Decision Support via Expert Systems

6.872/HST950

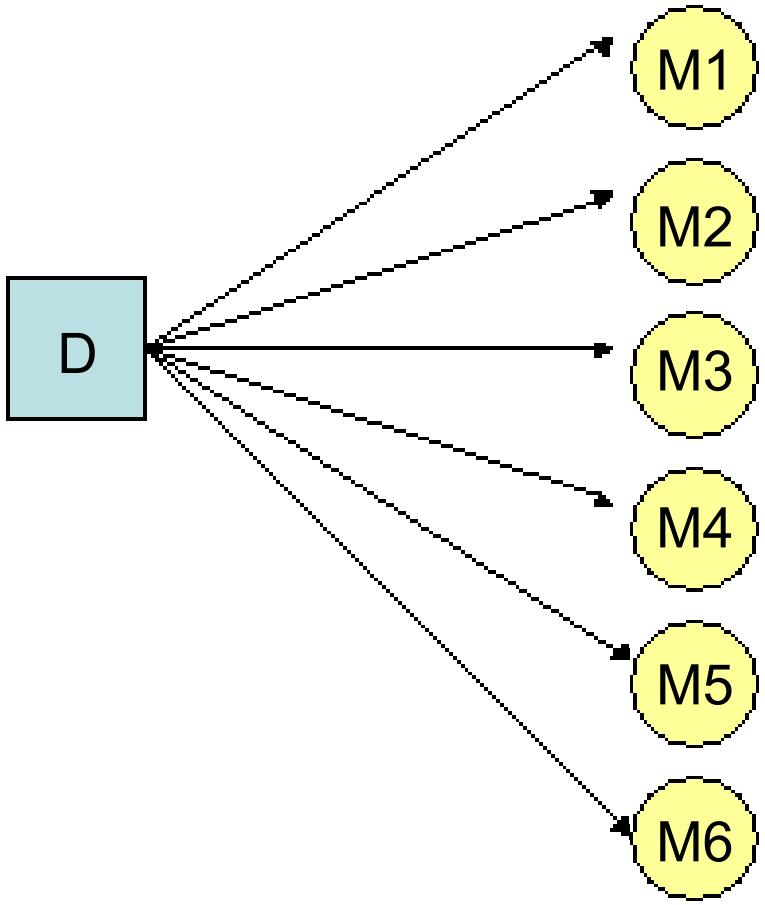
Harvard-MIT Division of Health Sciences and Technology
HST.950J: Medical Computing
Peter Szolovits, PHD

Components of an Expert System

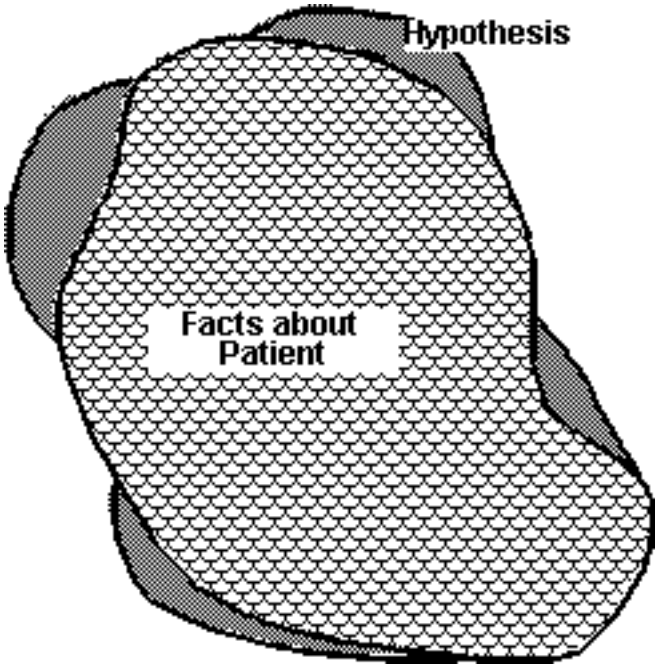
- Knowledge
  - In various forms: associations, models, etc.
- Strategy
  - Baconian, exhaustive enumeration, on-line, etc.
- Implementation
  - Programs, pattern matching, rules, etc.

Last Time

- Naïve Bayesian Inference
  - Exhaustive and Mutually Exclusive disease hypotheses (1 and only 1)
  - Conditionally independent observables (manifestations)
  - $P(D_i)$, $P(M_{ij}|D_i)$



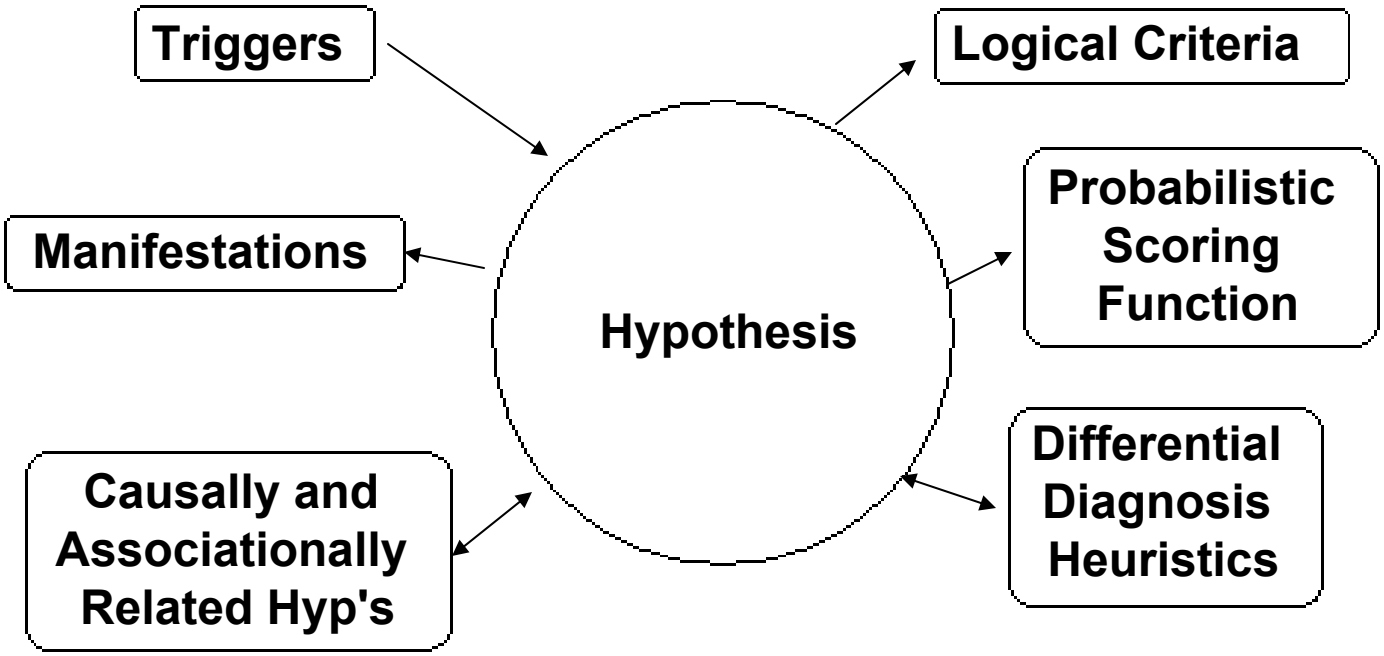
Taking the Present Illness—Diagnosis by Pattern Directed Matching



PIP's Theory of Diagnosis

- **From initial complaints, *guess* suitable hypothesis.**
- **Use current active hypotheses to guide questioning**
- **Failure to satisfy expectations is the strongest clue to a better hypothesis; *differential diagnosis***
- **Hypotheses are *activated, de-activated, confirmed* or *rejected* based on**
  - **(1) logical criteria**
  - **(2) probabilities based on:**
- **findings local to hypothesis**
- **causal relations to other hypotheses**

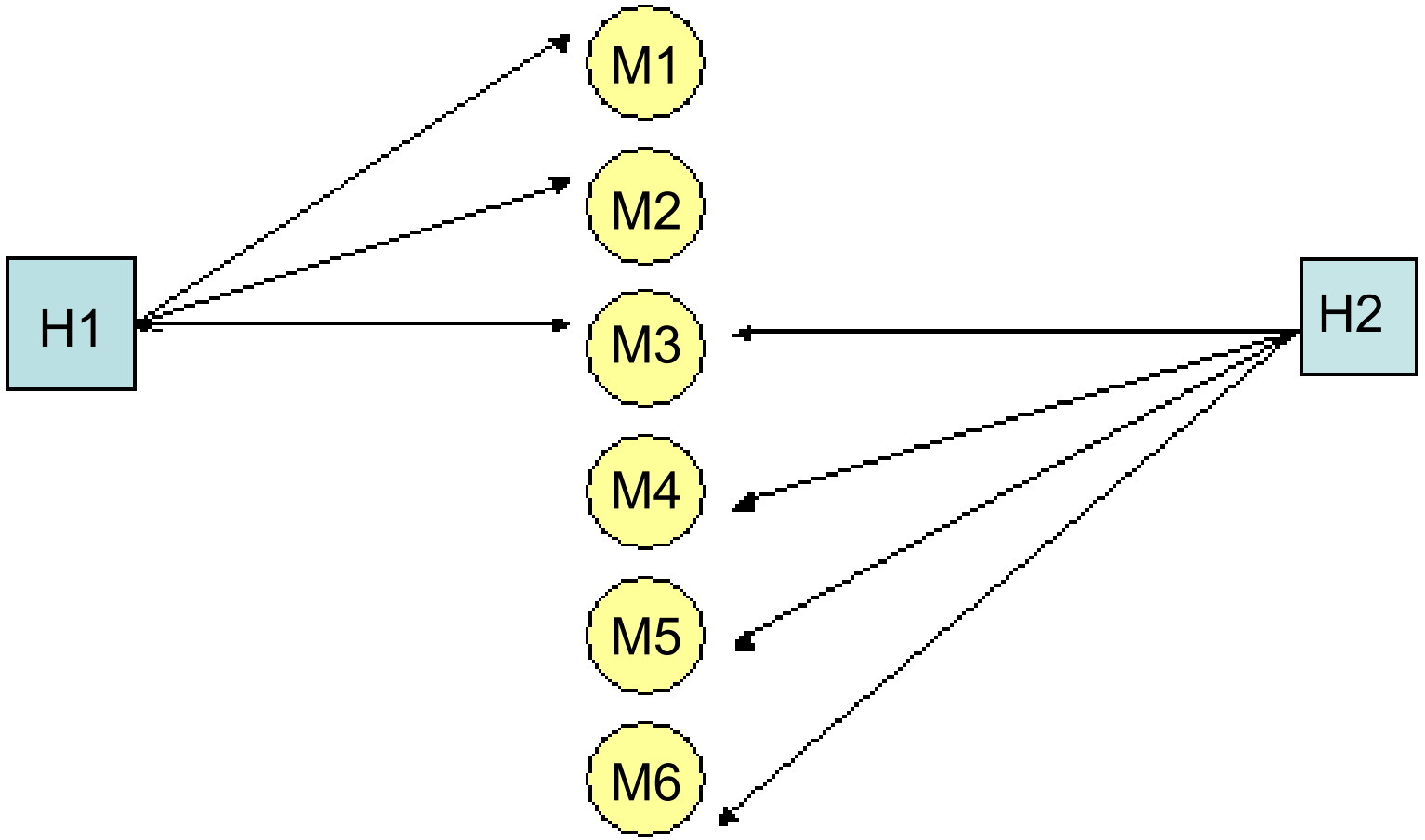
Memory Structure in PIP



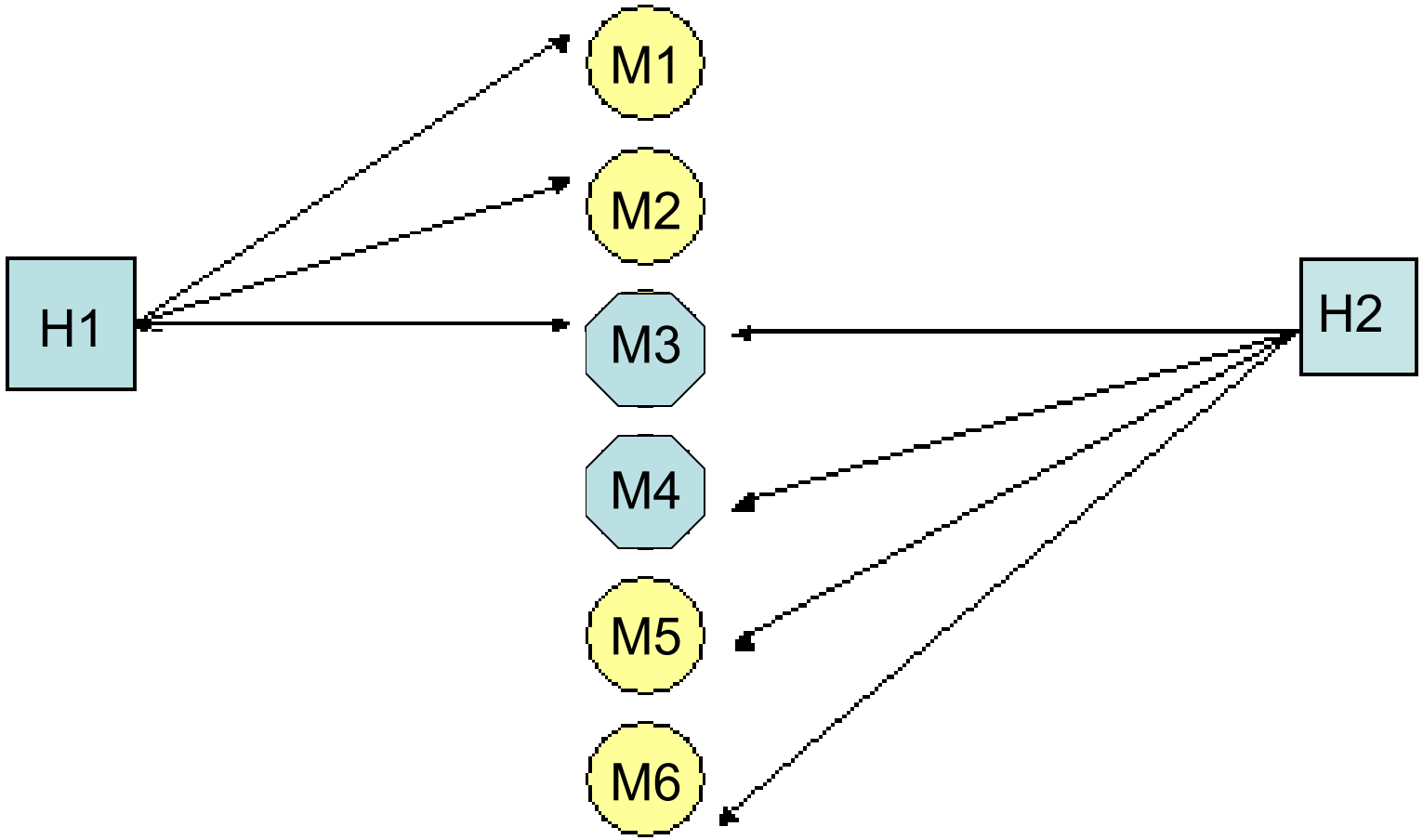
PIP's Model of Nephrotic Syndrome

- **NEPHROTIC SYNDROME, *a clinical state***
- ***FINDINGS:***
  - **1* Low serum albumin concentration**
  - **2. Heavy proteinuria**
  - **3* >5 gm/day proteinuria**
  - **4* Massive symmetrical edema**
  - **5* Facial or peri-orbital symmetric edema**
  - **6. High serum cholesterol**
  - **7. Urine lipids present**
- ***IS-SUFFICIENT:* Massive pedal edema & >5 gm/day proteinuria**
- ***MUST-NOT-HAVE:* Proteinuria absent**
- ***SCORING . . .***
- ***MAY-BE-CAUSED-BY:* AGN, CGN, nephrotoxic drugs, insect bite, idiopathic nephrotic syndrome, lupus, diabetes mellitus**
- ***MAY-BE-COMPLICATED-BY:* hypovolemia, cellulitis**
- ***MAY-BE-CAUSE-OF:* sodium retention**
- ***DIFFERENTIAL DIAGNOSIS:***
  - **neck veins elevated**

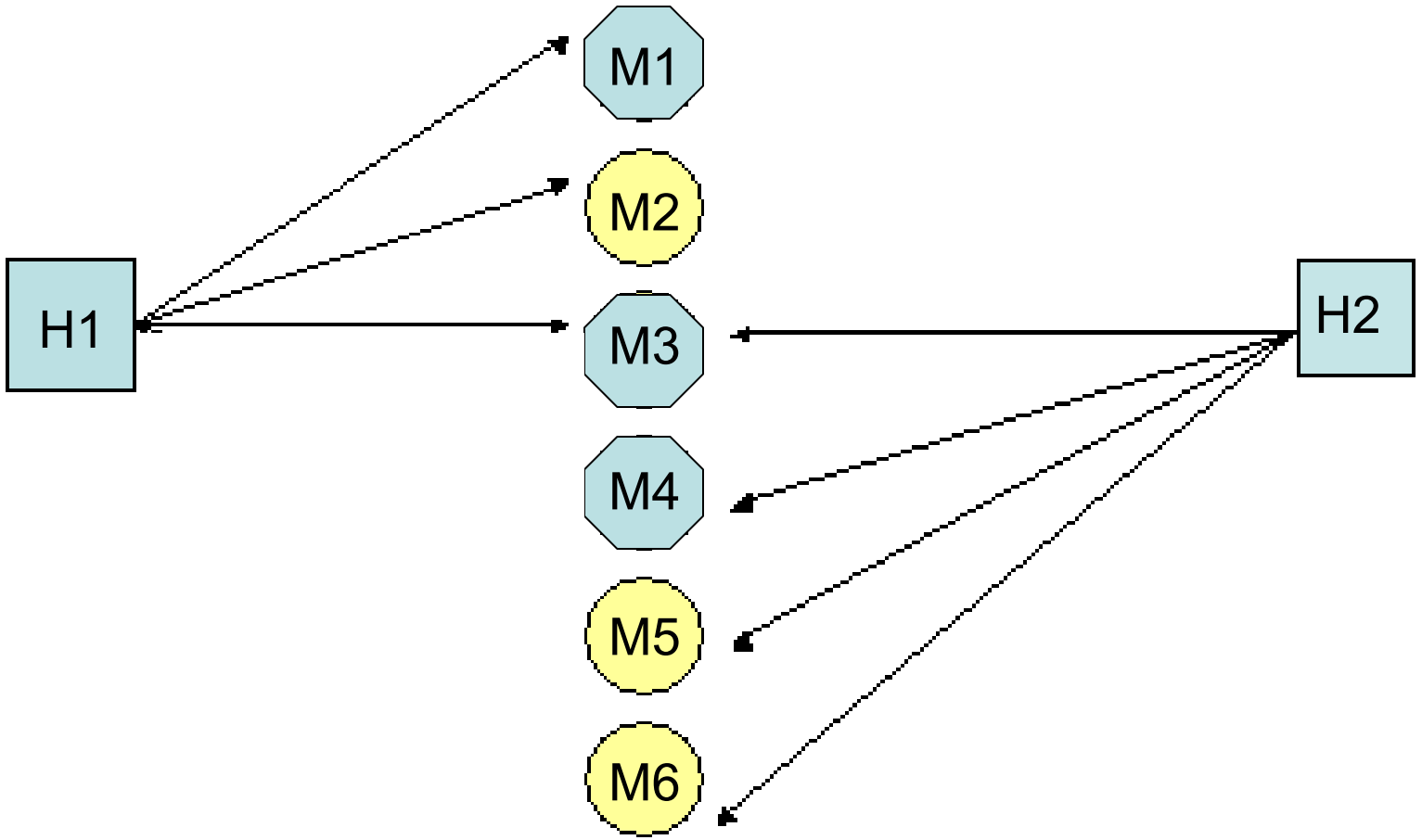
QMR Partitioning



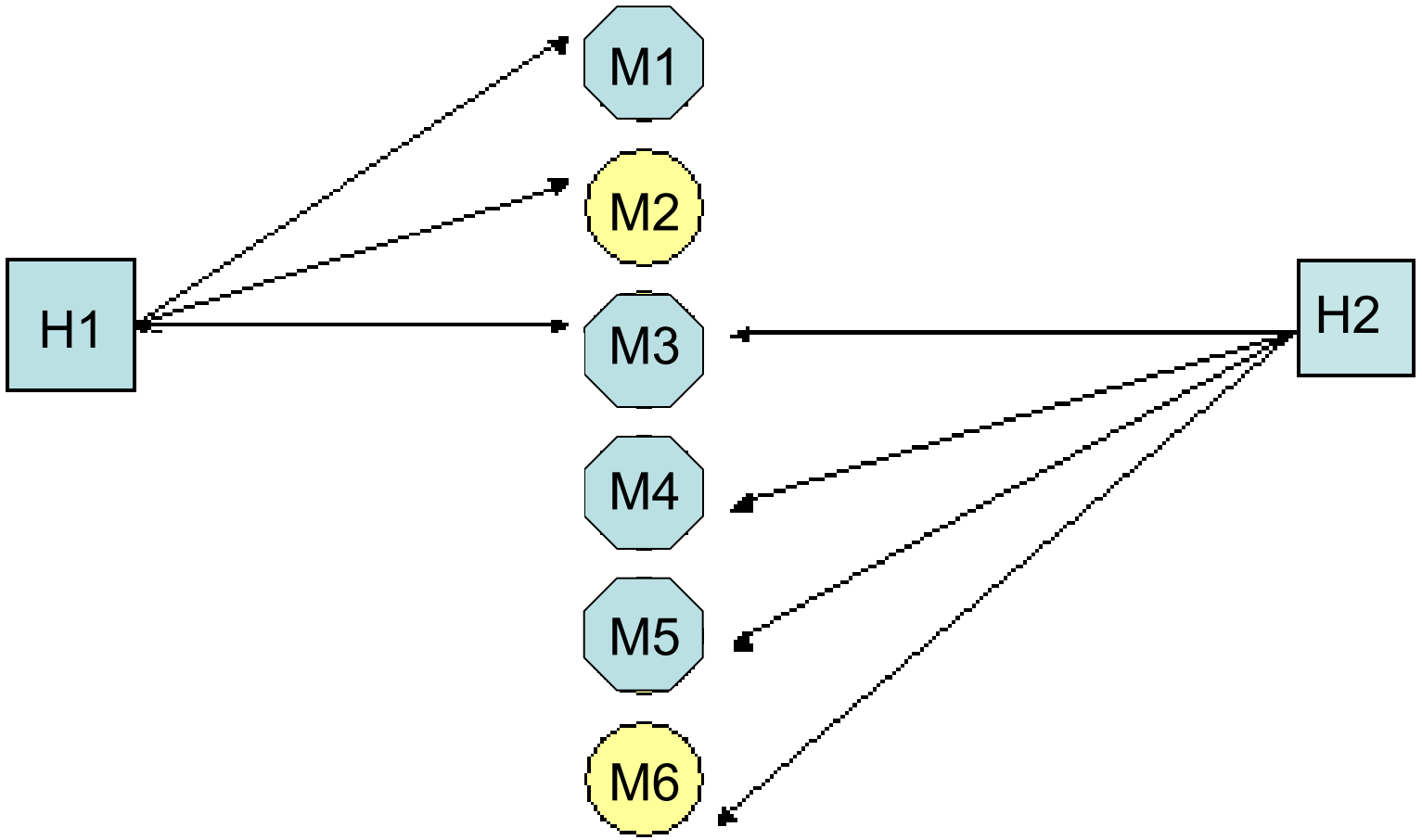
Competitors



Still Competitors



Still Competitors



Multi-Hypothesis Diagnosis

- Set aside complementary hypotheses
- … and manifestations predicted by them
- Solve diagnostic problem among competitors
- Eliminate confirmed hypotheses and manifestations explained by them
- Repeat as long as there are coherent problems among the remaining data

Internist/QMR

- Knowledge Base:
  - 956 hypotheses
  - 4090 manifestations (about 75/hypothesis)
  - *Evocation* like P(H|M)
  - *Frequency* like P(M|H)
  - *Importance* of each M
  - *Causal relations* between H's
- Diagnostic Strategy:
  - Scoring function
  - Partitioning
  - Several questioning strategies

QMR Scoring

- Positive Factors
  - Evoking strength of observed Manifestations
  - Scaled Frequency of causal links from
  - confirmed Hypotheses

- Negative Factors
  - Frequency of predicted but absent
  - Manifestations
  - Importance of unexplained Manifestations

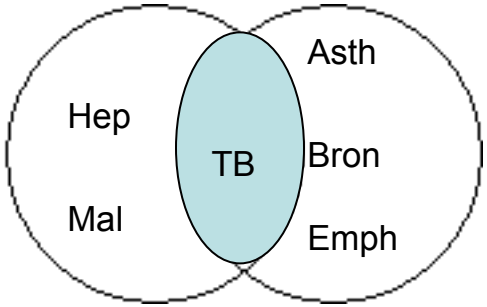
- Various scaling parameters (roughly exponential)

Symptom Clustering for Multi-Disorder Diagnosis

— Tom Wu, Ph.D. 1991

Clustering Alternatives

Symptom	Possible Causes
Fever	TB, Hepatitis, Malaria
Cough	TB, Asthma, Bronchitis, Emphysema

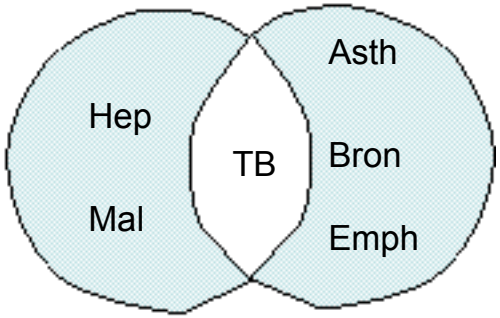


H1

Fever, Cough
TB

H2

Fever	Cough
Hep Mal	Asth Bron Emph



Symptom Clustering is Efficient

Like in any "planning island" approach, reducing an exponential problem to several smaller exponential problems vastly improves efficiency, *if it captures some insight into the problem.*

Wu's algorithm (SYNOPSIS) will keep a compact encoding even if it overgenerates slightly.

> **E.g., suppose that of the set of diseases represented by (d5, d6) x (d3, d7, d8, d9) x (d1, d2, d4), d6 x d8 x d1 is not a candidate. To represent this precisely would require enumerating the 23 valid candidates. Instead, the factored representation is kept.**

In a diagnostic problem drawn from a small subset of the Internist database, it is a *power of 3* faster and a *power of 5* more compact than standard symptom clustering.

Guide search via probabilities, if we have a reasonable model(!)