Decision Analysis & Decision Support
6.872/HST.950
Tasks?

- Mechanics
  - Record keeping
  - Administration
  - Scheduling
  - …
- Diagnosis
- Prognosis
- Therapy
Types of Decision Support

• “Doctor's Assistant” for clinicians at any level of training
• Expert (specialist) consultation for non-specialists
• Monitoring and error detection
• Critiquing, what-if
• Guiding patient-controlled care
• Education and Training
• Contribution to medical research
• …
Two Historical Views on How to Build Expert Systems

- Great cleverness
  - Powerful inference abilities
  - *Ab initio* reasoning

- Great stores of knowledge
  - Possibly limited ability to infer, but
  - Vast storehouse of relevant knowledge, indexed in an easy-to-apply form
Change over 30 years

- 1970’s: human knowledge, not much data
- 2000’s: vast amounts of data, traditional human knowledge (somewhat) in doubt

- Could we “re-discover” all of medicine from data? *I think not!*
- Should we focus on methods for reasoning with uncertain data? *Absolutely!*
Cancer Test

• We discover a cheap, 95% accurate test for cancer.
• Give it to “Mrs. Jones”, the next person who walks by 77 Mass Ave.
• Result is positive.
• What is the probability that Mrs. Jones has cancer?
Figuring out Cancer Probability

Assume Ca in 1% of general population:

\[
\frac{950}{950 + 4950} = .161
\]
At the Extremes

• If Ca probability in population is 0.1%,
  – Then post positive result, \( p(Ca) = 1.87\% \)

• If Ca probability in population is 50%,
  – Then post-positive result, \( p(Ca) = 95\% \)
Bayes’ Rule

\[ P(D \mid T) = \frac{P(D)P(T \mid D)}{P(D)P(T \mid D) + P(\overline{D})P(T \mid \overline{D})} \]
Odds/Likelihood Form

\[ P(D \mid T) = \frac{P(D)P(T \mid D)}{P(D)P(T \mid D) + P(\overline{D})P(T \mid \overline{D})} \]

\[ P(\overline{D} \mid T) = \frac{P(\overline{D})P(T \mid \overline{D})}{P(D)P(T \mid D) + P(\overline{D})P(T \mid \overline{D})} \]

\[
\frac{P(D \mid T)}{P(\overline{D} \mid T)} = \frac{P(D)}{P(\overline{D})} \frac{P(T \mid D)}{P(T \mid \overline{D})}
\]

\[ O(D \mid T) = O(D)L(T \mid D) \]

\[ W(D \mid T) = W(D) + W(T \mid D) \]
DeDombal, et al. Experience 1970’s & 80’s

• “Idiot Bayes” for appendicitis
• 1. Based on expert estimates -- lousy
• 2. Statistics -- better than docs
• 3. Different hospital -- lousy again
• 4. Retrained on local statistics -- good
Rationality

• Behavior is a continued sequence of choices, interspersed by the world’s responses
• Best action is to make the choice with the greatest expected value
• … decision analysis
Example: Gangrene

- From Pauker’s “Decision Analysis Service” at New England Medical Center Hospital, late 1970’s.
- Man with gangrene of foot
- Choose to amputate foot or treat medically
- If medical treatment fails, patient may die or may have to amputate whole leg.
- What to do? How to reason about it?
Decision Tree for Gangrene

- **amputate foot**
  - live (.99) 850
  - die (.01) 0

- **medicine**
  - full recovery (.7) 1000
    - worse (.25)
      - die (.05) 0
      - amputate leg
        - live (.98) 700
        - die (.02) 0
          - live (.6) 995
          - die (.4) 0
Evaluating the Decision Tree

- **amputate foot**: 871.5
  - live (.99): 850
  - die (.01): 0
- **medicine**: 871.5
  - full recovery (.7): 1000
    - worse (.25): 686
      - die (.05): 0
      - live (.98): 700
  - amputate leg: 686
    - live (.98): 700
    - die (.02): 0
  - medicine: 597
    - die (.4): 0
    - live (.6): 995
Decision Analysis:
Evaluating Decision Trees

- Outcome: directly estimate value
- Decision: value is that of the choice with the greatest expected value
- Chance: expected value is sum of (probabilities x values of results)
- “Fold back” from outcomes to current decision.
- Sensitivity analyses often more important than result(!)
HELP System uses D.A.

Intellectual Tools

Effect of age patient and $M_C$ (mortality for appendicitis without operation) on the probability threshold (point of crossing zero $\Delta u$ line) for decision to operate.

Utility Analysis of Appendectomy

4: Computer Representation of Medical Knowledge

Effect of patient's salary and assumed value of one day of good health ($70 or $140) on decision to operate for appendicitis.

PROB OF APPENDICITIS

A APPENDICITIS BY HISTORY
B REBOUND TENDERNESS IN RLQ
C PRIOR APPENDECTOMY
D IF C THEN EXIT
E WHITE BLOOD COUNT (WBC×100) TH/M3, LAST

F PROB B A 620 90
G PROB F 43 18 9, 74 23 7, 93 18 11, 108 10 11, 121 16 13, 134 6 16, 151 5 16, 176 4 14

FVAL G
UTILITY OF APPENDECTOMY IS ESTIMATED AS $----$

A  (A) AGE
B  SEX
C  (A) SALARY, GET A/365
D  JOB, PERCENT ACTIVITY NEEDED
E  LE A,B
F  DLOS D 30 1, 65 2, 80 4, 90 1, 100 – 0
G  DLOS D 40 1, 80 4, 95 5, 100 – 0 ...
I  COND E, F, 7, 1800, 0, C
J  COND E, G, 1, 900, 0, C ...
M  PROB OF APPENDICITIS
N  UTIL M, I, J, K, L
O  IF N LT 0, EXIT
FVAL N
“Paint the Blackboards!”

DECISION  PATIENT STATE  UTILITY

- Treat disease
- No disease (1-p)  Treat no disease
- treat
- Disease (p)
- No treat
- No disease (1-p)  No treat no disease
- No treat
Threshold

- Benefit $B = U(\text{treat dis}) - U(\text{no treat dis})$
- Cost $C = U(\text{no treat no dis}) - U(\text{treat no dis})$
- Threshold probability for treatment:

$$T = \frac{1}{\frac{B}{C} + 1}$$

Pauker, Kassirer, NEJM 1975
Test/Treat Threshold

Figure removed due to copyright restrictions.
Visualizing Thresholds

Figure removed due to copyright restrictions.
More Complex Decision Analysis Issues

- Repeated decisions
- Accumulating disutilities
- Dependence on history
- Cohorts & state transition models
- Explicit models of time
- Uncertainty in the uncertainties
- Determining utilities
  - Lotteries, …
- Qualitative models
Example: Acute Renal Failure

• Choice of a handful (8) of therapies (antibiotics, steroids, surgery, etc.)
• Choice of a handful (3) of invasive tests (biopsies, IVP, etc.)
• Choice of 27 diagnostic “questions” (patient characteristics, history, lab values, etc.)
• Underlying cause is one of 14 diseases
  – We assume one and only one disease
Decision Tree for ARF

• Choose:
  – Surgery for obstruction
  – Treat with antibiotics
  – Perform pyelogram
  – Perform arteriography
  – Measure patient’s temperature
  – Determine if there is proteinuria
  – …
Decision Tree for ARF

- Surgery for obstruction
- Treat with antibiotics
- Perform pyelogram
- Perform arteriography
- Measure patient’s temperature
- Determine if there is proteinuria

Value = ???
What happens when we act?

• Treatment: leads to few possible outcomes
  – different outcomes have different probabilities
    • probabilities depend on distribution of disease probabilities
  – value of outcome can be directly determined
    • value may depend on how we got there (see below)
    • therefore, value of a treatment can be determined by expectation

• Test: lead to few results, revise probability distribution of diseases, and impose disutility

• Questions: lead to few results, revise probability distribution
Full decision tree

P

A2

A1

P'

A1

A2

A3

A4

P''

A1

A2

A3

A4
Initial probability distribution

<table>
<thead>
<tr>
<th>Condition</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute tubular necrosis</td>
<td>0.250</td>
</tr>
<tr>
<td>Functional acute renal failure</td>
<td>0.400</td>
</tr>
<tr>
<td>Urinary tract obstruction</td>
<td>0.100</td>
</tr>
<tr>
<td>Acute glomerulonephritis</td>
<td>0.100</td>
</tr>
<tr>
<td>Renal cortical necrosis</td>
<td>0.020</td>
</tr>
<tr>
<td>Hepatorenal syndrome</td>
<td>0.005</td>
</tr>
<tr>
<td>Pyelonephritis</td>
<td>0.010</td>
</tr>
<tr>
<td>Atheromatous Emboli</td>
<td>0.003</td>
</tr>
<tr>
<td>Renal infarction (bilateral)</td>
<td>0.002</td>
</tr>
<tr>
<td>Renal vein thrombosis</td>
<td>0.002</td>
</tr>
<tr>
<td>Renal vasculitis</td>
<td>0.050</td>
</tr>
<tr>
<td>Scleroderma</td>
<td>0.002</td>
</tr>
<tr>
<td>Chronic glomerulonephritis, acute exacerbation</td>
<td>0.030</td>
</tr>
<tr>
<td>Malignant hypertension &amp; nephrosclerosis</td>
<td>0.030</td>
</tr>
</tbody>
</table>
## ARF’s Database: P(obs|D)

### Conditional probabilities for Proteinuria Diseases

<table>
<thead>
<tr>
<th>Diseases</th>
<th>Trace 0 to 2+</th>
<th>3+ to 4+</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATN</td>
<td>0.1</td>
<td>0.8</td>
</tr>
<tr>
<td>FARF</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>OBSTR</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>AGN</td>
<td>0.01</td>
<td>0.2</td>
</tr>
<tr>
<td>CN</td>
<td>0.01</td>
<td>0.8</td>
</tr>
<tr>
<td>HS</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>PYE</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>AE</td>
<td>0.1</td>
<td>0.8</td>
</tr>
<tr>
<td>RI</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td>RVT</td>
<td>0.0001</td>
<td>0.1</td>
</tr>
<tr>
<td>VASC</td>
<td>0.01</td>
<td>0.2</td>
</tr>
<tr>
<td>SCL</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>CGAE</td>
<td>0.0001</td>
<td>0.2</td>
</tr>
<tr>
<td>MH</td>
<td>0.0001</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Questions

- Blood pressure at onset
- proteinuria
- casts in urine sediment
- hematuria
- history of prolonged hypotension
- urine specific gravity
- large fluid loss preceding onset
- kidney size
- urine sodium
- strep infection within three weeks
- urine volume
- recent surgery or trauma
- age
- papilledema
- flank pain

- history of proteinuria
- symptoms of bladder obstruction
- exposure to nephrotoxic drugs
- disturbance in clotting mechanism
- pyuria
- bacteriuria
- sex
- transfusion within one day
- jaundice or ascites
- ischemia of extremities or aortic aneurism
- atrial fibrillation or recent MI
Invasive tests and treatments

- Tests
  - biopsy
  - retrograde pyelography
  - transfemoral arteriography

- Treatments
  - steroids
  - conservative therapy
  - iv-fluids
  - surgery for urinary tract obstruction
  - antibiotics
  - surgery for clot in renal vessels
  - antihypertensive drugs
  - heparin
Updating probability distribution

\[ P_{i+1}(D_j) = \frac{P_i(D_j)P(S|D_j)}{\sum_{k=1}^{n} P_i(D_k)P(S|D_k)} \]

Bayes’ rule
Value of treatment

• Three results: improved, unchanged, worsened
  – each has an innate value, modified by “tolls” paid on the way

• Probabilities depend on underlying disease probability distribution

```
      I       V(I)
   p_I  /      /
  ---  |      |
  Tx   V U     W
          V(U) V(W)
```

```jupyter
I V(I)
U V(U)
W V(W)
```
Modeling treatment

<table>
<thead>
<tr>
<th>Steroids</th>
<th>improved</th>
<th>unchanged</th>
<th>worse</th>
</tr>
</thead>
<tbody>
<tr>
<td>atrn</td>
<td>0.60</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>farf</td>
<td>0.05</td>
<td>0.35</td>
<td>0.60</td>
</tr>
<tr>
<td>obstr</td>
<td>0.05</td>
<td>0.60</td>
<td>0.35</td>
</tr>
<tr>
<td>agn</td>
<td>0.40</td>
<td>0.40</td>
<td>0.20</td>
</tr>
<tr>
<td>cn</td>
<td>0.05</td>
<td>0.75</td>
<td>0.20</td>
</tr>
<tr>
<td>hs</td>
<td>0.05</td>
<td>0.05</td>
<td>0.90</td>
</tr>
<tr>
<td>pye</td>
<td>0.05</td>
<td>0.05</td>
<td>0.90</td>
</tr>
<tr>
<td>ae</td>
<td>0.05</td>
<td>0.70</td>
<td>0.25</td>
</tr>
<tr>
<td>ri</td>
<td>0.01</td>
<td>0.14</td>
<td>0.85</td>
</tr>
<tr>
<td>rvt</td>
<td>0.10</td>
<td>0.30</td>
<td>0.60</td>
</tr>
<tr>
<td>vasc</td>
<td>0.15</td>
<td>0.25</td>
<td>0.60</td>
</tr>
<tr>
<td>scl</td>
<td>0.05</td>
<td>0.05</td>
<td>0.90</td>
</tr>
<tr>
<td>cgae</td>
<td>0.40</td>
<td>0.35</td>
<td>0.25</td>
</tr>
<tr>
<td>mh</td>
<td>0.05</td>
<td>0.05</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Utilities:

improved: 5000
unchanged: -2500
worse: -5000
Modeling test:
transfemoral arteriography

<table>
<thead>
<tr>
<th></th>
<th>p(clot)</th>
<th>cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>atn</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>farf</td>
<td>0.01</td>
<td>800</td>
</tr>
<tr>
<td>obstr</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>agn</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>cn</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>hs</td>
<td>0.01</td>
<td>800</td>
</tr>
<tr>
<td>pye</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>ae</td>
<td>0.03</td>
<td>800</td>
</tr>
<tr>
<td>ri</td>
<td>0.85</td>
<td>500</td>
</tr>
<tr>
<td>rvt</td>
<td>0.50</td>
<td>500</td>
</tr>
<tr>
<td>vasc</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>scl</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>cgae</td>
<td>0.01</td>
<td>500</td>
</tr>
<tr>
<td>mh</td>
<td>0.01</td>
<td>500</td>
</tr>
</tbody>
</table>
How large is the tree?

- Infinite, or at least \((27+3+8)^{(27+3+8)}\), \(\sim 10^{60}\)
- What can we do?
  - Assume any action is done only once
  - Order:
    - questions
    - tests
    - treatments
- \(27! \times 4 \times 3 \times 2 \times 8\), \(\sim 10^{30}\)
- Search, with a *myopic evaluation function*
  - like game-tree search; what’s the static evaluator?
  - Measure of certainty in the probability distribution
How many questions needed?

- How many items can you distinguish by asking 20 (binary) questions? $2^{20}$
- How many questions do you need to ask to distinguish among $n$ items? $\log_2(n)$
- *Entropy* of a probability distribution is a measure of how certainly the distribution identifies a single answer; or how many more questions are needed to identify it.
Entropy of a distribution

\[ H_i(P_1, \ldots, P_n) = \sum_{j=1}^{n} - P_j \log_2 P_j \]

For example:
\( H(.5, .5) = 1.0 \)
\( H(.1, .9) = 0.47 \)
\( H(.01, .99) = 0.08 \)
\( H(.001, .999) = 0.01 \)
\( H(.33, .33, .33) = 1.58 (!) \)
\( H(.005, .455, .5) = 1.04 \)
\( H(.005, .995, 0) = 0.045 \)

(!) -- should use \( \log_n \)
Interacting with ARF in 1973

Question 1: What is the patient's age?
1  0-10
2  11-30
3  31-50
4  51-70
5  Over 70
Reply: 5

The current distribution is:
Disease    Probability
FARF        0.58
IBSTR      0.22
ATN         0.09

Question 2: What is the patient's sex?
1  Male
2  Pregnant Female
3  Non-pregnant Female
Reply: 1

...
ARF in 1994
Local Sensitivity Analysis

[Image of a flowchart or diagram related to Acute Renal Failure Program with various parameters and their sensitivities indicated.]
Case-specific Likelihood Ratios
Therapy Planning Based on Utilities

The following facts are known about this patient:
- Age: Over 70; Sex: Male; Blood Pressure At Onset: Moderately Elevated;
- Urine Volume: 50 400 Cc Day; Kidney Size: Large; Large Fluid Loss Preceding Onset: No;
- Proteinuria: Zero; History Of Prolonged Hypotension: No.

This leads to the probability distribution over the diseases:
- ATH: 0.000; FARF: 0.006; OBSTR: 0.966; AGH: 0.000; CN: 0.000; HS: 0.000; PYE: 0.027;
- AE: 0.000; RI: 0.000; RUT: 0.000; VASC: 0.000; SCL: 0.000; CGAE: 0.000; MH: 0.000.

Plans for further testing and treatment (in descending value order) are:
Calculating full plan...
Determining best plan...
Plan number 1:
- Therapy SURGERY-FOR-URINARY-TRACT-OBSTRUCTION has ev=2862.9 (v=2862.9)

Plan number 2:
- Action RETROGRADE-PYEOGRAPHY, with possible outcomes giving ev=2400.1:
  - Outcome 0 (OBSTRUCTION), with p=0.9569
    - Best decision gives ev=2621.8:
      - Therapy SURGERY-FOR-URINARY-TRACT-OBSTRUCTION has ev=2621.8 (v=3122.3)
  - Outcome 1 (NO-OBSTRUCTION), with p=0.0431
    - Best decision gives ev=-2525.9:
      - Therapy ANTIBIOTICS has ev=-2525.9 (v=-1025.3)

Plan number 3:
- Action TRANSFEMORAL-ARTERIOGRAPHY, with possible outcomes giving ev=2361.0:
  - Outcome 0 (CLOT), with p=0.0100
    - Best decision gives ev=2359.4:
      - Therapy SURGERY-FOR-URINARY-TRACT-OBSTRUCTION has ev=2359.4 (v=2861.3)
  - Outcome 1 (NO-CLOT), with p=0.9900
    - Best decision gives ev=2361.0:
      - Therapy SURGERY-FOR-URINARY-TRACT-OBSTRUCTION has ev=2361.0 (v=2862.9)

Plan number 4:
- Action BIOPSY, with possible outcomes giving ev=1862.8:
Assumptions in ARF

- Exhaustive, mutually exclusive set of diseases
- Conditional independence of all questions, tests, and treatments
- Cumulative (additive) disutilities of tests and treatments
- Questions have no modeled disutility, but we choose to minimize the number asked anyway