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

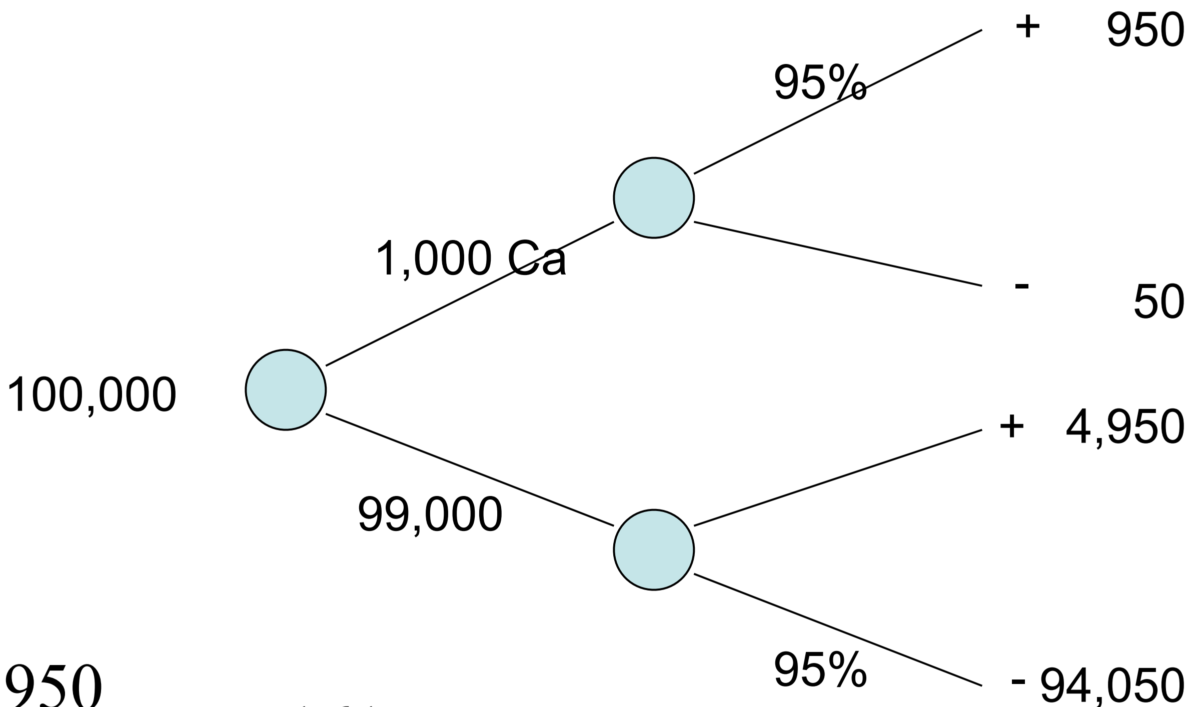
- Could we "re-discover" all of medicine from data? *I think not!*
- Should we focus on methods for reasoning with uncertain data? *Absolutely!*
- But: Feinstein, A. R. (1977). "Clinical Biostatistics XXXIX. The Haze of Bayes, the Aerial Palaces of Decision Analysis, and the Computerized Ouija Board." <u>Clinical Pharmacology and Therapeutics</u> **21**: 482–496.

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:



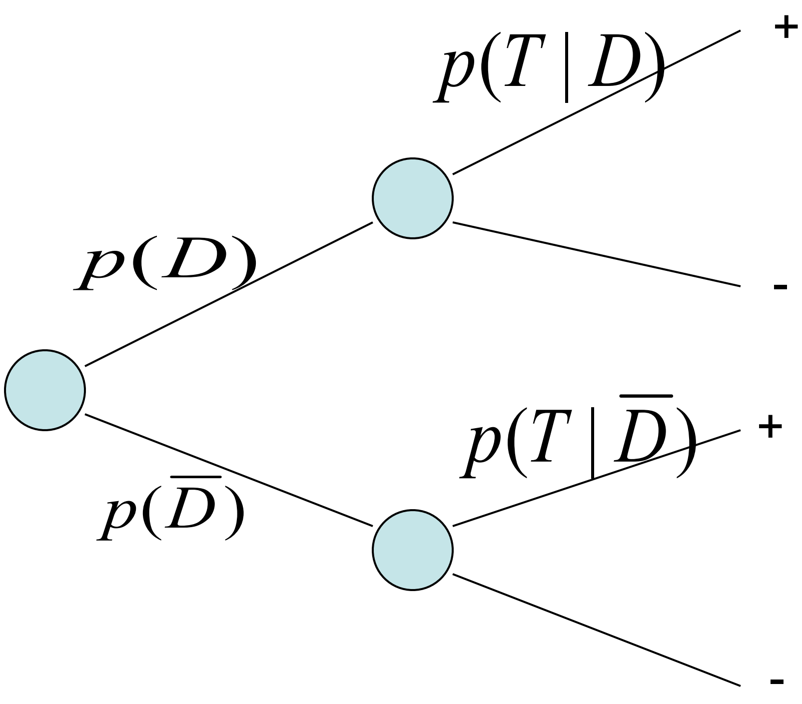
$$\frac{950}{950+4950} = .161$$

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*

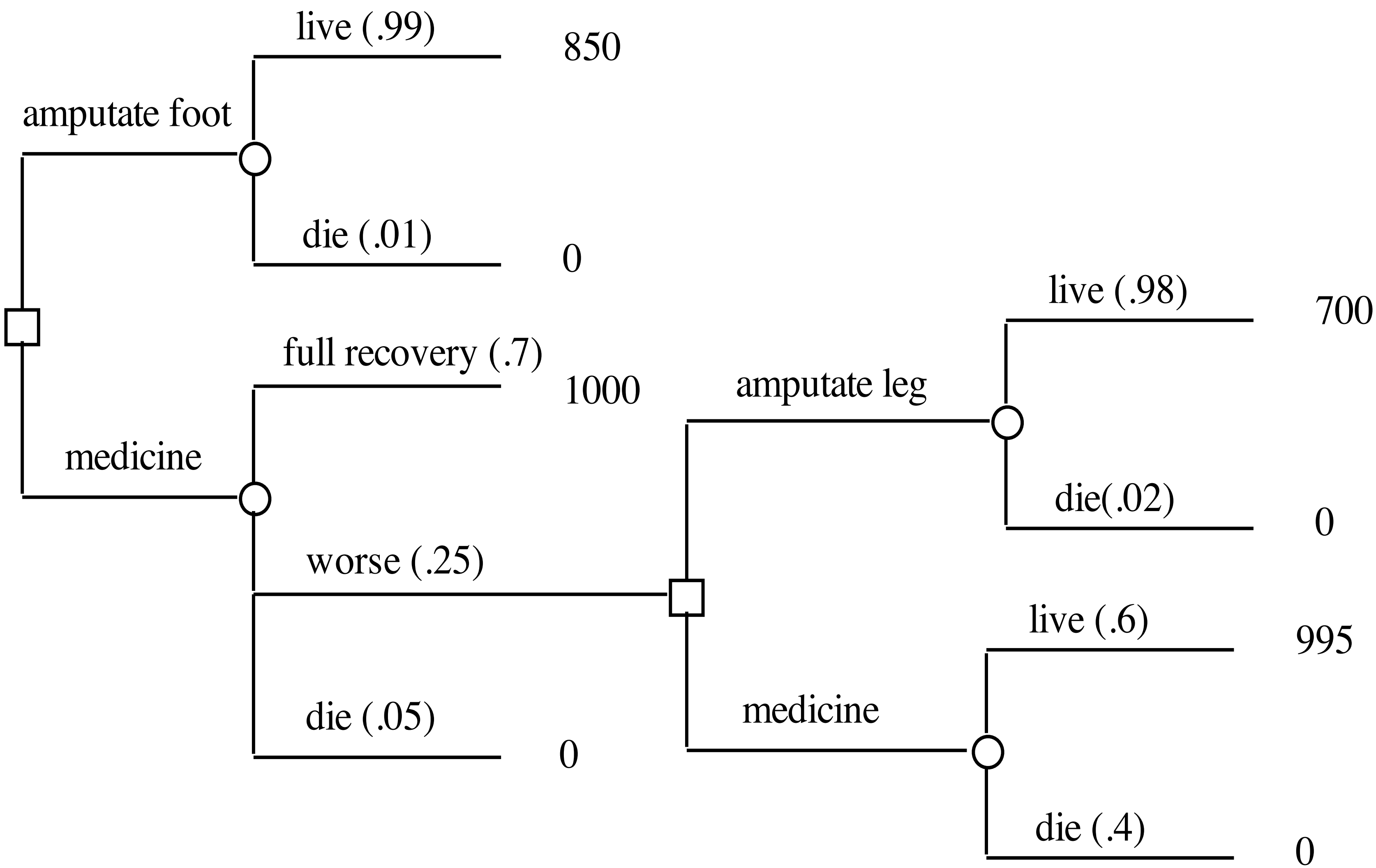
Rationality

- 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

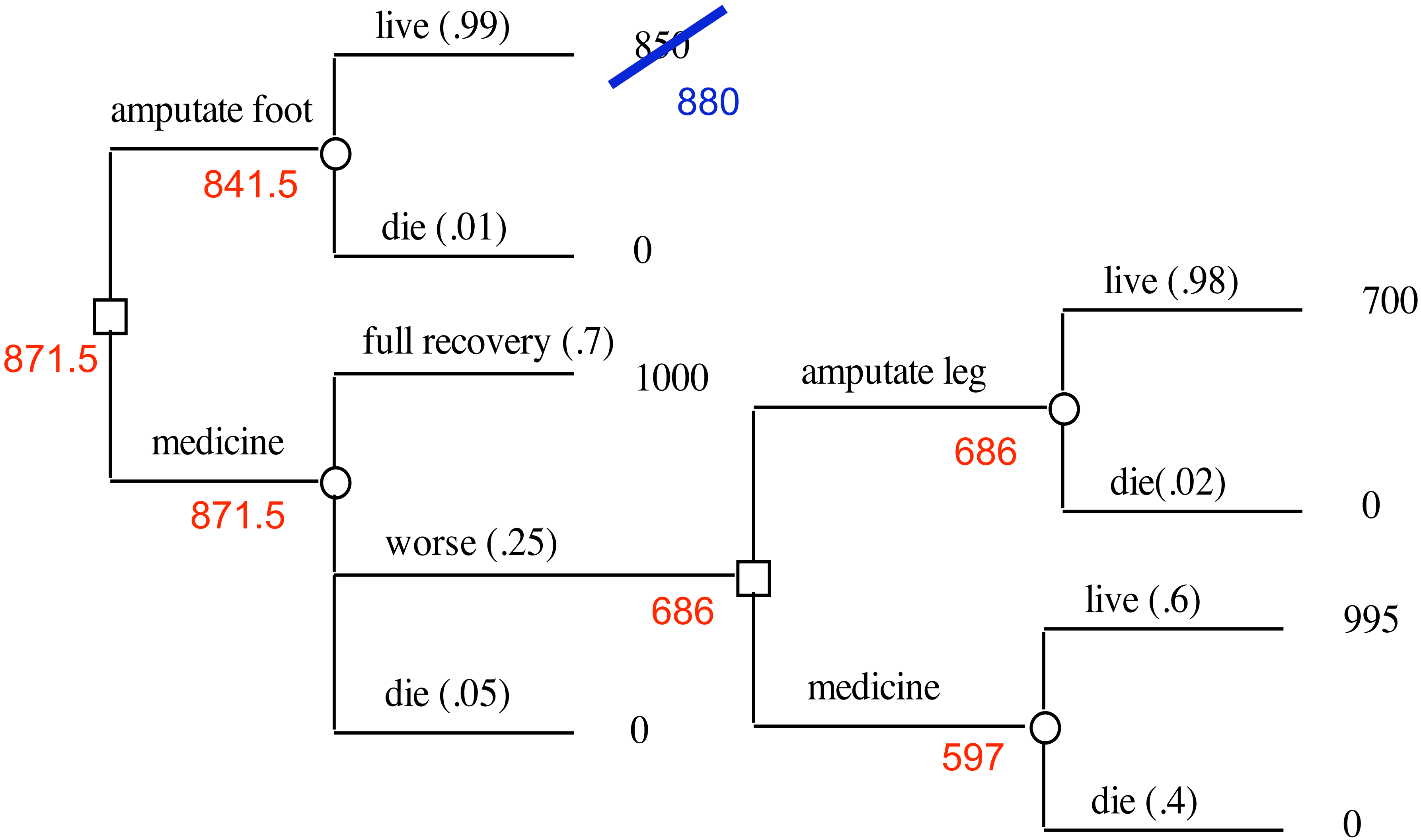
Example: Gangrene

- 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



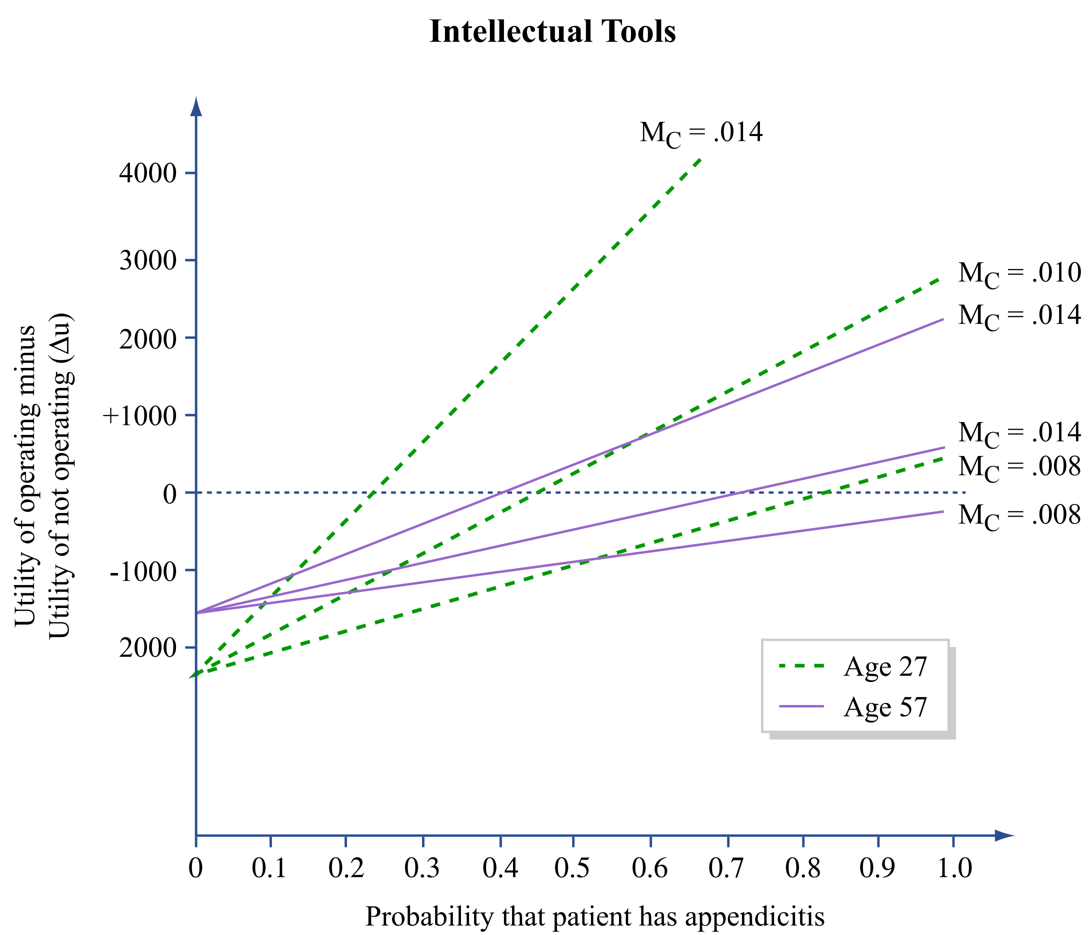
Evaluating the Decision Tree



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.



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

Warner HR, *Computer-Assisted Medical Decision Making,* Acad. Press 1979

Utility Analysis of Appendectomy

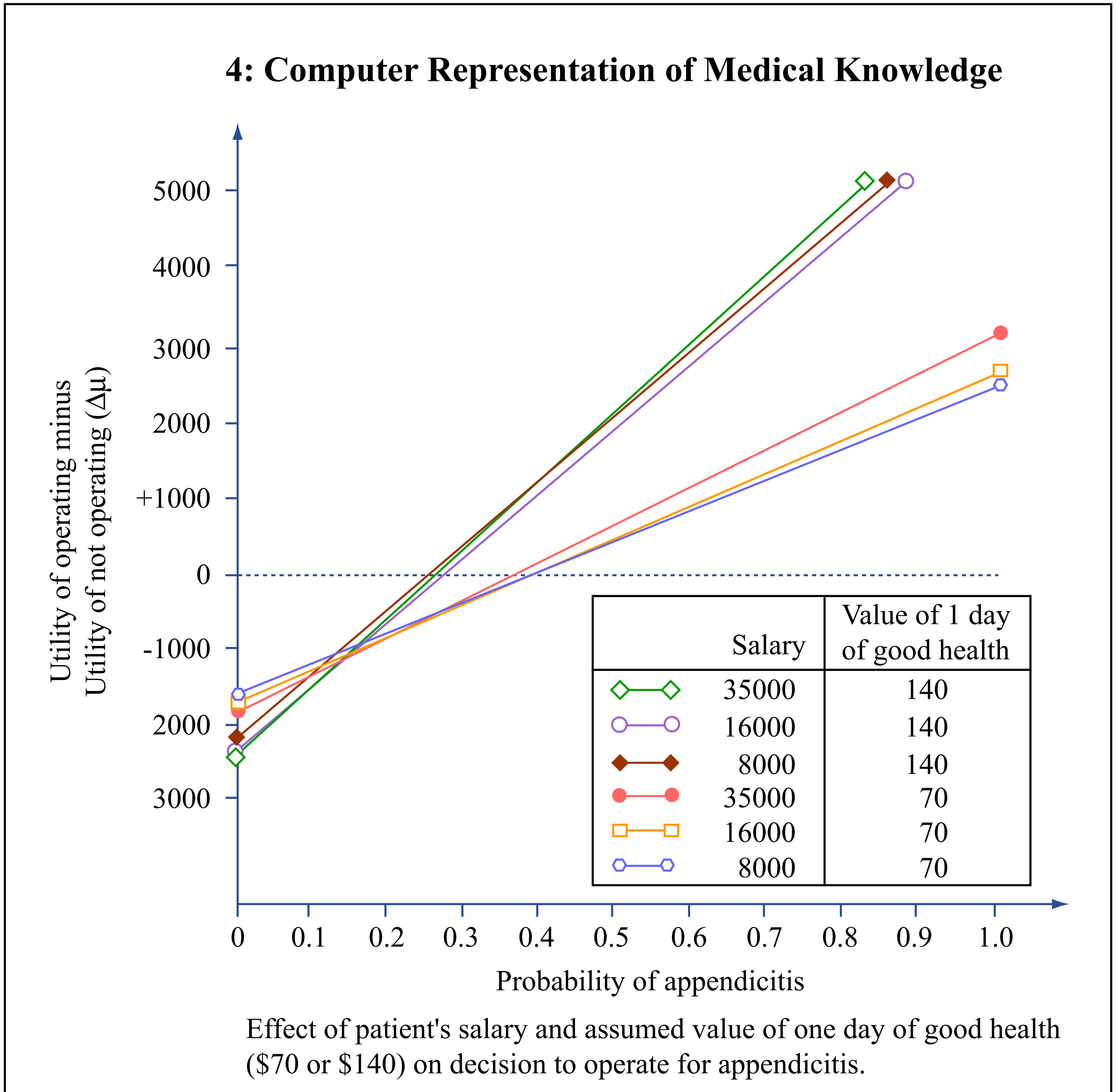


Image by MIT OpenCourseWare. Adapted from Warner, Homer R. "Computer-assisted medical decision making." Academic Press, 1979.

PROB OF APPENDICITIS

A APPENDICITIS BY HISTORY

B REBOUND TENDERNESS IN RLQ

C PRIOR APPENDECTOMY

D IF C THEN EXIT

E WHITE BLOOD COUNT (WBCx100) TH/M3, LAST

F PROB B A 620 90

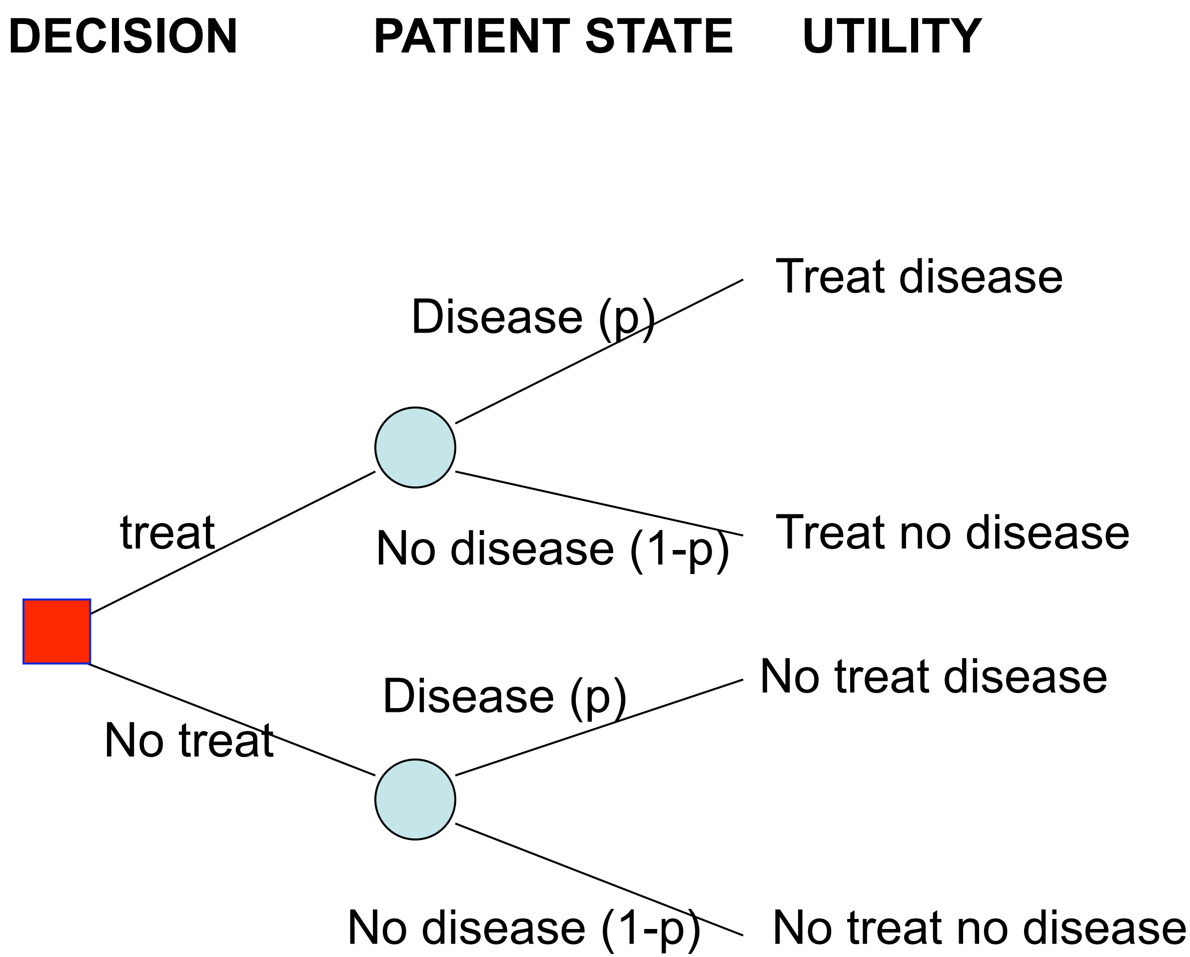
G PROB F 43 18 9, 74 23 7, 93 18 11, 108 10 11, 121 16 13, 134 6 16, 151 5 16, 176 4 14

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!”



Threshold

- Benefit B = U(treat dis) – U (no treat dis)
- Cost C = U(no treat no dis) – U(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.
See Kassirer, Jerome P., and Stephen G. Pauker. "Should Diagnostic Testing be Regulated?" *New England Journal of Medicine* (1978).

Visualizing Thresholds

Figure removed due to copyright restrictions.
See Kassirer, Jerome P., and Stephen G. Pauker. "Should Diagnostic Testing be Regulated?" *New England Journal of Medicine* (1978).

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

- Based on Gorry, *et al.*, AJM 55, 473-484, 1973.
- 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 — Value = ???

Treat with antibiotics

Perform pyelogram

Perform arteriography

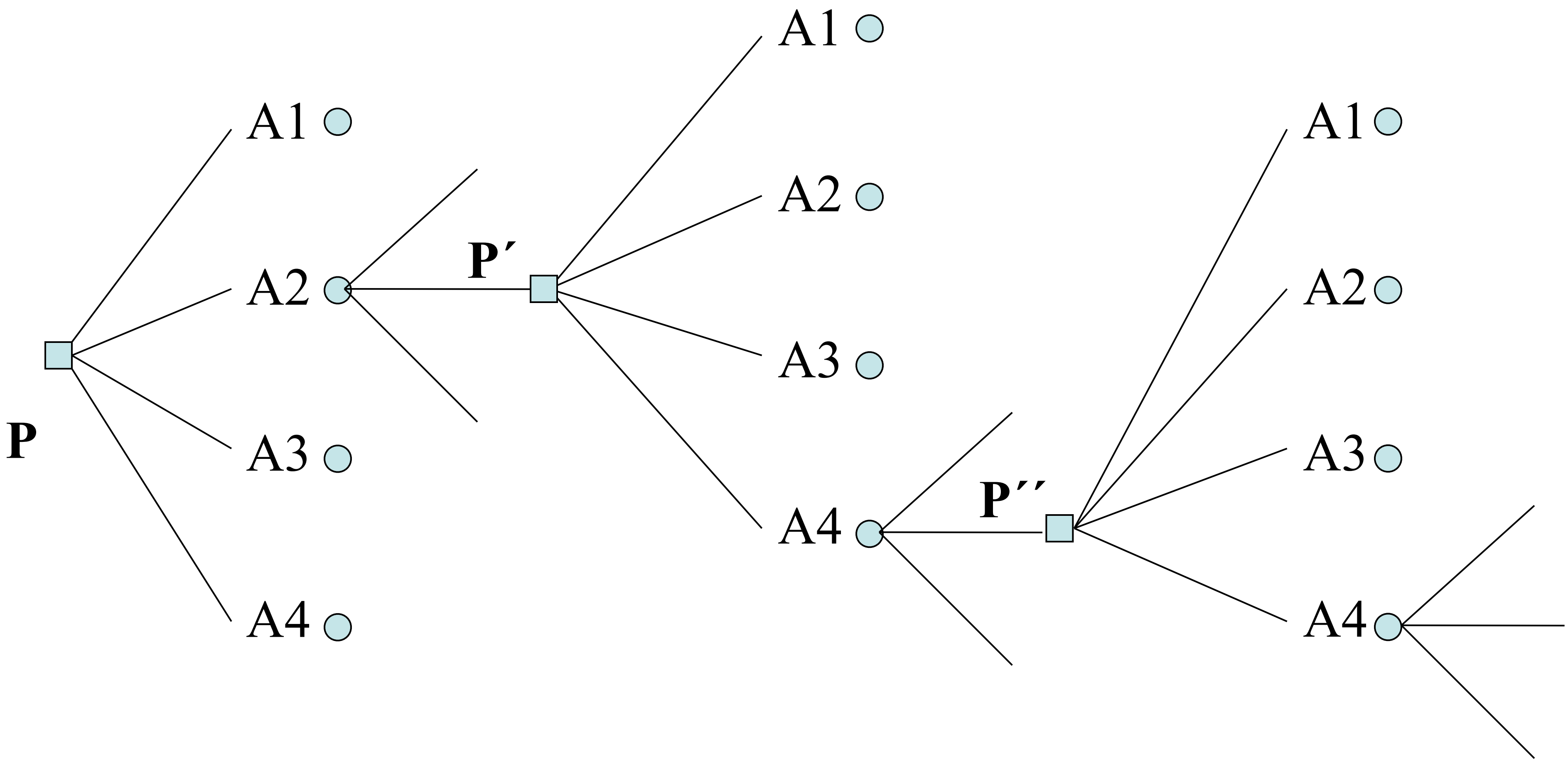
Measure patient's temperature

Determine if there is proteinuria

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



Initial probability distribution

ATN	Acute tubular necrosis	0.250
FARF	Functional acute renal failure	0.400
OBSTR	Urinary tract obstruction	0.100
AGN	Acute glomerulonephritis	0.100
CN	Renal cortical necrosis	0.020
HS	Hepatorenal syndrome	0.005
PYE	Pyelonephritis	0.010
AE	Atheromatous Emboli	0.003
RI	Renal infarction (bilateral)	0.002
RVT	Renal vein thrombosis	0.002
VASC	Renal vasculitis	0.050
SCL	Scleroderma	0.002
CGAE	Chronic glomerulonephritis, acute exacerbation	0.030
MH	Malignant hypertension & nephrosclerosis	0.030

ARF’s Database: P(obs|D)

Conditional probabilities for <u>Proteinuria</u> **Diseases**	**Probabilities** **0**	**Trace to 2+**	**3+ to 4+**
ATN	0.1	0.8	0.1
FARF	0.8	0.2	0.001
OBSTR	0.7	0.3	0.001
AGN	0.01	0.2	0.8
CN	0.01	0.8	0.2
HS	0.8	0.2	0.001
PYE	0.4	0.6	0.001
AE	0.1	0.8	0.1
RI	0.1	0.7	0.2
RVT	0.001	0.1	0.9
VASC	0.01	0.2	0.8
SCL	0.1	0.4	0.5
CGAE	0.001	0.2	0.8
MH	0.001	0.4	0.6

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

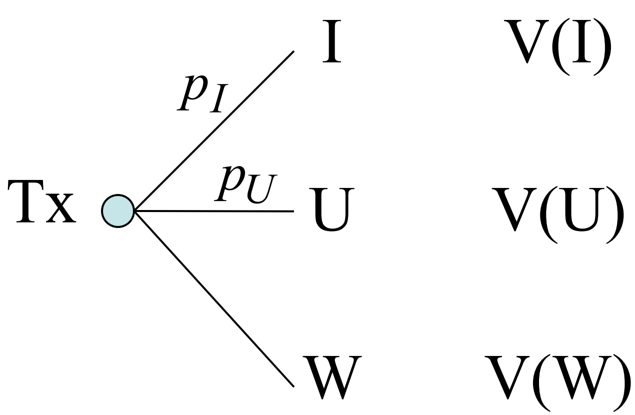
Updating probability distribution

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

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



Modeling treatment

		Steroids	
	improved	*unchanged*	*worse*
atn	0.60	0.20	0.20
farf	0.05	0.35	0.60
obstr	0.05	0.60	0.35
agn	0.40	0.40	0.20
cn	0.05	0.75	0.20
hs	0.05	0.05	0.90
pye	0.05	0.05	0.90
ae	0.05	0.70	0.25
ri	0.01	0.14	0.85
rvt	0.10	0.30	0.60
vasc	0.15	0.25	0.60
scl	0.05	0.05	0.90
cgae	0.40	0.35	0.25
mh	0.05	0.05	0.90

Utilities:

improved: 5000

unchanged: -2500

worse: -5000

Modeling test: transfemoral arteriography

	p(clot)	*cost*
atn	0.01	500
farf	0.01	800
obstr	0.01	500
agn	0.01	500
cn	0.01	500
hs	0.01	800
pye	0.01	500
ae	0.03	800
ri	0.85	500
rvt	0.50	500
vasc	0.01	500
scl	0.01	500
cgae	0.01	500
mh	0.01	500

How large is the tree?

- Infinite, or at least (27+3+8)^(27+3+8), ~10^60
- What can we do?
  - Assume any action is done only once
  - Order:
    - questions
    - tests
    - treatments
- 27! x 4 x 3 x 2 x 8, ~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

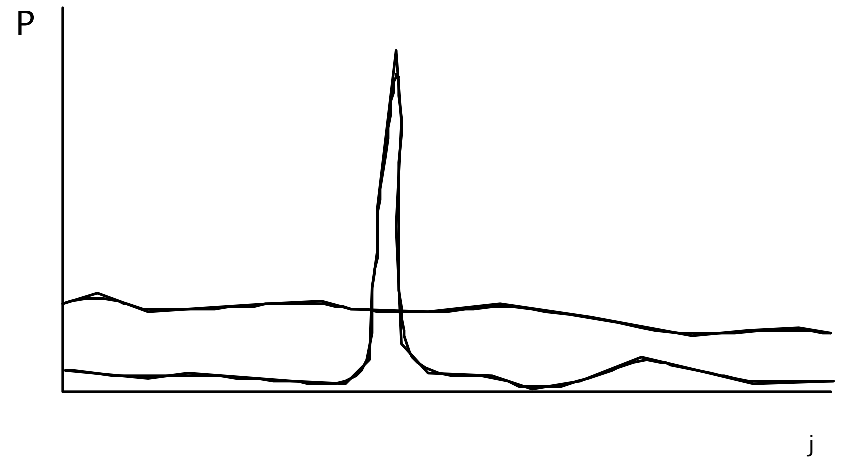
Entropy of a distribution

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

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

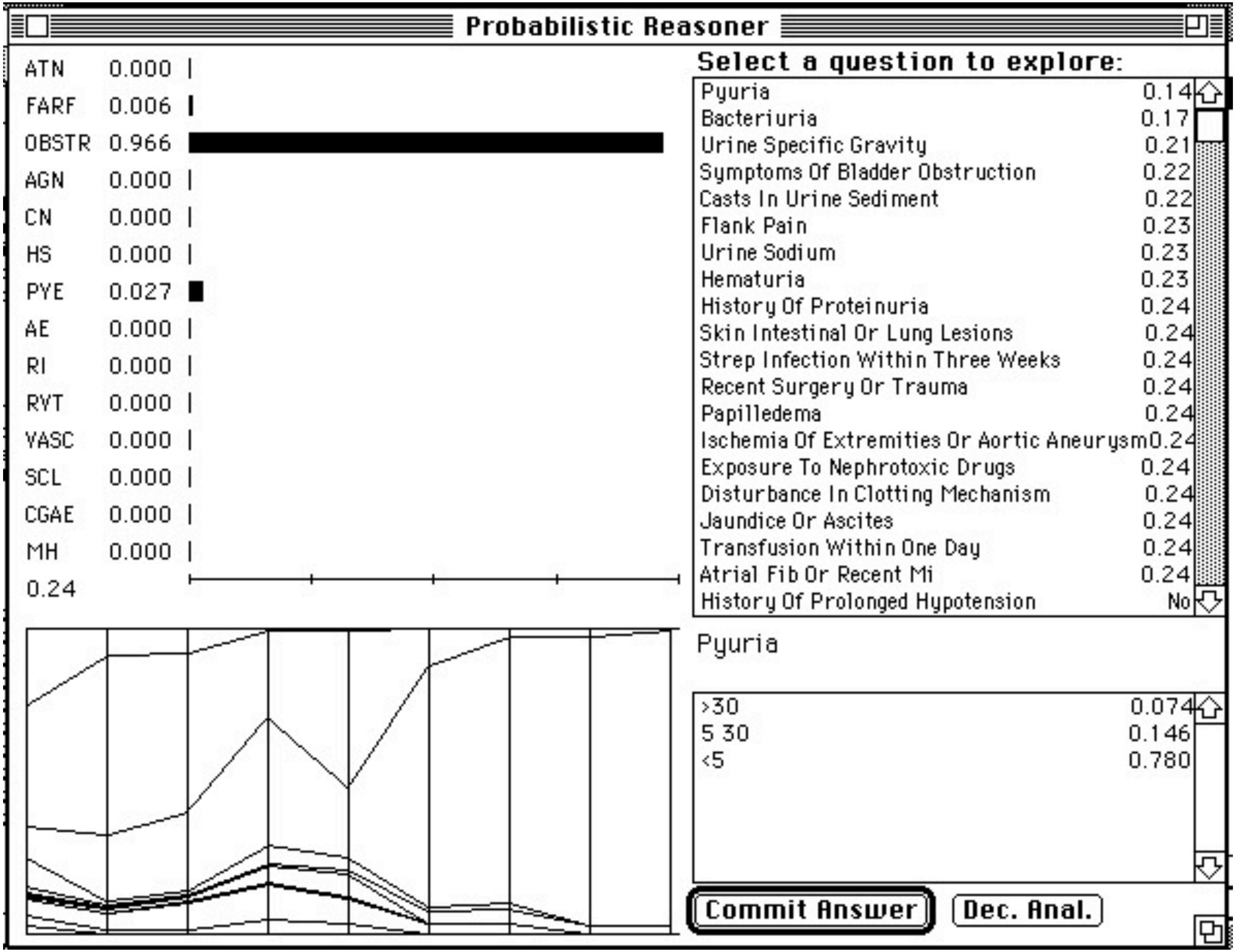
Question 1: What is the patient's age?
1 0-10
2 11-30
3 31-50
4 51-70
5 Over 70
Reply: <u>5</u>

The current distribution is:

Disease	Probability
FARF	0.58
IBSTR	0.22
ATN	0.09

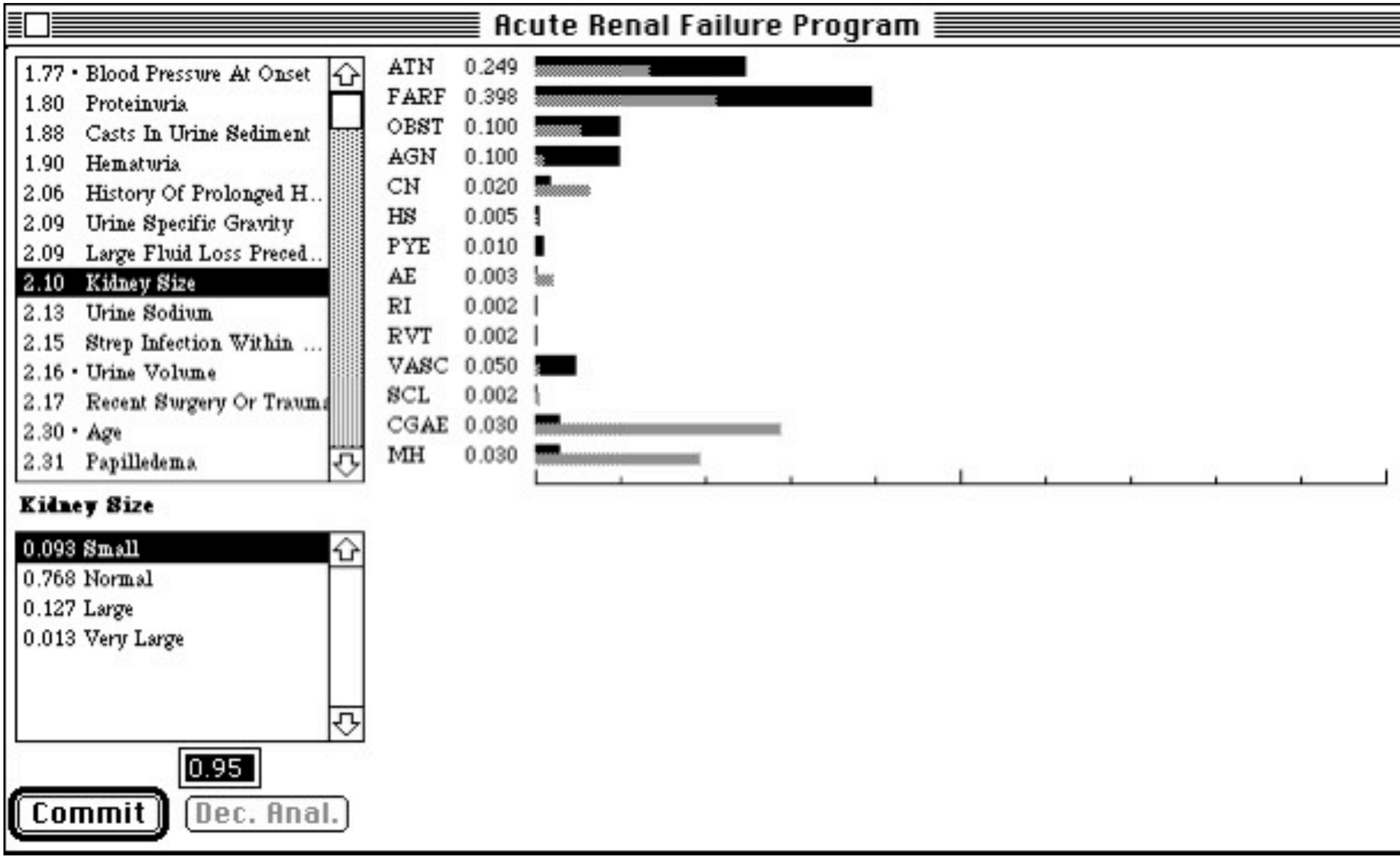
Question 2: What is the patient's sex?
1 Male
2 Pregnant Female
3 Non-pregnant Female
Reply: <u>1</u>
. . .

ARF in 1994

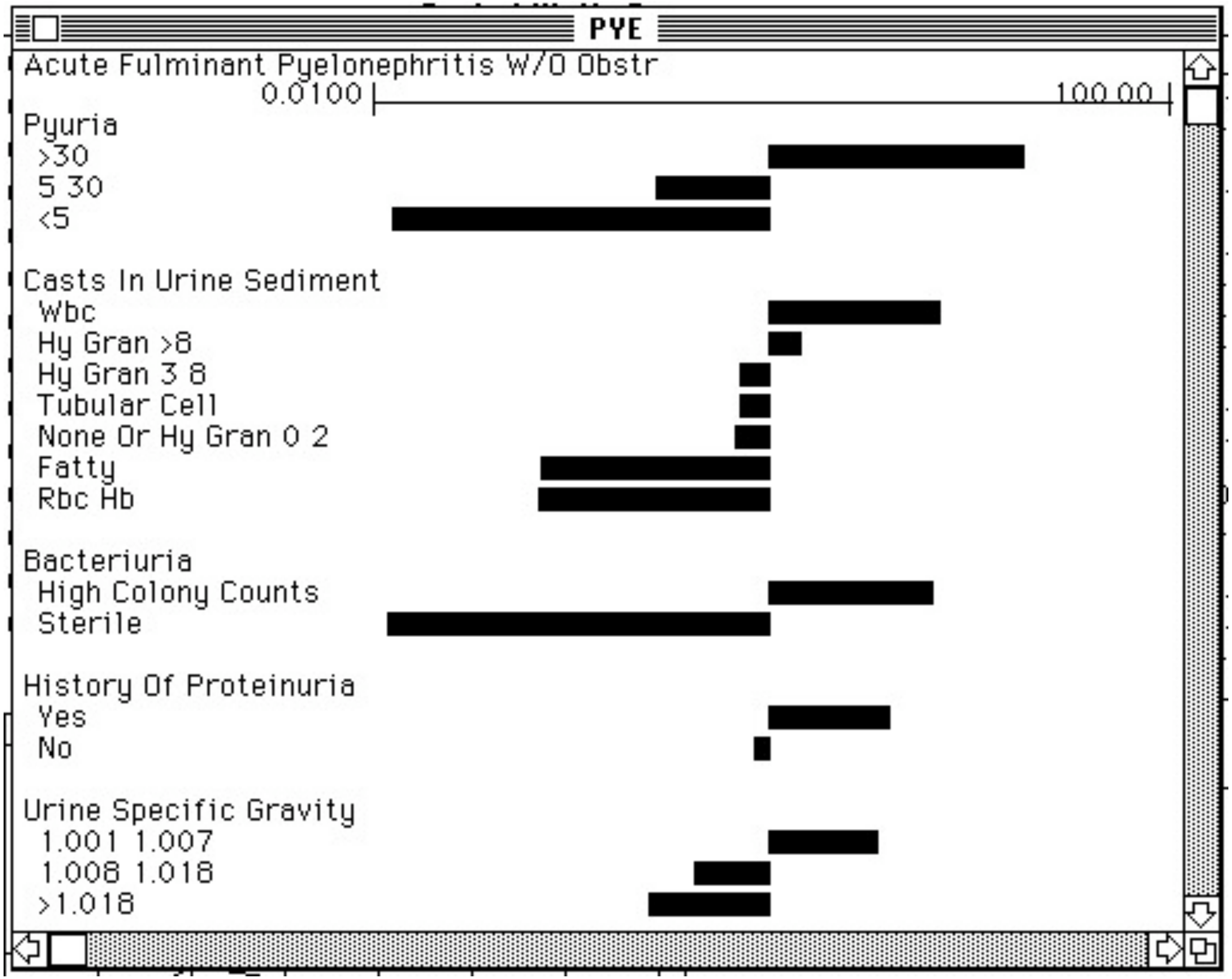


Probabilistic Reasoner
ATN 0.000
FARF 0.006
OBSTR 0.966
AGN 0.000
CN 0.000
HS 0.000
PYE 0.027
AE 0.000
RI 0.000
RVT 0.000
VASC 0.000
SCL 0.000
CGAE 0.000
MH 0.000
0.24
Select a question to explore:
Pyuria 0.14
Bacteriuria 0.17
Urine Specific Gravity 0.21
Symptoms Of Bladder Obstruction 0.22
Casts In Urine Sediment 0.22
Flank Pain 0.23
Urine Sodium 0.23
Hematuria 0.23
History Of Proteinuria 0.24
Skin Intestinal Or Lung Lesions 0.24
Strep Infection Within Three Weeks 0.24
Recent Surgery Or Trauma 0.24
Papilledema 0.24
Ischemia Of Extremities Or Aortic Aneurysm 0.24
Exposure To Nephrotoxic Drugs 0.24
Disturbance In Clotting Mechanism 0.24
Jaundice Or Ascites 0.24
Transfusion Within One Day 0.24
Atrial Fib Or Recent Mi 0.24
History Of Prolonged Hypotension No
Pyuria
>30 0.074
5 30 0.146
<5 0.780
Commit Answer
Dec. Anal.

Local Sensitivity Analysis



Case-specific Likelihood Ratios



Therapy Planning Based on Utilities

Therapy Plan 1

```
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:
  ATN: 0.000;  FARF: 0.006;  OBSTR: 0.966;  AGN: 0.000;  CN: 0.000;  HS: 0.000;  PYE: 0.027;
    AE: 0.000;  RI: 0.000;  RVT: 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-PYELOGRAPHY, 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