Nash equilibrium, are upper-hemicontinuous with respect to the payoff parameters within a simple model. What is unusual about the ICR is that it is upper-hemicontinuity with respect to beliefs in the sense (11.10), which allows the types $t_i(m)$ to come from different type spaces. For example, as it has been shown in Section 3.3.4, interim independent rationalizability and Bayesian Nash equilibrium are not not invariant to the way hierarchies are modeled. Hence, these solution concepts are not upperhemicontinuous under (11.10). Indeed, the structure theorem below will imply that there is no strict refinement of S^{∞} that is upper-hemicontinuous in that sense.

The meaning of upper-hemicontinuity to economic modeling is as follows. Consider the researcher at the beginning of Section 11.2, who has noisy information about finite orders of beliefs $\left(h_i^1(t_i), h_i^2(t_i), ..., h_i^k(t_i)\right)$. Suppose that a type $\hat{t}_i$ from some type space $\hat{T}$ is consistent with her information. Upper-hemicontinuity states that if k is sufficiently high and the noise is sufficiently small, then the researcher will be sure that all of the rationalizable actions of the actual type is in $S_i^{\infty}[t_i]$. That is, the predictions of the ICR for $\hat{t}_i$ (i.e. the propositions that are true for all actions in $S_i^{\infty}[t_i]$) remain true even if there is a small misspecification of interim beliefs due to lack of information, and the researcher can validate these predictions. I will call such predictions robust to misspecification of interim beliefs. The structure theorem implies the converse of the above statement, showing that the only robust predictions are those that follow rationalizability alone.

I will now present a structure theorem that establishes that without knowledge of infinite hierarchies one cannot refine interim correlated rationalizability. In that in order to verify any predictions that relies on a refinement, the researcher has to have the knowledge of infinite hierarchy of beliefs.

11.3.2 Structure Theorem

Theorem 11.2 (Structure Theorem) *Let (N, Θ, T, A, u, p) be any Bayesian game as described in Section 11.2, $t_i \in T_i$ be any type, and $a_i \in S_i^{\infty}[t_i]$ be any ICR action for type t_i. Then, there exists a sequence of Bayesian games $(N, \Theta^m, T^m, A, u, p^m)$ with types $t_i(m) \in T_i^m$ such that*

1. $S_i^{\infty}[t_i(m)] = \{a_i\}$;

2. $h_i^k(t_i(m)) \to h_i^k(t_i)$ *for all k as $m \to \infty$, and*

3. *each (Θ^m, T^m, p^m) is a finite type space with common prior p^m.*

Moreover, every open neighborhood of t_i under (11.10) contains an open subset on which a_i is the only rationalizable action.

Broadly speaking, the structure theorem establishes that any rationalizable action a_i can be made uniquely rationalizable by perturbing the interim beliefs of the type. Since ICR is upperhemicontinuous, the previous fact implies that a_i remains the unique rationalizable action under further small perturbations. That is, a_i remains as the unique rationalizable action over an open neighborhood of the perturbed type, as in the last statement of the structure theorem. It is also well known that any such open neighborhood contains a type coming from a finite type space with a common prior, as in Part 3.

Note that the theorem does not put any restriction on the original type space (Θ, T, p); it can be infinite or may not admit any common prior model. The only restrictions are the ones imposed in the modeling section. That is, A is finite, and Θ is represented as a subspace of Θ^* (without loss of generality), rendering u continuous in θ. Despite this, it imposes strong restrictions on the perturbed types, by requiring that they come from finite types with a common prior. This leads to a stronger result. The use of ICR as the solution concept also makes the structure theorem stronger because ICR is weaker than all of the solution concepts that will be analyzed in this course, including Bayesian Nash equilibrium and interim independent rationalizability. This is because the structure theorem on ICR implies the same structure theorem on its refinements as a corollary:

Corollary 11.1 *Let Σ be a non-empty solution concept such that $\Sigma[t] \subseteq S^\infty[t]$ for every type profile t in every Bayesian game. Let also (N, Θ, T, A, u, p) be any Bayesian game as described in the model, $t_i \in T_i$ be any type, and $a_i \in S_i^\infty[t_i]$ be any ICR action for type t_i. Then, there exists a sequence of Bayesian games $(N, \Theta^m, T^m, A, u, p^m)$ with types $t_i(m) \in T_i^m$ such that*

1. $\Sigma_i[t_i(m)] = \{a_i\}$;
2. $h_i^k(t_i(m)) \to h_i^k(t_i)$ *for all k as $m \to \infty$, and*
3. *each (Θ^m, T^m, p^m) is a finite type space with common prior p^m.*

Moreover, every open neighborhood of t_i under (11.10) contains an open subset V such that $\Sigma_i[t_i'] = \{a_i\}$ for all $t_i' \in V$.

In order to spell out the implications of the structure theorem for economic modeling, consider the researcher above, who can observe arbitrarily precise noisy signal about arbitrarily high but finite orders of beliefs. There are infinitely many types from various type spaces that are consistent with information. Suppose that she chooses to model the situation by one of these types, denoted by $\hat{t}_i$. Note that the set of possible types that is consistent with her information leads to an open neighborhood of $\hat{t}_i$. Consider any a_i that is rationalizable for $\hat{t}_i$. The structure theorem states that the set of alternatives types has an open subset on which a_i is uniquely rationalizable. Hence, she cannot rule out the possibility that a_i is the unique solution in the actual situation or in the alternative models that are consistent with her information. Moreover, if a_i is uniquely rationalizable in the actual situation, she could have learned that the actual situation is in the open set on which ai is uniquely rationalizable by obtaining a more precise information about higher orders of beliefs. Therefore, she could not rule out the possibility that she could have actually verify that ai is the unique ICR action.

Now suppose that the researcher uses a particular non-empty refinement Σ of ICR as her solution concept. Since Σ has to prescribe a_i to t_i when a_i is uniquely rationalizable for t_i, and since she cannot rule out the possibility that a_i is uniquely rationalizable, she cannot rule out the possibility that her solution concept prescribes a_i as the unique solution. Hence, in order to verify a prediction of her refinement, it must be the case that her prediction holds for a_i. Since a_i is an arbitrary ICR action, this implies that the only predictions of her solution concept that she can verify are those that she could have made without refining ICR.

Exercise 11.2 *Using the structure theorem, show that S^∞ does not have an upperhemicontinuous non-empty strict refinement, i.e., if F is non-empty, upperhemicontinuous and $F\ [t] \subseteq S^\infty[t]$ for all t, then $F = S^\infty$.*

A major assumption in the formulation of the structure theorem is that the space Θ^* of possible payoff parameters is so rich that every action can be made dominant under some parameter value. This richness assumption holds in the above treatment because Θ^* is taken to be space of all payoff function. While this assumption makes sense when the underlying game is static and one is willing to relax all common knowledge assumptions on payoffs, it does not hold otherwise. For example, in a dynamic game, the richness assumption fails typically because two distinct strategies often lead to the same outcome for some strategy of another player. If one is willing to make the assumption that it is common knowledge that some payoff function is not possible, this assumption fails again. There are many recent extensions in this and other directions, as I will discuss below.

11.3.3 Common p-Belief and Strategic Topology

The premise of the previous analyses is that two situations are similar if the players' kth-order beliefs are similar up to a large but finite k, reflecting the idea that a researcher may not be able to observe the entire hierarchy of beliefs. Of course this does not mean that the game theoretical predictions will also be similar. As the structure theorem shows, without having access to infinite-order beliefs, the predictive power of game theoretical models is bounded by the predictive power of ICR, which is assumed to be limited in many situations. Alternatively, one may try to determine when two situations are "strategically similar" in the sense that the game theoretical solutions are similar (under a given solution concept). This is helpful in determining the main features of belief hierarchies that lead to a certain prediction and determine what we need to know about the belief hierarchies in order to verify a given prediction for a given situation. This alternative approach is formulated by strategic topology, which describes how similar belief hierarchies are in the strategic sense above.

This section is devoted to exploring main features that makes two hierarchies strategically similar. The main issue turns out to be how to define "almost common knowledge". The e-mail game example takes "almost common knowledge" as mutual knowledge at very high orders, while structure theorem takes it as approximate mutual knowledge at arbitrarily high but finite orders. It turns out that one needs approximate mutual knowledge at *all* orders in order for game theoretical solutions to be similar to those in the common knowledge case. This notion, due to Monderer and Samet (1989), is called *common p-belief.* As the structure theorem suggests, this notion requires knowledge of infinite hierarchies of beliefs, even though it does not require that one knows the entire hierarchy precisely in order to verify a prediction.

Formally, consider a model (Ω, I, π) for a set of players $N = \{1, \ldots, n\}$. A player i is said to *p-believe in event* E at state ω if $\pi_{i,\omega}(E) \geq p$; that is player i assigns at least probability p on E. Write

$$B_i^p(E) = \{\omega | \pi_{i,\omega}(E) \geq p\}$$

for the set of states at which player i assigns at least probability p on event E. For any $\boldsymbol{p} = (p_1, \ldots, p_n)$, write

$$B^{\boldsymbol{p}} = \cap_i B_i^{p_i}(E) = \{\omega | \pi_{i,\omega}(E) \geq p_i \ \forall i\}$$

for the set of states at which every player i assigns at least probability p_i on event E. Towards defining common $\boldsymbol{p}$-belief, for any event E define sets $C^{m,\boldsymbol{p}}(E)$ for $m \geq 0$ by

$$\begin{aligned} C^{0,\boldsymbol{p}}(E) &= E; \\ C^{m,\boldsymbol{p}}(E) &= C^{m-1,\boldsymbol{p}}(E) \cap B^{\boldsymbol{p}}(C^{m,\boldsymbol{p}}(E)) \qquad (m > 0). \end{aligned}$$

The event *common p-belief* in E is defined as

$$C^{\boldsymbol{p}}(E) = \bigcap_m C^{m,\boldsymbol{p}}(E).$$

There is common p-belief in E at state ω if $\omega \in C^{\boldsymbol{p}}(E)$. It is customary to write common p-belief and $C^p(E)$ instead of common $(p, \ldots, p)$-belief and $C^{(p,\cdots,p)}(E)$, respectively.

A special case of p-belief arises for $p = 1$. This case is referred to as *certainty* and often used as a more realistic notion of knowledge than the one defined in Section **??**. *Common certainty* in event E is defined as common p-belief in E for $p = 1$, i.e., $C^1(E)$.

Remark 11.1 *The key distinction between p-belief and knowledge is that the Truth Axiom holds for knowledge but not p-belief or certainty. One cannot know a falsehood but can be certain about it. When a player i knows an event E, event E must be true, i.e., $K_i E \subset E$. In contrast, in real life, most people are certain about many things that are patently false according to others. In particular, $B_i^p(E)$ need not be contained in E (even for $p = 1$). That is why the above definition is somewhat nuanced, repeating lower-order events when assuming p-belief in higher orders:*

$$C^{\boldsymbol{p}}(E) = E \cap B^{\boldsymbol{p}}(E) \cap B^{\boldsymbol{p}}(E \cap B^{\boldsymbol{p}}(E)) \cap B^{\boldsymbol{p}}(E \cap B^{\boldsymbol{p}}(E) \cap B^{\boldsymbol{p}}(E \cap B^{\boldsymbol{p}}(E))) \cap \cdots$$

That is, E is true; everybody $\boldsymbol{p}$-believes in E; everybody $\boldsymbol{p}$-believes in that E is true and everybody $\boldsymbol{p}$-believes in E, and so on.

As in the case of common knowledge, common $\boldsymbol{p}$-belief is closely related to evident events. An event F is said to be $\boldsymbol{p}$-*evident* if

$$F \subset B^{\boldsymbol{p}}(F).$$

That is, whenever F occurs, every player i assigns at least probability p_i to event F.

Lemma 11.1 *$C^{\boldsymbol{p}}(E)$ is $\boldsymbol{p}$-evident. In particular, $\omega \in C^{\boldsymbol{p}}(E)$ if and only if $\omega \in F \subset E$ for some $\boldsymbol{p}$-evident F.*

For a type space (Θ, T, π), one can define the above concepts by taking $\Omega = \Theta \times T$ and $I_i(\theta, t) = \{t_i\} \times \Theta \times T_{-i}$, so that each player knows her own type and nothing more. Interestingly, since each player knows her own type, for simple events of the form $E = \Theta \times E_1 \times \cdots \times E_n \subset T$, for any $\boldsymbol{p} \gg 0$, mutual $\boldsymbol{p}$-belief implies truth, i.e., $B^{\boldsymbol{p}}(E) \subset E$, simplifying the iterative definition of common $\boldsymbol{p}$-belief:

$$C^{\boldsymbol{p}}(E) = E \cap B^{\boldsymbol{p}}(E) \cap B^{\boldsymbol{p}}(B^{\boldsymbol{p}}(E)) \cap \cdots$$

as in the case of common knowledge. Moreover for any $E = \Theta' \times E_1 \times \cdots \times E_n$, we can write

$$C^{\boldsymbol{p}}(E) = \Theta' \times \bar{C}^{\boldsymbol{p}}(E)$$

for some $\bar{C}^{\boldsymbol{p}}(E) = \bar{C}_1^{\boldsymbol{p}}(E) \times \cdots \times \bar{C}_n^{\boldsymbol{p}}(E)$ where $\bar{C}_i^{\boldsymbol{p}}(E) \subset T_i$ for each i.

The main attraction of common $\boldsymbol{p}$-belief is that the strategic behavior is somewhat continuous with respect to $\boldsymbol{p}$. In particular, if it is common $\boldsymbol{p}$-belief that the payoffs are as in a complete information game G, then any equilibrium of G remains as an approximate equilibrium for large $\boldsymbol{p}$ in that the payoff gain from deviation for a player i cannot exceed

$$(1 - p_i) \max_{\theta, a, a'} |u_i(\theta, a) - u_i(\theta, a')|.$$

This is often referred to as the Basic Lemma. In contrast, in the e-mail game, the payoff gain from deviation is independent of the order in which the payoffs are mutually known, and thus the players have substantial incentive to deviate as we approach common knowledge. The following result states some useful variations of the basic lemma. Note that ε-Bayesian Nash equilibrium means that the gain from a deviation cannot be greater than ε for any type; in particular, 0-Bayesian Nash equilibrium is simply a Bayesian Nash equilibrium.

Theorem 11.3 *For any complete information game $\hat{G} = (N, A, \hat{u})$ and Bayesian game $\mathcal{B} = (N, \Theta, A, u, T, \pi)$ with $u : \Theta \times A \to [0,1]^N$, write $\hat{\Theta} = \{\theta \in \Theta | u(\theta, \cdot) = \hat{u}\}$ for the set of states at which the payoffs are as in the complete information game. Assume that $\mathcal{B}$ satisfies the conditions for existence of equilibrium in Theorem 3.1. Then, the following are true.*

1. *for every equilibrium a^* of game G, there exists a $(1-p)$-Bayesian Nash equilibrium s^* of $\mathcal{B}$ such that $s^*(t) = a^*$ for every $t \in \bar{C}^p\left(\hat{\Theta} \times T\right)$.*

2. *for every $\boldsymbol{p}$-dominant equilibrium a^* of game G, there exists a Bayesian Nash equilibrium s^* of $\mathcal{B}$ such that $s^*(t) = a^*$ for every $t \in \bar{C}^{\boldsymbol{p}}\left(\Theta \times \hat{T}\right)$ where $\hat{T}_i = \{t_i | \pi_i(u_i(\theta, \cdot) = \hat{u}_i | t_i) = 1\}$ for each i.*

Proof. (Part 1) Modify $\mathcal{B}$ by setting the action space of each type $t_i \in \bar{C}_i^p\left(\hat{\Theta} \times T\right)$ as $\{a_i^*\}$. Consider any Bayesian Nash equilibrium s^* of the modified game, where $s^*(t) = a^*$ for every $t \in \bar{C}^p\left(\hat{\Theta} \times T\right)$. It is easy to see that s^* is a $(1-p)$-Bayesian Nash equilibrium of $\mathcal{B}$. Indeed, for any type $t_i \notin \bar{C}_i^p\left(\hat{\Theta} \times T\right)$, $s_i^*(t_i)$ is a best response s_{-i}^* in game $\mathcal{B}$ because s^* is an equilibrium in the modified game, where t_i can play any action in A_i. Now consider any type $t_i \in \bar{C}_i^{\boldsymbol{p}}\left(\hat{\Theta} \times T\right)$. Type t_i assigns at least probability p_i on the event that the payoffs are as in G and the other players play according to equilibrium a_{-i}^* of G:

$$\Pr(\theta \in \hat{\Theta}, s_{-i}^*(t_{-i}) = a_{-i}^* | t_i) \geq \pi_i(\hat{\Theta} \times \bar{C}_{-i}^p(\hat{\Theta} \times T) | t_i) = \pi_i(C^p(\hat{\Theta} \times T) | t_i) \geq p$$

where the first inequality is due to the fact $s_{-i}^*(t_{-i}) = a_{-i}^*$ for any $t_{-i} \in \bar{C}_i^p\left(\hat{\Theta} \times T\right)$; the equality is by definition, and the last inequality is because $C^{\boldsymbol{p}}(E)$ is a $\boldsymbol{p}$-evident event. Since a^* is a Nash equilibrium under $\hat{\Theta}$, the gain from any deviation is non-positive when $\theta \in \hat{\Theta}$ and $s_{-i}^*(t_{-i}) = a_{-i}^*$. Since the probability of the remaining set is at most $1-p$ and the payoffs are in $[0,1]$, the expected gain from deviation can be at most $1-p$.

(Part 2) Define s^* as in the first part. Once again, only types $t_i \in \bar{C}_i^{\boldsymbol{p}}\left(\hat{\Theta} \times \hat{T}\right)$ can have an incentive to deviate. But any such type t_i assigns at least probability p_i to types $\bar{C}_{-i}^{\boldsymbol{p}}\left(\hat{\Theta} \times \hat{T}\right)$ who play a_{-i}^*. But a_i^* is a best response to such a belief because a^* is $\boldsymbol{p}$-dominant equilibrium of game G and the payoffs of type t_i are as in game G. ■

Theorem 11.3 states that equilibria under common knowledge extend to model with high common p-belief. The first part states that all equilibria remain approximate equilibria under common p-belief for large p as $(1-p)$-equilibrium. The second part states that any $\boldsymbol{p}$-dominant equilibrium remains an equilibrium whenever it is common $\boldsymbol{p}$-belief the players know that their own payoffs are as in a complete information game—even if they may face uncertainty about other players' payoffs.

Strategic Topology In order to study similarity of belief hierarchies in terms of their strategic implications more generally, one extends the above ideas beyond the complete information case. This is formally done by imposing a topology on belief hierarchies under which the solution concept remains continuous. In general, if action space A and utility function $u : \Theta \times A \to [0,1]^N$ is fixed, for discrete A, one can simply declare two hierarchies similar if the set of solutions is identical in the two situations. It is not clear what kind of general insights one can obtain from such an analysis. Alternatively, one can require two hierarchies to lead to similar solutions for all A and u, as in the case of common $\boldsymbol{p}$-belief above. Dekel, Fudenberg and Morris (2006) define a strategic topology for ICR on the belief hierarchies. Formally, this is accomplished through a convergence notion for types as (11.10), as follows.

Fix N and Θ, and varying type spaces (Θ, T, π), where types can come different type spaces. Action space A and utility functions $u : \Theta \times A \to [0,1]^N$ also vary, where A is finite. A sequence of types $t_i(m)$, coming from type spaces (Θ, T^m, π^m), is said to *converge strategically* to a type t_i, from a type space (Θ, T, π), if for every game $\mathcal{B} = (N, \Theta, A, u, T, \pi)$ and every action $a_i \in A_i$, the following conditions are equivalent:

1. $a_i \in S_i^\infty [t_i|\mathcal{B}]$ (i.e., a_i is interim correlated rationalizable for type t_i in game $\mathcal{B}$);

2. for every $\varepsilon > 0$, there exists $\bar{m}$ such that for every $m \geq \bar{m}$, $a_i \in S_i^\infty [t_i(m)|\mathcal{B}^m]$ where $\mathcal{B}^m = (N, \Theta, A, u, T^m, \pi^m)$.

The requirement that (2) implies (1) is upperhemicontinuity, which holds under the convergence notion (11.10) by Theorem 11.1. The more onerous requirement is that (1) implies (2)—lowerhemicontinuity. More importantly these conditions are required for all games. As the structure theorem suggests such a convergence requirement is difficult to satisfy, in that there are few cases that are strategically similar under all possible games. Chen et. al. (2010,2017) characterize the strategic topology in terms of more familiar notions, and their results demonstrate that this is indeed the case.

For example, they show that strategic convergence implies that all common $\boldsymbol{p}$-belief assumptions under t_i hold approximately for large m.[2] When t_i comes from a finite type space, the converse is also true. In that case, there are several other characterizations of strategic convergence. One such characterization is that convergence in (11.10) is uniform in k; that is, $\sup_k d^k \left(h_i^k \left(t_i \left(m\right)\right), h_i^k \left(t_i\right)\right)$ goes to 0 as $m \to 0$, where d^k measures the distance between kth-order beliefs. (In the same vein, Ely and Peski (2011) show that any type that makes any non-trivial common p-belief assumption is critical under the strategic topology.)

In summary, as in the complete information case, the main feature that determines the strategic implications of a belief hierarchy is the set of events that are common $\boldsymbol{p}$-belief as a function of $\boldsymbol{p}$. In order to learn these events, one needs to know *all* hierarchies of beliefs sufficiently precisely. In particular, one cannot verify a prediction by looking at the finite orders of beliefs—unless it holds under all rationalizable solutions.

11.3.4 Ex-Ante Robustness

We have so far taken an interim view: at the time of modeling, the players have their private information already. In some situations, a researcher may know how players obtained their information at an ex-ante stage. A poker game would be a good example of such a situation: first there is an ex-ante stage at which no player has information, and each player learns her own hand at the interim stage after the game cards are dealt. For such a situation, a weaker notion of robustness may be more appropriate: *ex-ante robustness*. Under this notion, which is due to Kajii and Morris (1997), it suffices to consider only ex-ante perturbations, by assigning a very high ex-ante probability to the original situation, and considering only the behavior of likely types in the perturbed model. In that case, one can obtain a robust equilibrium selection, which is not possible from an interim perspective according to the structure theorem.

Formally, fix a finite complete information game $G = (N, A, \hat{u})$. An ε-*elaboration* of

[2] Formally, for any event E, $\varepsilon > 0$ and $k \geq 1$, define a neighborhood $E^{\varepsilon,k} = \left\{(\theta, t_{-i}) \,|\, d^k \left((\theta, t_{-i}), (\theta', t'_{-i})\right) \leq \varepsilon \text{ for some } (\theta', t'_{-i}) \in E\right\}$ of event E where d^k measures the distance according to kth-order beliefs using Prohorov metric on probability distributions. Under strategic onvergence if there is common $\boldsymbol{p}$-belief in E under t_i, then for every (ε, k) there is common $(\boldsymbol{p} - (\varepsilon, \ldots, \varepsilon))$-belief in $E^{\varepsilon,k}$ under $t_i(m)$ for sufficiently large m.

G is a Bayesian game $B = (N, \Theta, A, Tu, \pi)$ with common prior such that with probability at least $1 - \varepsilon$,

1. payoffs are as in $\hat{u}$ and

2. every player i knows that her own payoff function is $\hat{u}_i$.

An equilibrium $\hat{\sigma} \in \Delta(A)$ is *(ex-ante) robust* if for every $\lambda > 0$ there exists $\varepsilon > 0$ such that every every ε-elaboration has an equilibrium distribution μ on A such that $|\mu(a) - \hat{\sigma}(a)| < \lambda$. That is, equilibrium $\hat{\sigma}$ must extend to incomplete information games in which the complete information game has a high ex-ante probability.

There are two cases in which an equilibrium extends to elaborations. First, note that 0-elaborations are correlated equilibria, and the set of distributions induced by ε-elaborations are upperhemicontinuous with respect to ε. Hence, if there is a unique correlated equilibrium in the complete information game, then all equilibrium distributions induced by ε-elaborations will be nearby the unique correlated equilibrium, rendering that equilibrium ex-ante robust. Second, if an event has high ex-ante probability, then it will be also be common $\boldsymbol{p}$-belief with high probability. Then, by the second part of Theorem 11.3, each $\boldsymbol{p}$-dominant equilibrium of complete information extends to an elaboration where the $\boldsymbol{p}$-dominant equilibrium is played by the types with common $\boldsymbol{p}$-belief of event described in the definition of elaboration. The probability of those types approach 1 as $\varepsilon \to 0$ when $p_1 + \cdots + p_n < 1$. Therefore, any $\boldsymbol{p}$-dominant equilibrium with $p_1 + \cdots + p_n < 1$ is ex-ante robust. In particular, in 2×2 games, any risk-dominant equilibrium is ex-ante robust. This is formally stated next.

Theorem 11.4 *An equilibrium $\hat{\sigma}$ is ex-ante robust if it is the unique correlated-equilibrium distribution; a $\boldsymbol{p}$-dominant equilibrium is ex-ante robust whenever $p_1 + \cdots + p_n < 1$.*

11.3.5 Common Belief Foundations of Global Games

As in the previous chapter, global games literature takes a familiar signal generation process with additive noise and often proposes an equilibrium selection for complete information games according to risk dominance. The structure theorem shows that risk-dominance selection is derived by the form of incomplete information, one could select any equilibrium by devising a suitable form of small incomplete information. One may ask: what features of higher-order beliefs lead to risk-dominance? As it is explained in the definition of strategic topology, the main strategically relevant feature of higher-order beliefs is the set of common $\boldsymbol{p}$-belief events as a function of $\boldsymbol{p}$? Following Morris, Shin, and Yildiz (2016), this section shows that indeed it is common $\boldsymbol{p}$-belief of the event that the players have approximately uniform rank beliefs that leads to risk-dominance.

Consider a Bayesian game $\mathcal{B} = (N, \Theta, A, u, T, \pi)$ with payoff matrix

	a	b
a	$x_1(t_1), x_2(t_2)$	$x_1(t_1) - 1, 0$
b	$0, x_2(t_2) - 1$	$0, 0$

as in the canonical global games example. Here, the return $x_i(t_i)$ from investment for player i is written as a function of her own type so that she knows her payoffs. In the previous chapter this was taken as θ, but note that one can incorporate that case by setting $x_i(t_i) = E[\theta|t_i]$. The only assumption on the type space is that x_i and π_i are continuous functions of t_i, and the pre-images $x_i^{-1}([\underline{x}, \bar{x}])$ of compact intervals are sequentially compact. In particular, T can be the universal type space or a high-dimensional type space.

The first observation is that there are multiple equilibria whenever there is common p-belief of payoffs that support multiple equilibria under complete information. This is established next as a corollary to Theorem 11.3 (part 2).

Corollary 11.2 *Define*

$$M^{\varepsilon} = \{t|\varepsilon < x_i(t_i) < 1 - \varepsilon \text{ for all } i\}.$$

There exist Bayesian Nash equilibria s^* and s^{**} such that

$$s^*(t) = a \text{ and } s^{**}(t) = b \qquad \forall t \in C^{1-\varepsilon}(M^{\varepsilon}).$$

Towards providing a common belief foundation for risk dominance, next define rank belief functions. Since t_i can put point masses, one needs to define two distinct notions of rank beliefs:

$$\bar{R}(t_i) = \Pr(x_j(t_i) \leq x_i(t_i)|t_i) \text{ and } \underline{R}(t_i) = \Pr(x_j(t_i) < x_i(t_i)|t_i).$$

Rank beliefs are said to be approximately uniform (ε-uniform) if they are within ε-neighborhood of 1/2:

$$URB^{\varepsilon} = \left\{t|1/2 - \varepsilon < \underline{R}(t_i) \leq \bar{R}(t_i) < 1/2 + \varepsilon \text{ for all } i\right\}.$$

Similarly, risk dominance is also defined as a function of ε: action a is ε-risk-dominant if $x_i(t_i) > 1/2 + \varepsilon$:

$$RD^{\varepsilon} = \{t|x_i(t_i) \geq 1/2 + \varepsilon \text{ for all } i\}.$$

The next result establishes the common belief foundations of risk-dominance:

Proposition 11.1 *Assume $C^{1-\varepsilon}(URB^{\varepsilon})$ is closed. Action (a, a) uniquely rationalizable whenever a is 2ε-risk dominant and there is common $1-\varepsilon$-belief in ε-uniform rank beliefs, i.e., for any*

$$(t_1, t_2) \in C^{1-\varepsilon,1-\varepsilon}(URB^{\varepsilon}) \cap RD^{2\varepsilon}.$$

The proposition establishes risk dominance as a property of belief hierarchies (namely common belief in approximately uniform rank beliefs) without explicitly referring to general structure of the type space other than the assumption that the event $C^{1-\varepsilon}(URB^{\varepsilon})$ is a closed set. Under the usual information structures in global games it is straightforward to establish this result. The main attraction of this result is that it applies to arbitrary type spaces including multi-dimensional type spaces with little structure.

Note that the proposition does not assume existence of dominance regions explicitly. When $C^{1-\varepsilon}(URB^{\varepsilon})$ is closed, there will be dominance regions and there will be sufficient contagion from those regions to the hierarchies in $C^{1-\varepsilon}(URB^{\varepsilon})$. Existence of dominance regions does not lead to uniqueness, and there can be multiplicity under common certainty of uniform rank beliefs when $C^{1-\varepsilon}(URB^{\varepsilon})$ is not closed.

11.4 Linear Algebra of Higher-Order Beliefs

This section introduces a useful formulation of higher-order beliefs, building on Samet (1998). This formulation allows one to use the tools from linear algebra and Markow chains to study higher-order expectations.

Fix a finite Bayesian game $\mathcal{B} = (N, A, \Theta, u, T, p)$ and a finite *epistemic model* $M = ((\Omega, I, \pi), \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{a})$ for $\mathcal{B}$. Recall that Ω is a finite state space; I_i is a partition of Ω for each $i \in N$, where the cell $I_i(\omega)$ is the set of states that player i cannot distinguish from ω; $\pi_{i,\omega}$ is a probability distribution on Ω that puts probability 1 on $I_i(\omega)$; $\boldsymbol{\theta}$, $\boldsymbol{t}$, and $\boldsymbol{a}$ are mappings from Ω to Θ, T, and A, respectively, so that $\boldsymbol{\theta}(\omega)$, $\boldsymbol{t}(\theta)$, and $\boldsymbol{a}(\omega)$ are the payoff parameter, type profile and action profile, respectively, at state ω. For each player i, $\pi_{i,\omega}$, $\boldsymbol{t}_i(\omega)$, and $\boldsymbol{a}_i(\omega)$ are all constant over each information set $I_i(\omega)$ of player i, as player i knows her own belief, type and action. The players' beliefs about (θ, t_{-i}) in the epistemic model are identical to the beliefs in the Bayesian game:

$$\pi_{i,\omega} \circ (\boldsymbol{\theta}, \boldsymbol{t}_{-i})^{-1} = p_i(\cdot|\boldsymbol{t}_i(\omega)) \qquad \forall \omega, i.$$

The main focus in this section is on the higher-order beliefs of types, and the actions are ignored. A typical epistemic model is given by the type space itself, so that $\Omega = \Theta \times T$ and $I_i(\theta, t) = \{t_i\} \times \Theta \times T_{-i}$. In general, an epistemic model can be richer.

Players are said to be *non-delusional* if

$$\pi_{i,\omega}(\omega) > 0 \qquad \forall i, \omega. \tag{11.11}$$

Note that this implies that $\pi_{i,\omega}(\omega') > 0$ for each $\omega' \in I_i(\omega)$.

Recall that something is said to be common knowledge if it is true throughout the model. Since Ω may contain multiple "submodels" as subspaces, it will be useful to define the information partition associated with the notion of common knowledge. In general, each "submodel" contained in Ω will be a subset of Ω that could be viewed as a model in itself under the information partition I. Formally, let I_{CK} be the finest partition of Ω that is weakly coarser than each partition I_i. Writing $I_{CK}(\omega)$ for the cell in I_{CK} that contains ω, note that

$$I_{CK}(\omega) = \bigcup_{\omega' \in I_{CK}(\omega)} I_i(\omega')$$

for each player i. That is, each common knowledge information set can be partitioned using the information sets of each player. Hence, each $I_{CK}(\omega)$ can be viewed as a state space for a model in itself. Moreover, $I_{CK}(\omega)$ is the smallest such model that contains ω. That is, for any $\omega', \omega'' \in I_{CK}(\omega)$, there exists a player i who cannot distinguish ω' and ω'', i.e., $\omega' \in I_i(\omega'')$. In a well-written applied model, we would typically have $I_{CK}(\omega) = \Omega$. An event E is said to be *common knowledge at* ω if $I_{CK}(\omega) \subseteq E$.

The linear algebra of beliefs is as follows. Any belief $p \in \Delta(\Omega)$ on Ω can be written as a row vector $p \in \mathbb{R}^{\Omega}$ where $\sum_{\omega} p(\omega) = 1$. Likewise any function $f : \Omega \to \mathbb{R}$ can be written as a column vector $f \in \mathbb{R}^{\Omega}$. The expected value of f under belief p is simply the inner product of these two vectors:

$$E_p[f] = pf.$$

For any function f and a player i, one can define the state-dependent expectation of f for player i by

$$E_i[f](\omega) \equiv E_{i,\omega}[f] \equiv \pi_{i,\omega} f,$$

defining a function $E_i[f] \in \mathbb{R}^{\Omega}$.

The players' types can be represented by a "transition matrix" on states as follows. Define a matrix M_i by

$$M_i(\omega, \omega') = \pi_{i,\omega}(\omega'),$$

where the rows and columns are indexed by the states. Here, each type corresponds a row of matrix M_i. Since each of these rows are probability distribution, the matrix M_i is row-stochastic, i.e., its rows add up to 1. Multiplying any belief p on Ω with M_i, one can obtain another belief pM_i on Ω; hence the term "transition matrix".

One can use the matrices M_i to calculate players' expectations and higher-order expectations as follows. For any function f, the expectation of f for player i is simply the product of matrix M_i and f:

$$E_i[f] = M_i f.$$

These are interim expectations: $E_i[f]$ is a function (a column vector) that gives the expectation according to type $\boldsymbol{t}_i(\omega)$ at each ω. This immediately leads to a simple formula for calculating higher-order expectation of any function. Consider any sequence $i_1, i_2, \ldots, i_k$ of players, and let $E_{i_k \ldots i_2 i_1}[f]$ be the expectation by i_k of expectation by i_{k-1} of $\ldots$ expectation by i_1 of f. This is a kth-order expectation when $i_1 \neq i_2 \neq \cdots \neq i_k$. This expectation is computed simply by

$$E_{i_k \ldots i_2 i_1}[f] = M_{i_k} \cdots M_{i_2} M_{i_1} f$$

using matrix multiplication. Since a player knows her own beliefs, her expectation of her own expectation is simply her expectation, and therefore one can ignore the pairs $i_k = i_{k+1}$ in a chain. This fact is left as an easy exercise:

Exercise 11.3 *Prove the following:*

$$\begin{aligned} \boldsymbol{t}_i(\omega_1) &= \boldsymbol{t}_i(\omega_2) \Rightarrow M_i(\omega_1,\omega) = M_i(\omega_2,\omega) \quad \forall\omega \\ M_i M_i &= M_i. \end{aligned}$$

Here, the first equality states that player i knows her beliefs, and the second equality states that her expectation of her own expectations are themselves. The next example illustrates the above operations on a simple model.

Example 11.1 *Consider a model with two players and two states: $N = \{1,2\}$ and $\Omega = \{\omega_0, \omega_1\}$. Player 1 knows the state, i.e., $I_1 = \{\{\omega_0\}, \{\omega_1\}\}$, while Player 2 cannot distinguish the states, i.e., $I_2 = \{\Omega\}$. Player 2 assigns probability $p_2(\omega)$ to state ω. The matrices M_1 and M_2 can be written as*

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} p_2(\omega_0) & p_2(\omega_1) \\ p_2(\omega_0) & p_2(\omega_1) \end{bmatrix},$$

respectively. Since Player 1 knows the true state, M_1 is an identity mapping, assigning probability 1 to the true state at each state. Player 2 does not have any private information, and hence her beliefs are independent of the state. Define function f by $f(\omega_0) = 1$ and $f(\omega_1) = 0$. Then, the first-order expectation of f for players 1 and 2 are

$$\begin{aligned} E_1^1[f] &= M_1 f = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and} \\ E_2^1[f] &= M_2 f = \begin{bmatrix} p_2(\omega_0) & p_2(\omega_1) \\ p_2(\omega_0) & p_2(\omega_1) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} p_2(\omega_0) \\ p_2(\omega_0) \end{bmatrix}, \end{aligned}$$

respectively. The higher-order expectations are trivial. Along any sequence $i_1, i_2, \ldots, i_k$ in which $i_m = 2$ for some m, we have

$$E_{i_1,i_2,\ldots,i_k}[f] = M_{i_k} \cdots M_{i_1} f = E_2^1[f];$$

the higher-order expectations reduce to the first-order expectation of Player 2, who does not have any private information.

The size of each matrix M_i is given by the cardinality of the state space, which can be quite large in practice. Fortunately, one can use smaller operators in computing higher-order expectations. Since the beliefs of a player remain constant along each of her information set, one can reduce the size of the above matrices as follows. Suppose there are two players: $N = \{1, 2\}$, and index the information sets of players 1 and 2 as $\{1, \ldots, m_1\}$ and $\{1, \ldots, m_2\}$, respectively. For any player i and any function f, one can write the first-order expectation as

$$\bar{E}_i^1 [f] = \begin{bmatrix} E_i [f|1] \\ \vdots \\ E_i [f|m_i] \end{bmatrix}$$

where $E_i [f|m]$ is the expected value of f on the information set m of player i, using belief $\pi_{i,\omega}$ for some ω in that information set. One can also rewrite the transition matrix M_i as an $m_i \times m_j$ matrix $\bar{M}_i$ by setting

$$\bar{M}_i (k, l) = \pi_{i,\omega} (l) \quad \text{for some } \omega \text{ with } I_i (\omega) = k$$

for each information set k of player i and information set l of player $j \neq i$. Then, along any sequence $i_1, i_2, \ldots, i_k$ with $i_1 \neq i_2 \neq \cdots \neq i_k$, the kth-order expectation of f can be computed by

$$\bar{E}_{i_k,\ldots,i_2,i_1} [f] = \bar{M}_{i_k} \cdots \bar{M}_{i_2} \bar{E}_{i_1}^1 [f],$$

as a function of information sets of player i_k. The following example illustrates this and some other useful techniques for computing the higher-order expectations.

Example 11.2 *Consider the following commonly used binary type space with a common prior where* $N = \{1, 2\}$ *and* $\Theta = T_1 = T_2 = \{0, 1\}$*. Ex-ante, each* θ *is equally likely:* $\Pr(\theta = 0) = \Pr(\theta = 1) = 1/2$*. Conditional on* θ*, the types* t_1 *and* t_2 *are independently distributed with* $\Pr(t_i = \theta|\theta) = q$ *for each* i *and* θ *for some* $q > 1/2$*. (Each player observes a noisy binary signal about the state with switching probability* $1 - q$*.) The information sets are identified by types. Towards computing the* k*th-order expectation of* θ*, observe that the first-order expectation of* θ *is*

$$\bar{E}^1 \equiv \begin{bmatrix} E[\theta|t_i = 1] \\ E[\theta|t_i = 0] \end{bmatrix} = \begin{bmatrix} \Pr(\theta = 1|t_i = 1) \\ \Pr(\theta = 1|t_i = 1) \end{bmatrix} = \begin{bmatrix} q \\ 1 - q \end{bmatrix},$$

where the last equality relies on symmetry and the fact that the states are equally likely. Moreover, for any t_i*,*

$$\Pr(t_j = t_i|t_i) = q^2 + (1 - q)^2 \equiv r > 1/2$$

and

$$\Pr(t_j \neq t_i | t_i) = 2q(1-q) = 1-r.$$

Hence, the transition matrix for each player i is

$$\bar{M}_i = M \equiv \begin{bmatrix} r & 1-r \\ 1-r & r \end{bmatrix}.$$

Therefore, for any $k > 1$, the kth-order expectation of θ is given by

$$\bar{E}^k = M^{k-1}\bar{E}^1.$$

The calculation of M^{k-1} is the main step in calculating the higher-order beliefs. Towards this goal, first compute the eigenvalues of the matrix M using the characteristic equation

$$|M - \lambda I| = \begin{vmatrix} r-\lambda & 1-r \\ 1-r & r-\lambda \end{vmatrix} = (r-\lambda)^2 - (1-r)^2 = 0.$$

Thus, eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2r - 1 \in (0,1)$. Check that the eigenvectors associated with λ_1 and λ_2 are

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

respectively, where $Mv_k = \lambda_k v_k$. Since v_1 and v_2 are linearly independent, one can decompose M as

$$M = Q\Lambda Q^{-1}$$

where

$$Q = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; Q^{-1} = \frac{1}{2}Q; \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Thus,

$$\begin{aligned} M^{k-1} &= Q\Lambda^{k-1}Q^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \lambda_2^{k-1} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1+\lambda_2^{k-1} & 1-\lambda_2^{k-1} \\ 1-\lambda_2^{k-1} & 1+\lambda_2^{k-1} \end{bmatrix}. \end{aligned}$$

Therefore, the kth-order expectation of θ is

$$\bar{E}^k = M^{k-1}\bar{E}^1 = \begin{bmatrix} \frac{1}{2} + \left(q - \frac{1}{2}\right)\lambda_2^{k-1} \\ \frac{1}{2} + \left(q - \frac{1}{2}\right)\lambda_2^{k-1} \end{bmatrix}.$$

The ex-ante expectation of θ is $1/2$. When a player observes a high signal $t_i = 1$, she increases her kth-order expectation by $\left(q - \frac{1}{2}\right)\lambda_2^{k-1}$, where $q - \frac{1}{2}$ measures the informativeness of the signals. When she observes a low signal, she decreases her kth-order expectation by the same amount. The effect of the signal on the kth-order expectation is exponentially decreasing in k at rate $\lambda_2 = 2r - 1$.

One can apply the theory of Markov chains to study higher-order expectations as follows. Under any prior belief p on Ω, along a sequence $i_1, i_2, \ldots, i_k$, the ex-ante value of higher-order expectations of function f can be calculated as

$$pE_{i_k \ldots i_2 i_1}[f] = pM_{i_k} \cdots M_{i_2} M_{i_1} f.$$

Thus, the ex-ante higher-order expectation operator is

$$pM_{i_k} \cdots M_{i_2} M_{i_1}.$$

Note that this is the distribution after a Markov chain, applied starting from M_{i_k}. One can then use the existing theory of Markov chains to study higher-order beliefs, especially for large values of k. For an illustration, consider the two player case, $N = \{1, 2\}$, and assume without loss of generality that $I_{CK} = \{\Omega\}$, i.e., only the model is common knowledge. Now, since $M_i M_i = M_i$, any higher-order expectation can be written as $pM_1 M_2 \cdots M_1 M_2 \cdots M_i f$ or $pM_2 \cdots M_1 M_2 \cdots M_i f$. Towards computing such a multiplication, define

$$M_{12} \equiv M_1 M_2,$$

the operator that calculates the expectation of the expectation of player 2 by player 1. Thus, $2k$th-order expectation for a function f can be computed as[3]

$$E_{p,2k} f \equiv pE_{1212\cdots 12} f = pM_{12}^k f.$$

Now, when players are non-delusional, the Markov chain defined by the transition matrix M_{12} is irreducable and aperiodic, yielding an ergodic distribution that converges to the invariance distribution of M_{12} for any initial distribution p. Therefore, as $k \to \infty$, the

[3] Clearly, the expectations $E_{2121\cdots 21}$ are computed similarly using $M_{21} = M_2 M_1$; and odd-numbered expectations $E_{2121\cdots 212}$ can be computed using $(pE_2)\, M_{12}^k$.

$2k$th-order expectations converge to the expectation under the invariance distribution of M_{12}. That is,

$$\lim_{k\to\infty} E_{p,2k} f = p_{12} f$$

where p_{12} is the invariant distribution of M_{12}, defined by

$$p_{12} = p_{12} M_{12},$$

independent of the initial prior p one uses. Therefore, one can simply compute the invariance distribution of M_{12} via the last displayed equation to study very high orders of expectations. The independence from prior p implies that the limit of the higher-order expectations is common knowledge, i.e. the limit is the same at all states. For example, in Example 11.2, the higher-order expectations of both types converge to 1/2, the ex-ante expectation of θ. Samet (1998) shows that all these are generally true.

Here p can be any belief a researcher may be using to study the higher-order expectations. For example, when there is a common prior, she may use the common prior for p; or she may use what she views as the "true distribution". Of course, the higher-order expectations are independent of what she uses in the limit $k \to \infty$ (for any given sequence $i_1, i_2 \ldots$). As it turns out, the limiting invariant distribution will be closely related to the common-prior when there is one.

A prior for player i is a probability distribution p_i on Ω under which $\pi_{i,\omega}$ is a conditional probability system, i.e., for any ω with $p(I_i(\omega)) > 0$ we have $\pi_{i,\omega}(\omega') = p(\omega')/p(I_i(\omega))$ for each $\omega' \in I_i(\omega)$.[4] One can easily show that any prior for player i is an invariant distribution for M_i:

$$p_i M_i = p_i.$$

Therefore when there is a common prior p, we have

$$pM_i = p$$

for each player i, and it remains invariant under any sequence $M_{i_k} \cdots M_{i_2} M_{i_1}$:

$$pM_{i_k} \cdots M_{i_2} M_{i_1}.$$

[4] In other words, p can be computed from marginal distribution $p(I_i(\omega))$ on information sets of player i using the formula $p(\omega') = \pi_{i,\omega}(\omega')\, p(I_i(\omega))$ for each $\omega' \in I_i(\omega)$. Thus, p is a prior for player i if and only if there exists a probability distribution q on information sets $I_i(\omega)$ such that $p(\omega') = \pi_{i,\omega}(\omega')\, q(I_i(\omega))$ for each $\omega' \in I_i(\omega)$.

In particular, in the two player case above, the common prior p is the invariant distribution for M_{12} and M_{21}, and along all sequences $i_1, i_2 \cdots$, the higher order expectations converge to the expectation under the common prior p. Samet (1998) shows that this is generally true.

11.5 Common-Prior Assumption and Higher-Order Beliefs

11.5.1 Limits of Higher-Order Expectations

As described in the previous section, along any sequence $i_1, i_2, \ldots$, the higher-order expectations of any given function converges to a commonly known limit. Moreover, when there is a common prior p, the common prior p is an invariant distribution under each player's transition matrix M_i, implying that it remains invariant under arbitrary multiplications of different players' transition matrices. When players are non-delusional, for any permutation $\sigma : N \to N$, since $M_\sigma = M_{\sigma(1)} \cdot \cdots \cdot M_{\sigma(n)}$ is irreducible and an aperiodic, p is the unique invariance distribution. This implies that when there is a common-prior, the higher-order expectations converge to the expectation under the common prior p along all sequences $i_1, i_2, \ldots$, as shown by Samet (1998). Samet (1998) further shows that the converse is also true. When the common-prior assumption is violated, there are a function f on the state space and two sequences $i_1, i_2, \ldots$ and $i'_1, i'_2, \ldots$ along which the higher-order expectations of f converge to different limits. This gives a way to check if a hierarchy of beliefs is consistent with a common prior and determine what the prior is, by computing the limits of higher-order expectations.

11.5.2 Finite-order Implications of CPA

There is a sense in which the agreeing to disagree and no-trade theorems above relies on the common knowledge assumptions, as the common-prior assumption does put any restriction on the finite-order belief hierarchies.

To see this, consider a two player model with two states $\theta \in \{\theta_1, \theta_2\}$. Each player i assigns probability $2/3$ on state θ_i, and this is common knowledge, i.e., $T = \{(t_1, t_2)\}$ where $p_i(\theta_i|t_i) = 2/3$. The common-prior assumption clearly fails in this model. Moreover, the players agree to disagree: it is common knowledge that $p_1(\theta_1) = 2/3 \neq 1/3 = p_2(\theta_1)$. Likewise, if players are risk neutral, they would like to make infinite amount of bets, failing the conclusion of the no-trade theorem.

Nonetheless, at any finite order k, the kth-order belief hierarch associated with the above model can come from a common prior model with full support. Consider $\Omega = \Theta \times \left\{1, 2, \ldots, 2^{k+1}\right\}$ and assume that there is a uniform common prior on Ω. The first 2^k information sets of player 1 are of the form $I_{1,m} = \{(\theta_1, 2m-1), (\theta_1, 2m), (\theta_2, m)\}$, and the last information set of Player 1 is $I_{1,2^k+1} = \{\theta_2\} \times \left\{2^k+1, \ldots, 2^{k+1}\right\}$. Note that at each information set $m \leq k$, Player 1 assigns probability 2/3 on θ_1:

$$\Pr(\theta_1 | I_{1,m}) = 2/3.$$

At the last information set, she assigns probability 0 on θ_1:

$$\Pr\left(\theta_1 | I_{1,2^k+1}\right) = 0.$$

Symmetrically, set the information sets of Player 2 as $I_{2,m} = \{(\theta_1, m), (\theta_2, 2m-1), (\theta_2, 2m)\}$ for $m \leq k$ and $I_{2,2^k+1} = \{\theta_1\} \times \left\{2^k+1, \ldots, 2^{k+1}\right\}$. Player 2 assigns probability

$$\Pr(\theta_2 | I_{2,m}) = 2/3$$

on θ_2 at each information set $I_{2,m}$ with $m \leq 2^k$ and $\Pr\left(\theta_2 | I_{2,2^k+1}\right) = 0$ at $I_{2,2^k+1}$.

Consider the information sets $I_{1,1}$ and $I_{2,1}$. At these information set the first k orders of beliefs are as in the non-common prior model although we have a common prior model. If one sees only the first k orders of beliefs, she might not know whether there is a common prior. In particular, seeing that there is kth-order mutual knowledge of beliefs and they are different, she may falsely conclude the beliefs are common knowledge and different, concluding by the agreeing to disagree theorem that the common-prior assumption fails. But it might be that the true model is the latter model with Ω. In that model, there is a common prior, but the beliefs are not common knowledge. In particular, player i may assign probability 0 to θ_i—instead of 2/3, and players' $(k+1)$th-order beliefs will reflect this possibility. Note that No Trade Theorem states that there will not be a trade under the latter model (despite the mutual knowlegde of disagreement). The conclusion of this theorem is sensitive to infinite orders of beliefs. This is actually useful because by No Trade theorem we can rule out the common-prior model if we see trade (although we cannot see the infinite hierarchies).

To see the above claim, let

$$E = \Theta \times \left\{1, \ldots, 2^k\right\}$$

be the event at which each player i assigns probability $2/3$ on θ_i. The set of states at which player i knows event E is

$$K_i E = \bigcup_{m=1}^{2^{k-1}} I_{i,m}.$$

Hence, the event that there is mutual knowledge of E is

$$K^1 E = K_1 E \cap K_2 E = \Theta \times \left\{1, \ldots, 2^{k-1}\right\}.$$

By repeating this argument inductively, one can see that the event that there is mth-order mutual knowledge of E is

$$K^m E \equiv K_1 K^{m-1} E \cap K_2 K^{m-1} E = \Theta \times \left\{1, \ldots, 2^{k-m}\right\}.$$

In particular, there is kth-order mutual knowledge of E at $\Theta \times \{1\}$.

Lipman (2003) shows that this is true in general (for finite models in which the players are non-delusional). For every model (possibly without a common prior) and k, there is a model with common prior and a set of types such that the first k orders of beliefs of these types are as in the first k orders of beliefs of the types in the original model. As in the example above, the ex-ante probability of the set of those types is small (as a decreasing exponential function of k). Thus, having access to only finite hierarchies of beliefs one cannot know whether the common prior assumption fails or there was a common prior but a highly unlikely event under that prior occurred.

11.6 Notes on Literature

Modeling hierarchies of beliefs through type spaces has been proposed by Harsanyi (1967). The formalization of hierarchies is due to Mertens and Zamir (1985), who considered compact set Θ of parameters, and Brandenburger and Dekel (1993), who considered the case that Θ is a complete metric space. Endowing the universal type space with the product topology above, the above papers showed general technical properties of the universal type space, including the fact that the resulting space from hierarchies is indeed a type space (i.e. for each hierarchy there is a belief on θ and the hierarchies of others that leads to the hierarchy that we started with). Some of the useful facts they showed are

- if Θ is compact (resp., complete, metrizable), so is the universal type space;
- under the usual continuity assumption of beliefs and utilities, the mapping h_i of hierarchies is continuous; the inverse is also continuous whenever h_i is one-to-one (i.e. the type spaces are continuously embedded in the universal type space, and taking limit in the original type space is equivalent to taking limit in the universal type space);
- the images of finite type spaces are dense in the universal type space, so that one can perturb a given hierarchy to find a hierarchy that comes from a type in a finite type space.

Lipman (2003) has shown further that the images of finite type spaces with common prior are also dense in the universal type space, so that every open neighborhood in the universal type space contains a hierarchy that comes from a finite type space with a common prior. These results have been used the main text of the chapter.

The formulation of interim-correlated rationalizability is due to Dekel, Fudenberg, and Morris (2007), who also proved the upper-hemicontinuity of ICR with respect to belief hierarchies (Theorem 11.1). The e-mail game is due to Rubinstein (1989). In this example Rubinstein demonstrated that efficient equilibrium of (Attack, Attack) is sensitive to the specification of higher order beliefs. This was the first application of contagion argument to the best of my knowledge.

The Structure Theorem as stated above is due to Weinstein and Yildiz (2007a); Weinstein and Yildiz (2007a) assumes that the original game has a finite type space, and Chen (2012) drops this assumption. Chen (2012), Penta (2012), and Weinstein and Yildiz (2013) extend the structure theorem to dynamic games. Penta (2012) also characterizes the robust predictions under arbitrary common knowledge restriction on who knows which parameter. Weinstein and Yildiz (2011) characterize the robust predictions of equilibrium in nice games (with convex action spaces, continuous utility functions and unique best replies) under arbitrary common knowledge restrictions on payoffs. See also Oury and Tercieux (2012) for an interesting mechanism design application with small payoff perturbations. Weinstein and Yildiz (2007b) studies impact of higher-order beliefs under a stability condition; the examples with linear best-responses at the beginning of the chapter can be found there.

11.7 Exercises

Exercise 11.4 *Consider the following complete-information game*

	x	y	z
x	$3,0$	$0,3$	$0,2$
y	$0,3$	$3,0$	$0,2$
z	$2,0$	$2,0$	$2,2$

1. *Find the sets of Nash equilibria and the rationalizable strategies*

2. *Introducing payoff uncertainty in the above game, construct a Bayesian game such that for each player i, positive integer k, and action $a \in \{x, y, z\}$, there exists a type $t_{i,k,a}$ of player i for which the above payoff matrix is kth-order mutual knowledge and the only rationalizable action is a.*

Exercise 11.5 *Begin with a complete information game*

	a	b
a	θ, θ	$\theta - c, 0$
b	$0, \theta - c$	$0, 0$

where $c > 0$ and θ is equal to some known value $\hat{\theta} \in (0, c/2)$. Consider now an e-mail game scenario: There are two possible values of θ, namely $\hat{\theta}$ and θ', with some prior probabilities p and $1 - p$. Player 1 knows the value of θ, and if $\theta = \hat{\theta}$ then the e-mail exchange takes place, where each e-mail is lost with probability $\varepsilon \in (0, 1)$. If $\theta = \theta'$ then no e-mails are exchanged. For each action $a \in \{a, b\}$, find the range of ε for which there is some e-mail game (i.e. some choice of θ' and p) in which a is the unique rationalizable action for each type. Briefly discuss your finding.

Exercise 11.6 *Let* $N = \{0, 1, \ldots, n\}^2$ *be a two dimensional grid. Say that two points* (x, y) *and* (x', y') *in* N *are* neighbors *if* $|x - x'| + |y - y'| = 1$. *At each point* $i \in N$, *there is a firm, also denoted by* i. *As in a Cournot oligopoly, simultaneously, each firm* i *chooses a quantity* $q_i \in [0, 1]$ *to produce at zero marginal cost, and sells at price*

$$P_i(\theta, q, \alpha) = \theta - q_i - \sum_{k=1}^{\infty} \alpha^{k-1} \left(\sum_{j \in N_i^k} q_j / \left| N_i^k \right| \right)^k .$$

Here, $\theta \in [1, 2]$ *is a common demand parameter, and* $\alpha \in [0, 1)$ *is an interaction parameter with respect to distant neighbors.* N_i^k *is the* k*th iterated set of neighbors of* i*: thus* N_i^1 *is the immediate neighbors of* i *(e.g.,* $N_{(0,0)}^1 = \{(1, 0), (0, 1)\}$*),* N_i^2 *is the neighbors of neighbors of* i *(e.g.,* $N_{(0,0)}^2 = \{(0, 0), (0, 2), (2, 0), (1, 1)\}$*), and so on. The payoff of firm* i *is its profit:* $q_i P_i$. *The value of* α *is common knowledge, but* θ *is unknown, drawn from some finite set* $\Theta \subseteq [1, 2]$. *The players' information about* θ *is represented by a finite type space* T, *with some joint prior* $p \in \Delta(\Theta \times T)$.

1. *For any choice of a Bayesian Nash equilibrium* $q_\alpha^* : T \to [0, 1]^N$ *of the above Bayesian game (for each* α*), and for any* $t_i \in T_i$*, find* $\lim_{\alpha \to 0} q_\alpha^*(t_i)$.

   [It suffices to find a formula that consists of iterated expectations of the form $E_{ij_1 \ldots j_k}[\theta | t_i] \equiv E[E[\cdots E[\theta | t_{j_k}] \cdots | t_{j_1}] | t_i]$*, where* $i, j_1, \ldots, j_k \in N$. *Your formula does not need to be in closed form, but it should not refer to* q^**.]*

2. *Simplify your result in part (a) under the assumption that* $E_{ij}[\theta | t_i] = E[\theta | t_i]$ *for all* i, j, *and* t_i.

Exercise 11.7 *Fix* (N, A, u, Θ, T) *where* N *and* A *are finite,* $\Theta = T_i = R$ *for each* $i \in N$, *and* $u : \Theta \times A \to R^N$ *is continuous. Fix also a real number* y. *Consider a family of Bayesian games, indexed by real numbers* $\sigma > 0$, *and* $\tau > 0$, *where*

$$\begin{aligned} t_i &= \theta + \sigma \varepsilon_i \qquad (\forall i \in N) \\ \theta &= y + \tau \eta \end{aligned}$$

for some independently distributed random variables $\eta, \varepsilon_1, \ldots, \varepsilon_n$. *Assume that each* ε_i *has a density* f *and* η *has density* g *where* f *and* g *are positive, continuous, even (i.e.* $f(-\varepsilon) = f(\varepsilon) > 0$*) and decreasing on* R_+. *(You can make any other technical assumption that you may deem necessary.)*

1. *For any* $\sigma > 0$, $p < 1$ *and* $\delta > 0$, *show that there exist* $\bar{\tau} > 0$ *and* $\bar{z} > 0$ *such that event* $[y - \delta, y + \delta] \times [y - \bar{z}, y + \bar{z}]^N$ *is p-evident.*

2. *For any* $\sigma > 0$, *show that there exist* $\bar{\tau} > 0$ *and* $\bar{z} > 0$ *such that for any* $\tau < \bar{\tau}$ *and any strict Nash equilibrium* a^* *of complete information game* $(N, A, u(y, \cdot))$, *there exists a Bayesian Nash equilibrium* s^* *with* $s^*(t) = a^*$ *for all* $t \in [y - \bar{z}, y + \bar{z}]^N$.

3. *For any* $p > 1/2$, $\bar{z} > 0$, *and* $\tau > 0$, *show that there exist* $\bar{\sigma} > 0$ *such that the event* $\Theta \times [y - \bar{z}, y + \bar{z}]^N$ *is not p-evident whenever* $\sigma < \bar{\sigma}$.

4. *Briefly discuss your findings in relation to global games literature.*

Exercise 11.8 *Consider a two-player incomplete information game with types*

$$x_i = y + \eta + \varepsilon_i$$

where y *is a known parameter and* η, ε_1, *and* ε_2 *are independently-distributed real-valued random variables. Each noise term* ε_i *has a symmetric distribution* F *on* $[-1, 1]$ *with density* f, *and* η *has a symmetric distribution* G *with density* g, *where* $f(-\varepsilon) = f(\varepsilon)$ *and* $g(-\eta) = g(\eta)$. *There are two possible cases for the tail densities of the common shock:*

Fat Tails *There exist* $\bar{\eta}$, $\alpha > 1$, *and* γ *such that*

$$g(\eta) = \gamma |\eta|^{-\alpha} \qquad (\forall \eta > \bar{\eta}).$$

Light Tails *There exist* $\bar{\eta}$, $\alpha > 0$, *and such that*

$$g(\eta) = \gamma \exp\left(-\alpha \eta^2\right) \qquad (\forall \eta > \bar{\eta}).$$

In the following exercises you are asked to compute the role of the prior belief y *on the rationalizable actions of some type* x_i. *(Make any technical assumptions that you may need to make.)*

1. *Assume* F *is the uniform distribution on* $[-1, 1]$. *For* $y \ll x_i - \bar{\eta}$, *compute the conditional distributions of* η *and* x_{-i} *and the rank belief* $R(x_i)$ *for each case above. What happens as* $y \to -\infty$?

2. *For more general distributions F on $[-1, 1]$, compute the limit of interim beliefs of each x_i as $y \to \infty$ or $y \to -\infty$ for each case above. Compute the limit of the belief hierarchy of each type x_i for the above limits (it suffices to describe a type from a type space whose belief hierarchy is as in the limit).*

3. *Suppose that, in a game, x_i has a unique rationalizable action, a_i^*, under an improper uniform prior for η. Compute the rationalizable action of type x_i in the limits $y \to \infty$ or $y \to -\infty$ for the fat tails case.*

4. *Consider a monotone supermodular game where the lowest action 0 is dominant for very low x_i and the highest action 1 is dominant for very high x_i. Compute the rationalizable action of a given type x_i in the limits $y \to \infty$ and $y \to -\infty$ for the light tails case. (The limit may depend on the details of the game, which are intentionally left blank for you to fill in—towards making the problem tractable.)*

Exercise 11.9 *For each payoff profile below, compute the rationalizable strategies when the payoff profile is common knowledge. Then, construct a type space with some payoff relevant parameter θ, such that (i) the payoffs are continuous functions of θ, (ii) the payoffs are as in the given profile at some θ^*, and (iii) for each rationalizable action a_i under common knowledge there is a sequence of types whose belief hierarchies converge to the hierarchy with common certainty of θ^* but a_i is the unique rationalizable action for the types in the sequence.*

Matching Penny:

	a	b
a	$1, -1$	$-1, 1$
b	$-1, 1$	$1, -1$

Dominant Strategy Equilibrium:

	a	b
a	$1, 1$	$0, 0$
b	$0, 0$	$0, 0$

Exercise 11.10 *Consider the following variation of Example 6.2, where $N = \{1, 2\}$, $\Theta = T_1 = T_2 = \{0, 1\}$. It is common knowledge that Player 1 knows the true state*

(i.e., $\Pr(t_1 = \theta|\theta) = 1$ *for each* θ*), while player two is imperfectly informed as before:* $\Pr(t_2 = \theta|\theta) = q$ *for each* θ *for some* $q \in (1/2, 1)$*. The players have non-common prior on* θ*: Player 1 assigns probability* $p \neq 1/2$ *on* $\theta = 0$ *while Player 2 assign probability 1/2 on each state as before. For any sequence* $i_k \neq i_{k-1} \neq \cdots \neq i_1$*, compute the* k*th-order expectation*

$$E^k_{i_k \cdots i_1} f,$$

where $f(\theta) = 1 - \theta$*. Compute the limit as* $k \to \infty$*, and briefly discuss your findings.*

Challenge: *What if player 1 were also imperfectly informed but her signal were more informative than player 2? Can you come up with a formalization of the idea that less informed players have stronger impacts on higher-order expectations?*

Exercise 11.11 *Consider the following two-player information structure—with differing priors. There is a real-valued parameter* θ*. Ex-ante, each player* i *believes that* θ *is normally distributed with mean* μ_i *and variance* σ^2*. Each player* i *privately observes a noisy signal*

$$x_i = \theta + \varepsilon_i,$$

where $(\theta, \varepsilon_1, \varepsilon_2)$ *are independently distributed and* $\varepsilon_i \sim N(0, 1)$ *according to both players.*

1. Assuming that μ_1 and μ_2 are common knowledge, compute the kth-order expectation $E^k_i(x_i)$ of each type, and difference of expectations $E^k_1(x_1) - E^k_2(x_2)$ for each type profile (x_1, x_2). Compute the limit of $|E^k_1(x_1) - E^k_2(x_2)|$ as $k \to \infty$.

2. Now assume that each player i privately knows μ_i and μ_1 and μ_2 are independently distributed with expected value $\bar{\mu}$; that is, it is common knowledge that each player i, independent of (x_i, μ_i), believes that μ_j and θ are independently distributed where the distribution of μ_j does not depend on (x_i, μ_i) and has mean $\bar{\mu}$. Compute the kth-order expectation $E^k_i(x_i, \mu_i)$ of each type, and difference of expectations $E^k_1(x_1, \mu_1) - E^k_2(x_2, \mu_2)$ for each type profile (x_1, μ_1, x_2, μ_2). Compute the limit of $|E^k_1(x_1, \mu_1) - E^k_2(x_2, \mu_2)|$ as $k \to \infty$.

3. Briefly discuss your finding, focusing on how differing priors affect higher-order expectation differences.

Bibliography

[1] Aumann, Robert (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1-18.

[2] Bernheim, D. (1984): "Rationalizable Strategic Behavior," Econometrica, 52, 1007-1028.

[3] Brandenburger, A. and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1391-1402.

[4] Brandenburger, A. and E. Dekel (1993): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 59, 189-198.

[5] Chen, Y. (2012): "A Structure Theorem for Rationalizability in Dynamic Games", *Games and Economic Behavior*, forthcoming.

[6] Chen, Y. C., Di Tillio, A., Faingold, E., & Xiong, S. (2010). "Uniform topologies on types," *Theoretical Economics*, 5(3), 445-478.

[7] Chen, Y. C., Di Tillio, A., Faingold, E., & Xiong, S. (2017). "The strategic impact of higher-order beliefs," *Review of Economic Studies*.

[8] Dekel, E. D. Fudenberg, S. Morris (2006): "Topologies on Types," *Theoretical Economics*, 1, 275-309.

[9] Dekel, E. D. Fudenberg, S. Morris (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15-40.

[10] Ely, J. and M. Peski (2006): "Hierarchies of belief and interim rationalizability", *Theoretical Economics* 1, 19–65.

[11] Ely J. and M. Peski (2011): "Critical Types". *Review of Economic Studies*, 78, 907-937.

[12] Harsanyi, J. (1967): “Games with Incomplete Information played by Bayesian Players. Part I: the Basic Model,” *Management Science* 14, 159-182.

[13] Kajii, A. and S. Morris (1997): “The Robustness of Equilibria to Incomplete Information,” *Econometrica*, 65, 1283-1309.

[14] Mertens, J. and S. Zamir (1985): “Formulation of Bayesian Analysis for Games with Incomplete Information,” *International Journal of Game Theory*, 10, 619-632.

[15] Monderer, D. and D. Samet (1989): “Approximating Common Knowledge with Common Beliefs,” *Games and Economic Behavior*, 1, 170-190.

[16] Morris, Stephen, Hyun Song Shin, and Muhamet Yildiz (2016): "Common belief foundations of global games." *Journal of Economic Theory* 163, 826-848.

[17] Oury, M. and O. Tercieux (2012): “Continuous Implementation,” *Econometrica*, forthcoming.

[18] Pearce, D. (1984): “Rationalizable Strategic Behavior and the Problem of Perfection,” *Econometrica*, 52, 1029-1050.

[19] Penta, A. (2012): “Higher Order Uncertainty and Information: Static and Dynamic Games,” *Econometrica*, 80, 631-661.

[20] Rubinstein, A. (1989): “The Electronic Mail Game: Strategic Behavior Under ‘Almost Common Knowledge’,” *The American Economic Review*, 79, No. 3, 385-391.

[21] Tan, T. and S. Werlang (1988): “The Bayesian foundations of solution concepts of games,” *Journal of Economic Theory* 45, 370-391.

[22] Weinstein, J. and M. Yildiz (2007a): “A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements,” *Econometrica*, 75, 365-400.

[23] Weinstein, Jonathan, and Muhamet Yildiz (2007b): "Impact of higher-order uncertainty." *Games and Economic Behavior* 60, 200-212.

[24] Weinstein, J. and M. Yildiz (2011): "Sensitivity of Equilibrium Behavior to Higher-order Beliefs in Nice Games," *Games and Economic Behavior*, 72, 288-300.

[25] Weinstein, J. and M. Yildiz (2009): "A Structure Theorem for Rationalizability in Infinite-horizon Games," Working Paper.

[26] Yildiz, Muhamet (2015): "Invariance to representation of information." *Games and Economic Behavior* 94 142-156.

Appendix A

Mathematical Tools

In this appendix, I will present the three fundamental mathematical theorems that are commonly used in the proofs: the Separating-Hyperplane Theorem, Berge's Maximum Theorem, and Kakutani's Fixed-point Theorem. There are many extensions of these results, but I will focus on the versions that will be used in the proofs.

The Separating-Hyperplane Theorem is used throughout economic theory, often in order to establish existence of prices or beliefs. The Maximum Theorem is used in order to establish the continuity properties of the solutions. Finally, the Fixed-Point Theorems are often used in order to establish the existence of a solution.

A.1 Separating Hyperplane Theorem

Separating Hyperplane Theorems simply state that two convex sets can be separated by a plane. Using this simple fact creatively, one can prove many interesting results in economics. There are different versions of this theorem, varying with respect to their generality, regularity assumptions, and the notion of separation. I will first present a very general version, which considers arbitrary topological vector spaces.

Theorem A.1 (Hahn-Banach Separating Hyperplane Theorem) *Let V be a topological vector space, and A and B be convex, non-empty subsets of V. Assume that $A \cap B = \emptyset$ and A is open. Then, there exist a continuous linear map $\lambda : V \to \mathbb{R}$ and $t \in \mathbb{R}$ such that $\lambda(a) < t \leq \lambda(b)$ for all $a \in A$, $b \in B$.*

Note that, since λ is linear, the isocurves $\{v|\lambda(v) = x\}$ with varying x are hyperplanes in V. For arbitrary disjoint, non-empty sets A and B, the theorem states that there exists a hyperplane $\{v|\lambda(v) = x\}$ such that A lies on one side and B lies on the other side. When A is open, A will not intersect the hyperplane. From this, one can easily derive different versions of this result with varying the topological assumptions on A and B. For example, if A has non-empty interior, one can conclude that there exists a continuous linear map $\lambda : V \to \mathbb{R}$ and $t \in \mathbb{R}$ such that $\lambda(a) \leq t \leq \lambda(b)$ for all $a \in A$, $b \in B$, and moreover $\lambda(a) < t$ whenever a is in the interior. (Here, one uses the interior of A, which is an open set, instead of A in the theorem.) Similarly, if B is closed (but A is not necessarily open), then one can conclude, once again, that a continuous linear map $\lambda : V \to \mathbb{R}$ and $t \in \mathbb{R}$ such that $\lambda(a) < t \leq \lambda(b)$ for all $a \in A$, $b \in B$, by substituting the complement of B in place of A in the theorem.

The next result states a special case of this theorem for finite-dimensional Euclidean spaces, where any linear mapping λ is simply an inner-product with a vector.

Theorem A.2 (Separating Hyperplane Theorem in $\mathbb{R}^n$) *Let A and B be convex, non-empty, and disjoint subsets of $\mathbb{R}^n$. Assume that either A is open, or B is closed. Then, there exist a non-zero $\lambda \in \mathbb{R}^n$ and $t \in \mathbb{R}$ such that $\lambda \cdot a < t \leq \lambda \cdot b$ for all $a \in A$, $b \in B$.*

Note that when $\lambda = (\lambda_1, \ldots, \lambda_n) > 0$ (i.e. $\lambda_x \geq 0$ for each $x \in X = \{1, \ldots, n\}$ with at least one strict inequality), one can obtain a probability measure $\mu = (\mu_1, \ldots, \mu_n)$ on the set X by normalization

$$\mu_x = \frac{\lambda_x}{\lambda_1 + \cdots + \lambda_1}.$$

In that case, the linear mapping $v \mapsto \mu \cdot v$ is the expected value of the function v with respect to μ. More generally, let V be a set of continuous functions $v : X \to \mathbb{R}$ for some possibly infinite but compact metric space X. If λ in Hahn-Banch Theorem above is increasing, then one can define a probability distribution P_λ on X by setting

$$P_\lambda(A) = \frac{\lambda(1_A)}{\lambda(1_X)}$$

on each event A, where 1_A is the characteristic function of set A, taking value of 1 on A and 0 outside of A. In that case, the mapping λ is the expectation operator on functions with respect to P_λ:

$$\lambda(v) = \int v dP_\lambda.$$

A.2 Continuity of Correspondences

Consider any metric spaces (X, d) and (Y, d) where d denotes the metric. A correspondence F from X to Y is an arbitrary subset of $X \times Y$. Conventionally, the correspondence F is denoted as $F : X \rightrightarrows Y$, writing $F(X)$ for the set $\{y : (x, y) \in F\}$. The defining subset of $X \times Y$ is called the *graph* of the correspondence and denoted by $G(F) = \{(x, y) : y \in F(x)\}$ for clarity. A correspondence F is said to be *non-empty* if $F(x)$ is always non-empty, *closed-valued* if $F(x)$ is always closed, *convex-valued* if $F(x)$ is always convex, and so on.

For correspondences, continuity concepts relate to the metric used to measure the distances between the sets. Here, I will introduce a standard notion of continuity. A correspondence F is said to be *upperhemicontinuous* at $x \in X$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every x' with $d(x, x') < \delta$ and for every $y' \in F(x')$, there exists some $y \in F(x)$ with $d(y, y') < \varepsilon$.

The set Y is often the domain of optimization, such as the strategy set, and is assumed to be compact. In that case, upperhemicontinuity is closely related to the following simpler concept. A correspondence F is said to have the *closed-graph property* if $G(f)$ is closed, i.e., for any sequence (x_m, y_m) with limit point (x, y), if $y_m \in F(x_m)$ for each m, then $y \in F(x)$. In general, the closed-graph property is weaker than upper-hemicontinuity, but the two concepts coincide when F is closed-valued and Y is bounded.

Proposition A.1 *If Y is compact and a correspondence $F : X \rightrightarrows Y$ has the closed-graph property, then F is closed-valued and upperhemicontinuous.*

Since the correspondences considered here are closed-valued and the ranges of the correspondences are compact, I will often use the closed-graph property and upperhemicontinuity interchangeably.

Exercise A.1 *Find a correspondence that has a closed graph but fails upperhemicontinuity. Find also an upperhemicontinuous correspondence that does not have a closed graph.*

A correspondence $F : X \rightrightarrows Y$ is said to be *lowerhemicontinuous* if for any sequence $x_m \to x$ and for any $y \in F(x)$, there exists a sequence $y_m \to y$ with $y_m \in F(x_m)$ for each m. A correspondence is said to be *continuous* if it is both upperhemicontinuous and lowerhemicontinuous.

The next section shows that the set of optimal choices is upperhemicontinuous but not lowerhemicontinuous. The solution concepts often inherit this property.

A.3 Berge's Maximum Theorem

The Maximum Theorem states that the best-response correspondence is upperhemicontinuous in parameters when the payoff function and the domain of optimization vary continuously in all relevant parameters. The next result presents a version of Berge's maximum theorem for metric spaces. There are also more general versions for general topological spaces.

Theorem A.3 (Berge's Maximum Theorem) *Let X and Y be metric spaces and $f : X \times Y \to \mathbb{R}$ be a continuous function, and let $D : X \to Y$ be a continuous correspondence that is nonempty- and compact-valued. Then:*

1. *The function $M : X \to \mathbb{R}$, defined by*

$$M(x) = \max_{y \in D(x)} f(x, y),$$

is continuous and

2. *the correspondence $F : X \rightrightarrows Y$,*

$$F(x) = \arg\max_{y \in D(x)} f(x, y),$$

is nonempty, compact-valued, and upperhemicontinuous.

That is, if the objective function f is continuous both with respect to the object y of maximization and the parameter x and if the domain $D(x)$ is continuously varying with respect to the parameter x in the sense of correspondences above, then the maximum payoff $M(x)$ that can be obtained is continuous in parameter x, and the set $F(x)$ of maximizers is upperhemicontinuous with respect to x. In that case, in optimization problems, the limits of the solutions is a solution to the optimization problem in the limit. Hence, one can find *a* solution by considering approximate problems and taking the limit. Nevertheless, $F(x)$ often fails to be lowerhemicontinuous, and hence there can be other solutions in the limit problems, solutions that cannot be approximated as the solutions to approximate problems.

Example A.1 *Take $x \in [0,1]$, $Y = \{L, R\}$ and*

$$f(x, L) = x \text{ and } f(x, R) = 0.$$

Then, the set of optimizers is

$$F(x) = \begin{cases} \{L, R\} & \text{if } x = 0, \\ \{L\} & \text{if } x > 0. \end{cases}$$

In the limit $x \to 0$, the unique solution L remains to be a solution (i.e. F is upperhemicontinuous), but a new solution R appears at the limit (i.e. F is not lowerhemicontinuous).

In general, the most difficult condition of the Maximum Theorem is continuity requirement for the domain $D(x)$ of optimization, as it involves lowerhemicontinuity. Fortunately, the domain of maximization is often fixed (e.g. the entire strategy set). In that case, the result simplifies to:

Theorem A.4 (Berge's Maximum Theorem) *Let X and Y be metric spaces where Y is compact and $f : X \times Y \to \mathbb{R}$ be a continuous function. Then:*

1. *The function $M : X \to \mathbb{R}$, defined by*

$$M(x) = \max_{y \in Y} f(x, y),$$

is continuous and

2. *the correspondence $F : X \rightrightarrows Y$,*

$$F(x) = \arg\max_{y \in Y} f(x, y),$$

is nonempty, compact-valued, and upperhemicontinuous.

An important application of the Maximum Theorem is continuity of best responses with respect to beliefs:

Lemma A.1 *Let* (S, d) *and* (X, d) *be separable metric spaces, and* $u : S \times X \to \mathbb{R}$ *be continuous. Then,*

$$U(s, \mu) = \int u(s, x)\, d\mu(x)$$

is continuous in $(s, \mu) \in S \times \Delta(X)$ *and*

$$B_i(\mu) = \arg\max_{s \in S} \int u(s, x)\, d\mu(x)$$

is upperhemicontinuous in μ*.*

Proof. By the Maximum Theorem, it suffice to show that $U(s, \mu)$ is continuous. This is implied by the continuity of the mapping $(s, \mu) \mapsto \delta_s \times \mu$, where δ_s is the Dirac measure on s. I will give a more explicit elementary proof for compact metric spaces. To this end, take any sequence (s^n, μ^n) with limit (s, μ) and any $\varepsilon > 0$. First note that since u is continuous and $S \times X$ is compact, by Heine-Cantor Theorem, u is uniformly continuous. Hence, there exists $\delta > 0$ such that, whenever $d(s, s') < \delta$,

$$|u(s, x) - u(s', x)| < \varepsilon/2,$$

yielding

$$|U(s, \mu^n) - U(s', \mu^n)| \leq \int |u(s, x) - u(s', x)|\, d\mu^n(x) \leq \varepsilon/2$$

for every n. Since $s^n \to s$, there then exists $\bar{n}_1$ such that

$$|U(s, \mu^n) - U(s^n, \mu^n)| \leq \varepsilon/2$$

for all $n > \bar{n}_1$. On the other hand, since $u(s, \cdot)$ is a continuous function of x and $\mu^n \to \mu$, $\int u(s, x)\, d\mu^n(x) \to \int u(s, x)\, d\mu(x)$. Hence, there exists $\bar{n} > \bar{n}_1$ such that

$$|U(s, \mu) - U(s, \mu^n)| < \varepsilon/2$$

for all $n > \bar{n}$. Combining the last two displayed inequalities, one obtains that

$$|U(s, \mu) - U(s^n, \mu^n)| < \varepsilon$$

for all $n > \bar{n}$ as desired. ■

A.4 Kakutani's Fixed-Point Theorem

Fixed-point theorems conclude that a correspondence F with some given properties has a fixed-point, in that there exists some x with $x \in F(x)$. There are other such theorems, with varying generality. Kakutani's fixed-point theorem applies for Euclidean spaces (i.e. $\mathbb{R}^n$). It has been used by Nash to prove the existence of Nash equilibrium.

Theorem A.5 (Kakutani's Fixed-Point Theorem) *Let X be a non-empty, compact and convex subset of some Euclidean space. Let also $F : X \rightrightarrows X$ be a non-empty, convex-valued correspondence with a closed-graph. Then, there exists $x \in X$ with $x \in F(x)$.*

In game theoretical applications, X is often the strategy space, which is often assumed to be compact and convex when one includes all mixed strategies, and F is often the set of solutions to an optimization problem, which would often be non-empty and have a closed graph by the Maximum Theorem. In that case, convexity of $F(x)$ is ensured by having a quasiconcave objective function. One can use this theorem creatively, by devising a correspondence F with the above properties and such that the fixed points of F have the desired property (e.g. they are Nash equilibria).

A.5 Theory of Lattices and Supermodularity

This section presents the basic concepts in lattice theory.

A.5.1 Lattices

Lattices are partially-ordered sets where any two elements have a greatest lower bound (sup) and smallest upper bound (sup).

Definition A.1 *A partially-ordered set $(X, \geq)$ is said to be lattice if for all $x, y \in X$*

$$\begin{aligned} x \vee y &\equiv inf\,\{z \in X | z \geq x, z \geq y\} \in X \\ x \wedge y &\equiv sup\,\{z \in X | x \geq z, y \geq z\} \in X. \end{aligned}$$

Operators $\vee$ and $\wedge$ are called *join* and *meet*, respectively. Note that $x \vee y \in X$ is weakly greater than both x and y (i.e., $x \vee y \geq x$ and $x \vee y \geq y$). Moreover, if $z \geq x$ and $z \geq y$, then $z \geq x \vee y$. That is, $x \vee y$ is the smallest upper bound for $\{x, y\}$. Similarly, $x \wedge y$ is the greatest lower bound for $\{x, y\}$ in the sense that $x \geq x \wedge y, y \geq x \wedge y$ and if $x \geq z$ and $y \geq z$, then $x \wedge y \geq z$.

Example A.2 *Let $X = 2^S$ be the set of all subsets of a set S, and order X by set inclusion, i.e., $A \geq B \Longleftrightarrow A \supseteq B$. For any $A, B \in X$, note that $A \cup B \supseteq A$, $A \cup B \supseteq B$ and if $C \supseteq A$ and $C \supseteq B$, then $C \supseteq A \cup B$. Therefore, $A \vee B = A \cup B \in X$. Similarly, $A \wedge B = A \cap B \in X$. Therefore, $(X, \supseteq)$ is a lattice.*

Example A.3 *Endow $\mathbb{R}^n$ with the usual coordinate-wise order:*

$$(x_1, ..., x_n) \geq (y_1, ..., y_n) \Longleftrightarrow \ x_i \geq y_i \quad \forall i.$$

$(\mathbb{R}^n, \geq)$ is a lattice with

$$\begin{aligned} x \vee y &= (max\{x_1, y_1\}, ..., max\{x_n, y_n\}) \\ x \wedge y &= (min\{x_1, y_1\}, ..., min\{x_n, y_n\}). \end{aligned}$$

Definition A.2 *A lattice $(X, \geq)$ is said to be* complete *if for every $S \subseteq X$, a greatest lower bound* $\inf(S)$ *and a least upper bound* $\sup(S)$ *exist in X, where* $\inf(\varnothing) = \sup(X)$ *and* $\sup(\varnothing) = \inf(X)$.

Note that, in the above examples,$(2^S, \supseteq)$ is complete because for any family $A_\alpha \subseteq S, \vee_\alpha A_\alpha = \cup_\alpha A_\alpha \in 2^S$and $\wedge_\alpha A_\alpha = \cap_\alpha A_\alpha \in 2^S$. On the other hand, $(\mathbb{R}^n, \geq)$ is not complete because $\sup(\mathbb{R}^n)$ does not exist.

A.5.2 Strong Set Order and Sublattices

Given a lattice $(X, \geq)$, one can extend the order $\geq$ to subsets of X as follows.

Definition A.3 (Strong Set Order) *Given any lattice $(X, \geq)$, for any $A, B \subseteq X$, write $A \geq B$ if*

$$x \vee y \in A, x \wedge y \in B \qquad (\forall x \in A, y \in B).$$

With the usual order on $\mathbb{R}$, note that

$$\{1, 2, 3, 4\} \geq \{0, 1, 2, 3\}$$

but

$$\{1, 2, 3, 4\} \ngeq \{-0.5, 0.5, 1, 5, 2, 5\}$$

although the set $\{-0.5, 0.5, 1, 5, 2, 5\}$ is obtained by subtracting 0.5 from each element of $\{0, 1, 2, 3\}$. This is a very strong notion of order as it implies many other natural orders on sets. For example, if $A \geq B$, then $\max A \geq \max B$ and $\min A \geq \min B$.

A lattice may have a subset that is a lattice in itself according to original order. Such subsets are called sublattices.

Definition A.4 *Given any lattice $(X, \geq)$, any $S \subseteq X$ is said to be* sublattice *if for any $x, y \in S$, $x \vee y \in S$ and $x \wedge y \in S$.*

The following gives an equivalent definition for sublattices.

Fact A.1 *Given any lattice $(X, \geq)$ and any $S \subseteq X$, S is a sublattice if and only if $S \geq S$.*

For example, under the usual order, $S = \{(x_1, x_2) : x_1 + x_2 \leq 1\}$ is not a sublattice of $\mathbb{R}^2$ because $(1, 0) \vee (0, 1) = (1, 1) \notin S$. On the other hand, $[0, 1]^2$ and $S' = \{(x_1, x_2) : x_1 - x_2 \leq 1\}$ are sublattices.

A.5.3 Functions on Lattices—Supermodularity

I will next introduce important properties of functions to or from lattices. The first property is an elementary monotonicity property, requiring that the order is preserved.

Definition A.5 *Given any partially ordered sets $(T, \geq)$ and $(X, \geq)$, a function $f : T \to X$ is said to be* isotone *(or* weakly increasing*) if*

$$t \geq t' \Rightarrow f(t) \geq f(t').$$

Throughout the lecture, we will take t to be a parameter and investigate how it effects the outcomes according to a solution concept. Since our solution concepts, such as argmax and Nash equilibrium, are set valued, the above definition will be often applied to set-valued functions. Note that, for any lattice $(Y, \geq)$, $(2^Y, \geq)$ is a partially ordered set (with the strong set order).

The second property formalizes the idea of complementarity in terms of functions:

Definition A.6 *Given any lattice $(X, \geq)$, a function $f : X \to \mathbb{R}$ is said to be supermodular if for all $x, y \in X$*

$$f(x \vee y) + f(x \wedge y) \geq f(x) + f(y).$$

A function f is said to be submodular if $-f$ is supermodular.

Note that if X is linearly ordered (as $\mathbb{R}$), then every function $f : X \to \ R$ is supermodular because the above inequality is vacuously satisfied as equality.

When $X = X_1 \times X_2$, ordered coordinate-wise, supermodularity captures the idea of complementarity between X_1 and X_2 precisely. Indeed, if we take $x = (x_1, x_2)$ and $y = (y_1, y_2)$ with $x_1 \geq y_1$ with $y_2 \geq x_2$, we have $x \vee y = (x_1, y_2)$ and $x \wedge y = (y_1, x_2)$.Then, we can write the inequality in the definition of supermodularity as

$$f(x_1, y_2) - f(x_1, x_2) \geq f(y_1, y_2) - f(y_1, x_2).$$

That is, the marginal contribution of increasing the second input from x_2 to y_2 increases when we increase the first input from y_1 to x_1. In other words, marginal contribution of an input is increasing with the other input, capturing the usual meaning of complementarity (as in production theory).

One can also write the above inequality as a condition on the mixed differences:

$$[f(x_1, y_2) - f(x_1, x_2)] - [f(y_1, y_2) - f(y_1, x_2)] \geq 0.$$

This condition reduces to a usual restriction on the cross-derivatives for smooth functions on $\mathbb{R}^2$:

$$\frac{\partial^2 f}{\partial_{x_1} \partial_{x_2}} \geq 0$$

Supermodularity turns out to be closely related to monotone comparative statistics, an ordinal property. Despite this, supermodularity is a cardinal property, as it is not preserved under monotone transformations. The next example illustrates this.

Example A.4 *Let $X = \{0,1\}^2$ and endow it with the usual order. Consider the following supermodular function:*

$$f\ (1,1) = 3, f(1,0) = f(0,1) = 1, f(0,0) = 0.$$

Note that $\sqrt{f}$ is not supermodular.

A.5.4 Increasing Differences and Supermodularity in Product Spaces

In game theoretical applications, the lattices are of the product form as the space of strategy profile is a product set. In such lattices supermodularity reduces to a simpler condition.

For a family of lattices $(X_1, \geq_1), ..., (X_n, \geq_n)$, let $X = X_1 \times \cdots \times X_n$ and endow X with the coordinate-wise order:

$$(x_1, ..., x_n) \geq (y_1, ..., y_n) \iff x_i \geq_i y_i \qquad \forall i.$$

For $x \in X$ and any i and j, define $x_{-ij} = (x_k)_{k \notin \{i,j\}}$. For any function $f : X \to \mathbb{R}$, define $f(\cdot|x_{-ij}) : X_i \times X_j \to \mathbb{R}$ by setting $f(x_i, x_j|x_{-ij}) = f(x_i, x_j, x_{-ij})$. Note that $f\ (\cdot|x_{-ij})$ is the restriction of f to vectors where the entries other than i and j are fixed at x_{-ij}.

Definition A.7 *A function $f : X \to\ \mathbb{R}$ is said to have increasing differences if for any $\left(i, j, x_{-ij}, x_i, x_i', x_j, x_j'\right)$,*

$$\left[x_i \geq x_i' \text{ and } x_j \geq x_j'\right] \Rightarrow f(x_i, x_j, x_{-ij}) - f\ (x_i', x_j, x_{-ij}) \geq f(x_i, x_j', x_{-ij}) - f\ \left(x_i', x_j', x_{-ij}\right).$$

That is, ceteris peribus, the marginal contribution of ith entry (obtained by changing x_i' to higher x_i) is higher when the jth entry is fixed at the higher level of x_j rather than x_j'. When $X = \mathbb{R}^n$, the condition of increasing differences can be called pairwise supermodularity, because the above condition can be written as a supermodularity condition on function $f(\cdot|x_{-ij}) : X_i \times\ X_j \to\ \mathbb{R}$, defined by setting $f(x_i, x_j|x_{-ij}) = f(x_i, x_j, x_{-ij})$. That is,

$$f((x_i, x_j) \vee (x_i', x_j'), x_{-ij}) - f((x_i, x_j), x_{-ij}) \geq f((x_i', x_j'), x_{-ij}) - f((x_i, x_j) \wedge (x_i', x_j'), x_{-ij}).$$

Both increasing differences and pair-wise supermodularity are weaker forms of supermodularity as they are restrictions of supermodularity condition to special sets of cases. (Supermodularity is weakly stronger than pair-wise supermodularity, and pair-wise supermodularity is weakly stronger than increasing difference condition.)

It turns out that supermodularity can be decomposed into increasing differences and supermodularity within each X_i. The following lemma is a main step towards establishing this fact. Its proof also exhibit a common technique of using telescopic equation.

Lemma A.2 *If f has increasing differences and $x_j \geq y_j$ for each j, then for every i,*

$$f(x_i, x_{-i}) - f(y_i, x_{-i}) \geq f(x_i, y_{-i}) - f(y_i, y_{-i}).$$

Proof. Take $i = 1$ without loss of generality. Then,

$$\begin{aligned} f(x_1, x_{-1}) - f\ (x_1, y_{-1}) &= \sum_{j>1} [f(x_1, ..., x_{j-1}, x_j, y_{j+1}, ..., y_n) - f(x_1, ..., x_{j-1}, y_j, ..., y_n)] \\ &\geq \sum_{j>1} [f(y_1, ..., x_{j-1}, x_j, y_{j+1}, ..., y_n) - f(y_1, ..., x_{j-1}, y_j, ..., y_n)] \\ &= f(y_1, x_{-1}) - f(y_1, y_{-1}). \end{aligned}$$

Here, the first and the last equalities are telescopic equations, writing the whole difference as a sum of one step changes. The inequality is by increasing differences: for any j, by increasing differences between 1 and j,

$$\begin{aligned} & f(x_1, ..., x_{j-1}, x_j, y_{j+1}, ..., y_n) - f(x_1, ..., x_{j-1}, y_j, ..., y_n) \\ \geq \quad & f(y_1, ..., x_{j-1}, x_j, y_{j+1}, ..., y_n) - f(y_1, ..., x_{j-1}, y_j, ..., y_n), \end{aligned}$$

and one obtains the inequality by summing up both sides. Of course, this is equivalent to the statement in the lemma. ■

Lemma extends the increasing differences condition from comparison of two entries to the comparison of two vectors, establishing a (seemingly) stronger increasing difference condition. This further implies that, in product spaces, supermodularity can be decomposed into increasing differences and supermodularity within each X_i, as established next.

Proposition A.2 *For any product lattice* $(X, \geq)$ *(with* $X = X_1 \times \cdots \times X_n$ *and coordinate-wise order) and for any function* $f : X \to \mathbb{R}$, f *is supermodular if and only if*

1. f *has increasing differences and*

2. f *is supermodular within* X_i *for each* i *(i.e.,*

$$f(x_i \vee y_i, x_{-i}) + f(x_i \wedge y_i, x_{-i}) \geq f(x_i, x_{-i}) + f(y_i, x_{-i})$$

for all $x_i, y_i \in X_i$ *and* $x_{-i} \in X_{-i}$*).*

Proof. Supermodularity implies increasing differences (1) and supermodularity within each coordinate (2) by definition. To prove the converse, take any $x, y \in X$ and assume the conditions (1) and (2) in the proposition. Then,

$$\begin{aligned}
&f(x \vee y) - f(y) \\
&= \sum_{i=1}^{n} [f(x_1 \vee y_1, ..., x_{i-1} \vee y_{i-1}, x_i \vee y_i, y_{i+1}, ..., y_n) - f(x_1 \vee y_1, ..., x_{i-1} \vee y_{i-1}, y_i, y_{i+1}, ..., y_n)] \\
&\geq \sum_{i=1}^{n} [f(x_1 \vee y_1, ..., x_{i-1} \vee y_{i-1}, x_i, y_{i+1}, ..., y_n) - f(x_1 \vee y_1, ..., x_{i-1} \vee y_{i-1}, x_i \wedge y_i, y_{i+1}, ..., y_n)] \\
&\geq \sum_{i=1}^{n} [f(x_1, ..., x_{i-1}, x_i, x_{i+1} \wedge y_{i+1}, ..., x_n \wedge y_n) - f(x_1, ..., x_{i-1}, x_i \wedge y_i, x_{i+1} \wedge y_{i+1}, ..., x_n \wedge y_n)] \\
&= f(x) - f(x \wedge y)]
\end{aligned}$$

Here, the first and the last equalities are telescopic equations. To obtain the first inequality, observe that for every i, by part 2 in the hypothesis, the term in the second line is at least as high as the term in the third line. To obtain the second inequality, for each i, use the increasing differences property to apply Lemma A.2 and conclude that the the term in the third line is at least as high as the term in the fourth line. Each inequality is then obtained by summing up the inequalities over i. ■

An immediate corollary to the proposition is that supermodularity reduces to pairwise-supermodularity on $\mathbb{R}^n$:

Corollary A.1 *For any* $f : \mathbb{R}^n \to \mathbb{R}$, *the following are equivalent:*

1. f *is supermodular*

2. *f has increasing differences;*

3. *f is pair-wise supermodular.*

Proof. Since $\mathbb{R}$ is linearly ordered, every function is supermodular on $\mathbb{R}$. The corollary then follows from Proposition 1 and the definitions. ■

Hence, supermodularity is a pair-wise concept on $\mathbb{R}^n$—and in fact on any product of linearly-ordered sets. This is because it reflects the pair-wise concept of complementarity. As stated in terms of constraints, complementarity is also captured by sublattices. It turns out that sublattices can be reduced to pair-wise constraints on $\mathbb{R}^n$.

Proposition A.3 *Let X be a sublattice of $\mathbb{R}^n$ (under coordinate-wise order). For all i, j, define*

$$\begin{aligned} C_{ij} &= \{(x_i, x_j) | (x_i, x_j, x_{-ij}) \in X \textit{ for some } x_{-ij}\} \\ S_{ij} &= C_{ij} \times \prod_{k \neq i,j} \mathbb{R} \end{aligned}$$

Then,

$$X = \bigcap_{i,j} S_{ij}.$$

That is, a sublattice X can be written as a set of pair-wise constraints: $x \in X$ if

$$(x_i, x_j) \in C_{i,j} \qquad \forall i, j.$$

This limits the applicability of lattice theory and the analyses in these lectures substantially because many important constraints cannot be stated as sublattices. For example, when there are three or more goods, a budget set $\{x | \ (x - \bar{x}) \cdot p \leq 0\}$ cannot be a lattice under usual order or its reverse.

A.5.5 Order Topology and Continuity

In a lattice, the order induces the relevant concepts of continuity and convergence. I will conclude this section by describing these concepts, which will be used in the analyses of supermodular games.

Consider a complete lattice $(X, \geq)$. Consider any monotone sequence x_n in X. Since $(X, \geq)$ is complete,

$$\sup\{x_n|n \in \mathbb{N}\} \text{ and } \inf\{x_n|n \in \mathbb{N}\}$$

exist. For any weakly increasing sequence x_n (with $x_{n+1} \geq x_n$ for all n), it is natural to think that x_n converges to sup $\{x_n|n \in \mathbb{N}\}$. Similarly, for any weakly decreasing sequence x_n (with $\{x_n \geq x_{n+1}$for all n), it is natural to think that x_n converges to inf $\{x_n|n \in \mathbb{N}\}$. The *order topology* is the smallest topology in which every weakly increasing sequence x_n converges to its supremum

$$\lim x_n \equiv \sup x_n,$$

and every weakly decreasing sequence x_n converges to its infimum

$$\lim x_n \equiv \inf x_n,$$

Definition A.8 *A function $f : X \to Y$ (where Y is any topological space, such as $\mathbb{R}$), f is said to be continuous (in the order topology) if for every monotone sequence x_n,*

$$\lim f(x_n) = f(\lim x_n).$$

That is, for every weakly increasing sequence x_n, lim $f(x_n) = f(\sup\ x_n)$, and for every weakly decreasing x_n, lim $f(x_n) = f(\inf\ x_n)$.

A.5.6 Exercises

Exercise A.2 *For some lattice $(X, \geq)$, consider supermodular functions $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$. Prove or disprove the following.*

1. *For any $a, b \geq 0, af + bg$ is supermodular.*

2. *If f and g are isotone and nonnegative (i.e. $f(x) \geq 0$ and $g(x) \geq 0$ for all x), then fg is supermodular.*

3. *Under the conditions in part (b), f^α is supermodular for any $\alpha \geq 1$.*

Exercise A.3 *This question asks you to prove some basic facts.*

1. *For any lattice $(X, \geq)$ show that if S and T are sublattices, so is $S \cap T$.*

2. *Consider a function* $f : X \times \Omega \to \mathbb{R}$, *where* $(X, \geq)$ *is a lattice and* Ω *is a probability space with expectation operator* E. *Show that if* $f\ (\cdot, \omega)$ *is supermodular for all* $\omega \in \Omega$, *then* $E[f]$ *is also supermodular.*

3. *Let* X *be a set of sets that is closed under union (i.e.* $A \cup B \in X$ *for all* $A, B \in X$*) and with* $\varnothing \in X$. *Define order* $\geq$ *on* X *by* $A \geq B \iff A \supseteq B$. *Show that* $(X, \geq)$ *is a complete lattice. What are* $A \vee B$ *and* $A \wedge B$*?*

4. *On the set* $\mathbb{R}_+^{\mathbb{R}+}$ *of functions* $f : \mathbb{R}_+ \to \mathbb{R}_+$, *define order* $\geq$ *by* $f \geq g \iff f(x) \geq g\ (x)$ *for all* $x \in \mathbb{R}$. *Show that* $(\mathbb{R}_+^{\mathbb{R}+}, \geq)$ *is a lattice. Show also that the following are sublattices:*

   (a) *all continuous functions,*

   (b) *all non-increasing functions, and*

   (c) *all functions* f *with* $f \leq g$ *for some fixed function* g.

Exercise A.4 *Prove the following statements.*

1. *If* f *and* g *are supermodular, so is* $f + g$.

2. *If* f *is supermodular and* $a > 0$, *then* af *is also supermodular.*

3. *If* $f : \Theta \times X \to \mathbb{R}$, *where*

   - X *is a lattice,*
   - $\theta \in \Theta$ *is not known,*
   - $f(\theta, \cdot) : X \to \mathbb{R}$ *is supermodular for each* $\theta \in \Theta$,

   then $E[f] : X \to \mathbb{R}$ *is supermodular, where* E *is an expectation operator on* Θ.

Exercise A.5 *Let* X *be a complete lattice and* $f : X \to X$ *be isotone. (Do not assume that* f *is continuous.) Define*

$$\underline{x} = \inf\{x | f(x) \leq x\}$$
$$\overline{x} = \sup\{x | f(x) \geq x\}.$$

Show that $\overline{x}$ *and* $\underline{x}$ *are fixed points of* f *, i.e.,* $\overline{x} = f\ (\overline{x})$ *and* $x = f\ (\underline{x})$. *Show also that, if* $f\ (x) = x$, *then* $\underline{x} \leq x \leq \overline{x}$.

Exercise A.6 *For each case below, show that $(X, \geq)$ is a lattice. Determine the join and meet operators and check whether it is complete.*

1. *X is the set of all probability distributions on the real line; $\geq$ is the relation of first-order stochastic dominance.*

   Hint: You can take X as the set of CDFs $F : \mathbb{R} \to [0,1]$ and write

   $$F \geq G \iff [F(x) \leq G(x) \qquad \forall x].$$

   If you feel more comfortable, you can confine X to continuous CDFs and/or restrict the domain to $[0,1]$.

2. *X is the set of all partitions of a fixed set A. $\geq$ is the refinement ordering: for any $P, P' \in X$, $P \geq P'$ if and only if P is finer than P', i.e., for any $S \in P$, $S' \in P'$, if $S \cap S' \neq \varnothing$, then $S \subseteq S'$. (You can take A finite if you feel more comfortable.)*

3. *Fix a finite type space (Θ^*, T^*, p), where $T^* = T_1^* \times \cdots \times T_n^*$ and each type $t_i \in T_i^*$ is associated a belief $p_{t_i} \in \Delta(\Theta^* \times T_{-i}^*)$. A* belief-closed subspace *is a pair (Θ, T), with T a nonempty set of the form $T_1 \times \cdots \times T_n$, where $\Theta \subseteq \Theta^*$ and $T_i \subseteq T_i^*$ for each i, and such that $p_{t_i}(\Theta \times T_{-i}) = 1$ for each i and each $t_i \in T_i$. Take X to be the set of all belief-closed subspaces, together with $(\varnothing, \varnothing)$, and the ordering to be set inclusion: $(\Theta, T) \geq (\Theta', T')$ if $\Theta \supseteq \Theta'$ and $T \supseteq T'$.*