Decision Support via Expert Systems

6.872/HST950 Peter Szolovits

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.

Flowchart

BI/Lincoln Labs Clinical Protocols 1978

U.T.I./ VAGINITIS PROTOCOL (12/73)	Unitf: Date:
Chief complaint(a)	Name :
Chief complaint(s)	Birthdate: Phone:
yes no SUBJECTIVE	Provider:
Vaginal discharge, unusual Days duration Vaginal/vulvar Itch/irritation Days duration Pain/burning on urination Inside urethra Outside on a raw area Days duration Unusually frequent urination Days duration Rx for any of above in past 3 mo Age245 Pregnant now Diabetic New pain side/back/belly/pelvis Severe Any blue boxes checked Gyn procedure in past 2 mo Meds inserted into vagina in past few days Any grey boxes checked Incontinence (prior to UTI Sx) Vemiting/too nauseated to eat Fever by Hx in past 48 hrs Chills, teeth chatter Hx of hospitalization for UT prof Kidney X-ray (IVP) Bladder/kidney stones Cystoscopy/in-dwelling catheter High blood pressure Had a UTI before age 12 Past UTI 's25 Actibiotic taken in past 3 weeks OBJECTIVE De urinalysis and culture Bact NBC RBC 23+ protein Any sugar	Any vag dx? Dx non-specific vaginitis Any dx yet? Any greys? Dx urethritis Any reds? Consult MD Mull consult MD for other reasons PLAN (also see back of protocol) Dx of trichomonas? Rx Flagyl Dx of monilia? Rx Mycostatin Dx of non-specific vaginitis? Dx of urethritis Dx of urethritis Dx of urethritis Dy suria so bad pt can hardly urinate
Bact2+ or WBC220? Dx UTI 210 RBC	
A ≥2+ protein	

Codifying Human Knowledge

- Decomposition into "chunks" of knowledge, chaining of inferences
- Matching of case data to prototypical situations
- Using causal models (pathophysiology) to figure out cases

Mycin—Rule-based Systems

- Task: Diagnosis and prescription for bacterial infections of the blood (and later meningitis)
- Method:
 - Collection of modular rules
 - Backward chaining
 - Certainty factors

RULE037

IF the organism 1) stains grampos 2) has coccus shape 3) grows in chains THEN

There is suggestive evidence (.7) that the identity of the organism is streptococcus.

Mycin consult

-----PATIENT-1------

1) Patient's name: FRED SMITH

2) Sex: MALE

3) Age: 55

4) Have you been able to obtain positive cultures from a site at which Fred Smith has an infection? **YES**

-----INFECTION-1------

5) What is the infection? **PRIMARY-BACTEREMIA**

6) Please give the date when signs of INFECTION-1 appeared. 5/5/75

The most recent positive culture associated with the primarybacteremia will be referred to as:

-----CULTURE-1-----

7) From what site was the specimen for CULTURE-1 taken? **BLOOD**

8) Please give the date when this culture was obtained. 5/9/75

The first significant organism from this blood culture will be called:

-----ORGANISM-1------

9) Enter the identity of ORGANISM-1. UNKNOWN

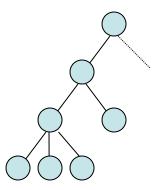
10) Is ORGANISM-1 a rod or coccus (etc.)? ROD

11) The gram stain of ORGANISM-1: GRAMNEG

Davis, et al., Artificial Intelligence 8: 15-45 (1977)

How Mycin Works

- To find out a fact
 - If there are rules that can conclude it, try them
 - Ask the user
- To "run" a rule
 - Try to find out if the facts in the premises are true
 - If they all are, then assert the conclusion(s), with a suitable certainty
- Backward chaining from goal to given facts
- Dynamically traces out behavior of (what might be) a flowchart
 - ▲ Information used everywhere appropriate
 - ▲ Single expression of any piece of knowledge



Explore Mycin's Use of Knowledge

** Did you use RULE 163 to find out anything about ORGANISM-1?

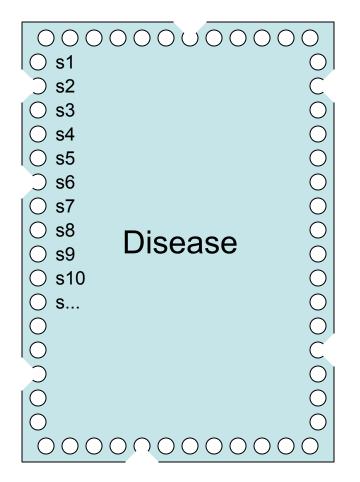
RULE163 was tried in the context of ORGANISM-1, but it failed because it is not true that the patient has had a genito-urinary tract manipulative procedure (clause 3).

** Why didn't you consider streptococcus as a possibility? The following rule could have been used to determine that the identity of ORGANISM-1 was streptococcus: RULE033
But clause 2 ("the morphology of the organism is coccus") was already known to be false for ORGANISM-1, so the rule was never tried.

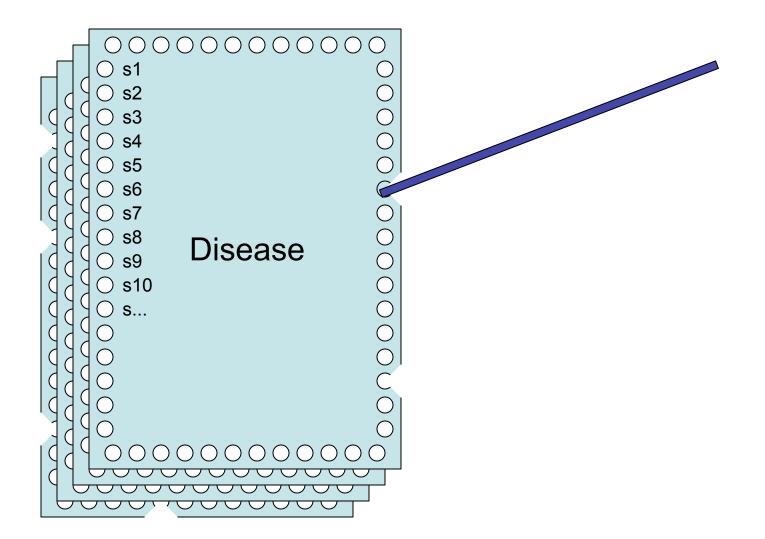
Davis, et al., Artificial Intelligence 8: 15-45 (1977)

Even Simpler Representation

000000000000	
○ s1 ○	
○ s2 ○	
○ s3 ○	
○ s4 ○	
○ s5 ○	
○ s6 ○	
○ s7 ○	
\bigcirc s8 Disease \bigcirc	
○ s10 ○	
○ s ○	
0 0	
0 0	
0 0	
0 0	
0 0	
000000000000	



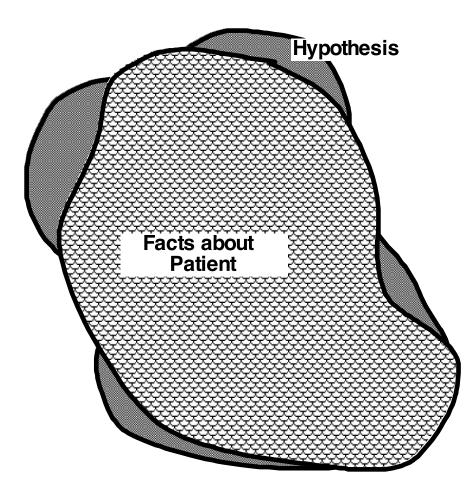
Diagnosis by Card Selection



Diagnosis by Edge-Punched Cards

- Dx is intersection of sets of diseases that may cause all the observed symptoms
- ➤ Difficulties:
 - ➢Uncertainty
 - Multiple diseases
- ~ "Problem-Knowledge Coupler" of Weed

Taking the Present Illness—Diagnosis by Pattern Directed Matching



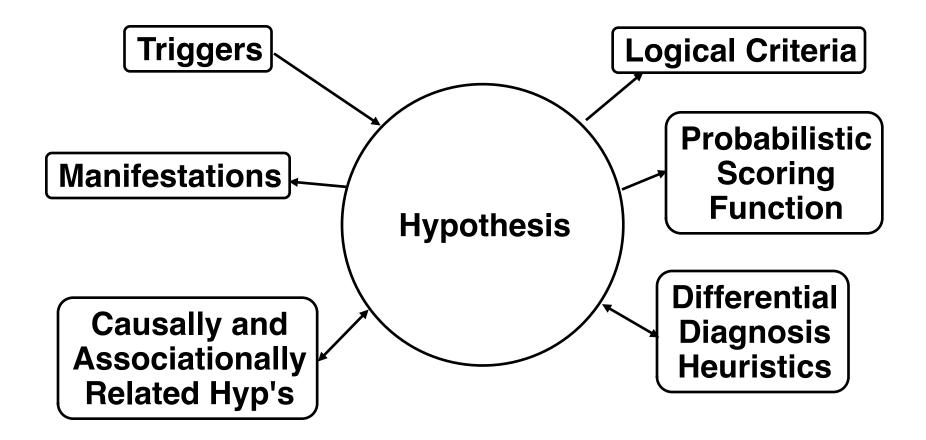
PIP's Theory of Diagnosis

- 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

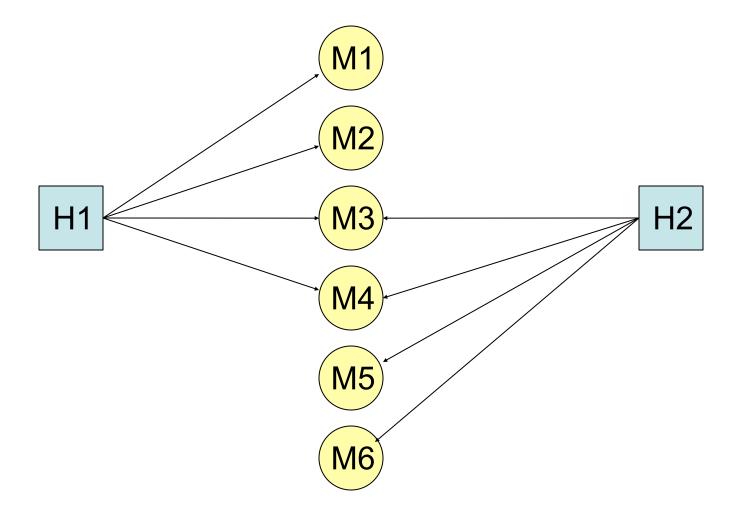
IS-SUFFICIENT: Massive pedal edema & >5 gm/day proteinuria *MUST-NOT-HAVE:* Proteinuria absent *SCORING...*

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 met constrictive pericarditis

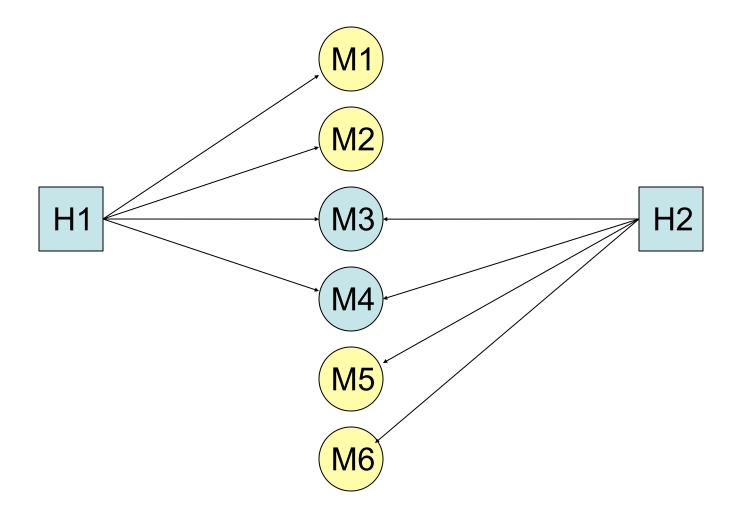
neck veins elevated in constrictive pericarditis ascites present in cirrhosis

pulmonary emboli present i renal vein thrombosis

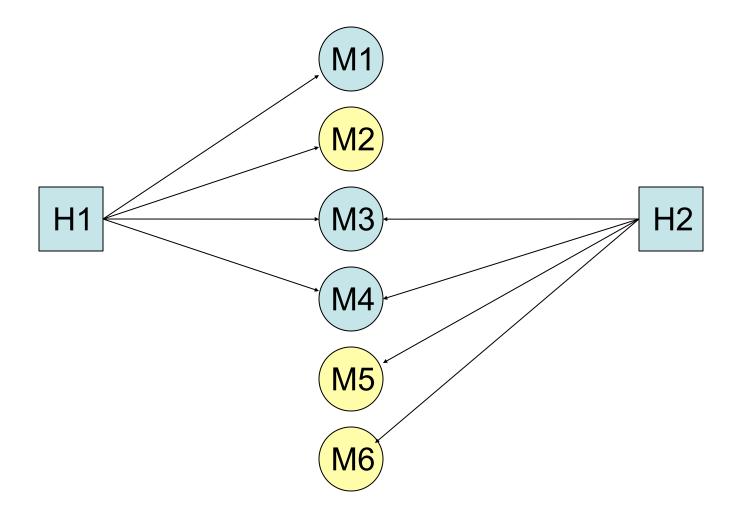
QMR Partitioning



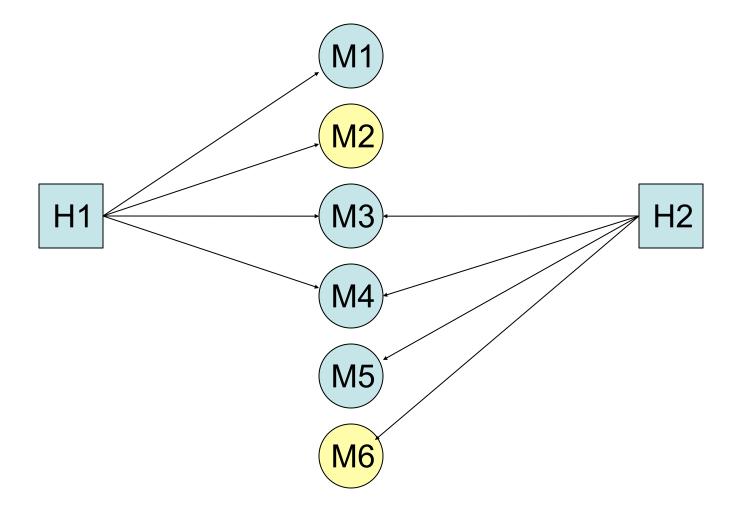
Competitors



Still Competitors



Probably Complementary



Multi-Hypothesis Diagnosis

- Set aside complementary hypotheses
- \succ ... 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 Database

Disease	Finding
ANEMIA DUE TO ABNORMAL MATURATION ANEMIA OF CHRONIC DISEASE ANEMIA OF DECREASED VITAMIN B12 ABSORPTION ANEMIA OF FOLATE DEFICIENCY ANEMIA OF VITAMIN B12 DEFICIENCY ANEMIA SECONDARY TO MARROW DAMAGE ANGINA PECTORIS ANGINA VARIANT <prinzmetal> ANGIOIMMUNOBLASTIC LYMPHADENOPATHY ANKYLOSING SPONDYLITIS</prinzmetal>	
Findings: 1 3 TRIGLYCERIDE <s> SERUM INCREASED 0 2 TACHYCARDIA 0 3 SKIN SWEATING INCREASED GENERALIZED 1 1 SHOULDER PAIN RIGHT 1 1 SHOULDER PAIN LEFT 0 4 SEX MALE 0 2 SEX FEMALE 0 2 SEX FEMALE 1 SHOULDER PAIN LEFT 1 1 HEMORRHAGE GASTROINTESTINAL ACUTE RECEN 1 1 HEMORRHAGE ACUTE RECENT HX</s>	1 2 PEDIATRIC DRUG HYPERSENSITIVITY CHOLESTATIC REF 1 2 PEDIATRIC EXTRAHEPATIC BILIARY ATRESIA 1 2 PEDIATRIC BILIARY CIRRHOSIS SECONDARY 1 2 PEDIATRIC BILIARY CIRRHOSIS PRIMARY 1 2 PEDIATRIC FATTY LIVER SECONDARY 1 2 OBESITY 1 1 WEBER CHRISTIAN DISEASE 1 2 ATHEROMATOUS EMBOLISM 1 4 DIABETIC KETOACIDOSIS 2 3 DIABETES MELLITUS 1 3 GOUTY ARTHRITIS CHRONIC 1 4 GOUTY ARTHRITIS CHRONIC 1 4 GOUTY ARTHRITIS ACUTE 1 3 ABDOMINAL AORTIC ANEURYSM <uncomplicated> 1 3 VENTRICULAR ANEURYSM LEFT 1 3 ARTERIOSCLEROTIC HEART DISEASE 1 3 MYOCARDIAL INFARCTION ACUTE 1 3 CRESCENDO ANGINA</uncomplicated>

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)

Example Case

掌 Internist Data Summary		_ 🗆 🗡
Internist Reconstruction Data Manifestations PRESENT:	Summary	Diagnose
ABDOMEN DISTENTION ABDOMEN FLUID WAVE AGE GTR THAN 55 ALKALINE PHOSPHATASE BLOOD GTR THAN 2 TIMES NORMAL AMMONIA BLOOD INCREASED ANOREXIA ARTHRITIS HX ASCITIC FLUID PROTEIN 3 GRAM <s> PER DL OR LESS ASCITIC FLUID WBC 100 TO 500 ASTERIXIS BILIRUBIN BLOOD CONJUGATED INCREASED BILIRUBIN BLOOD CONJUGATED INCREASED BILIRUBIN URINE PRESENT CHEST PAIN LATERAL EXACERBATION WITH BREATHING CHEST PAIN LATERAL SHARP</s>	L	1
DEPRESSION HX DYSPNEA ABRUPT ONSET		-
	Remove Present	
Manifestations ABSENT:		
ALCOHOLISM CHRONIC HX ASCITIC FLUID AMYLASE INCREASED ASCITIC FLUID CYTOLOGY POSITIVE ASCITIC FLUID LDH GTR THAN 500 DIARRHEA CHRONIC ESOPHAGUS BARIUM MEAL VARICES FECES BLACK TARRY FEVER HEMATOCRIT BLOOD LESS THAN 35 PRESSURE VENOUS CERVICAL INCREASED ON INSPECTION STOMACH BARIUM MEAL ULCER CRATER <s> T3 RESIN UPTAKE INCREASED T4 FREE BLOOD INCREASED UREA NITROGEN BLOOD 30 TO 59 URIC ACID BLOOD INCREASED</s>		*
	Remove Absent	

Initial Solution

Diagnostic Results	
voblem: -94 HEPATITIS CHRONIC ACTIVE -119 PEDIATRIC HEPATITIS CHRONIC ACTIVE -136 MACRONODAL CIRRHOSIS <postnecrotic> -158 BILIARY CIRRHOSIS PRIMARY -178 PEDIATRIC BILIARY CIRRHOSIS PRIMARY</postnecrotic>	Complementary: -143 MICRONODAL CIRRHOSIS <laennecs> -162 HEPATITIS ACUTE VIRAL -170 CHOLANGIOCARCINOMA <intrahepatic hilar="" non=""> -178 HEPATIC AMYLOIDOSIS</intrahepatic></laennecs>
	Shelf: ABDOMEN DISTENTION ARTHRITIS HX
xplained: AGE GTR THAN 55 ALKALINE PHOSPHATASE BLOOD GTR THAN 2 TIMES NORMAL ANOREXIA BILIRUBIN BLOOD CONJUGATED INCREASED BILIRUBIN URINE PRESENT FECES LIGHT COLORED HAND <s> PALMAR ERYTHEMA</s>	CHEST PAIN LATERAL EXACERBATION WITH BREATHING CHEST PAIN LATERAL SHARP FECES GUAIAC TEST POSITIVE PLEURAL FRICTION RUB WEIGHT INCREASE RECENT HX
IMMUNOELECTROPHORESIS SERUM IGA INCREASED IMMUNOELECTROPHORESIS SERUM IGG INCREASED bsent:	Askable:
DIARRHEA CHRONIC FEVER HEMATOCRIT BLOOD LESS THAN 35	ABDOMEN PAIN CHRONIC ABDOMEN PAIN EPIGASTRIUM ABDOMEN PAIN EPIGASTRIUM UNRELIEVED BY ANTACID ABDOMEN PAIN EXACERBATION WITH MEAL <s> ABDOMEN PAIN NON COLICKY ABDOMEN PAIN DEPERTY</s>
nexplained:	ABDOMEN PAIN PRESENT ABDOMEN PAIN RIGHT UPPER QUADRANT
ABDOMEN DISTENTION ABDOMEN FLUID WAVE AMMONIA BLOOD INCREASED ARTHRITIS HX ASCITIC FLUID PROTEIN 3 GRAM <5> PER DL OR LESS ASCITIC FLUID WBC 100 TO 500	ABDOMEN TENDERNESS PRESENT ABDOMEN TENDERNESS RIGHT UPPER QUADRANT ACTIVATED PARTIAL THROMBOPLASTIN TIME INCREASED AGE 16 TO 25 AGE 26 TO 55 ALBUMIN SERUM DECREASED AI KALINE PHOSPHATASE BOOD INCREASED NOT OVER 2 TIMES NORMA

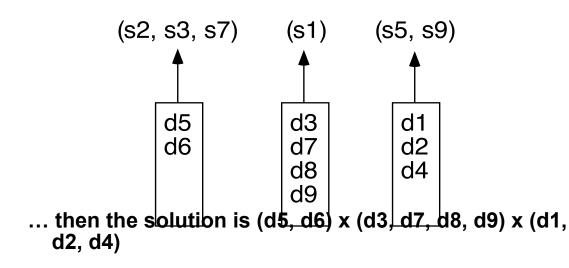
Symptom Clustering for Multi-Disorder Diagnosis

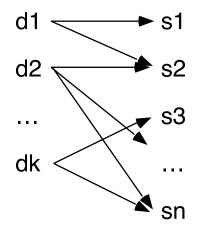
— Tom Wu, Ph.D. 1991

Assume a bipartite graph representation of diseases/ symptoms

Given a set of symptoms, how to proceed?

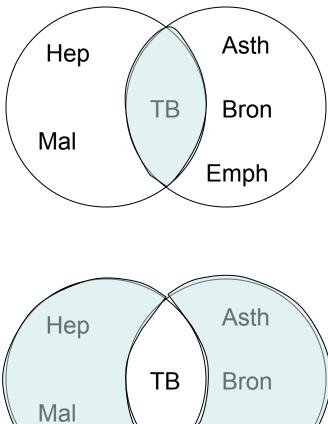
If we could "guess" an appropriate clustering of the symptoms so that each cluster has a single cause ...





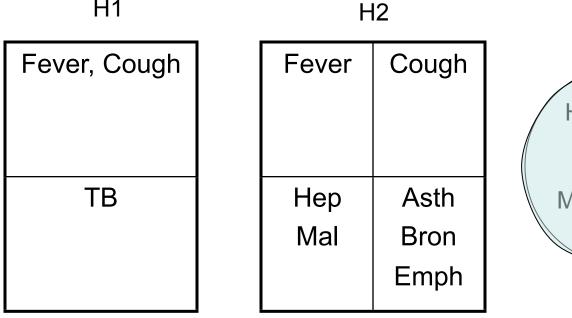
Clustering Alternatives

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



Emph

H1

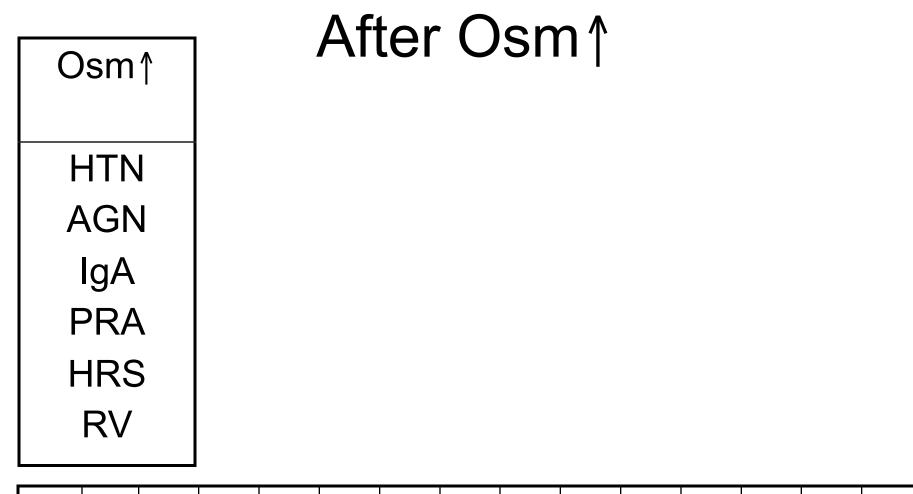


Synopsis in Renal Disease

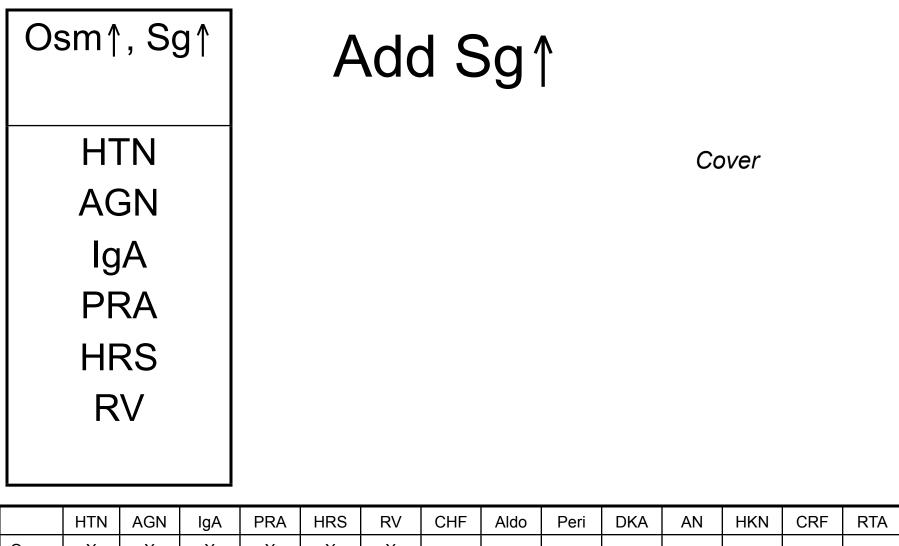
- Diseases
 - Hypertension (HTN)
 - Acute glomerulonephritis (AGN)
 - IgA nephropathy (IgA)
 - Prerenal azotemia (PRA)
 - Hepatorenal syndrome (HRS)
 - Renal vasculitis (RV)
 - Congestive heart failure (CHF)
 - Aldosteronism (Aldo)
 - Constrictive pericarditis (Peri)
 - Diabetic ketoacidosis (DKA)
 - Analgesic nephropathy (AN)
 - Hypokalemic nephropathy (HKN)
 - Chronic renal failure (CRF)

- Symptoms
 - High urine osmolality (Osm↑)
 - High urine specific gravity (Sg↑)
 - Low urine sodium (Na↓)
 - Low urine pH (pH↓)

	HTN	AGN	lgA	PRA	HRS	RV	CHF	Aldo	Peri	DKA	AN	HKN	CRF	RTA
Osm	X	Х	Х	Х	Х	Х								
Sg↑	Х	Х	Х	Х	Х	Х	Х							
Na↓				X	Х		Х	Х	Х					
pH↓		Х		Х						Х	Х	Х	Х	Х



	HTN	AGN	lgA	PRA	HRS	RV	CHF	Aldo	Peri	DKA	AN	HKN	CRF	RTA
Osm	X	Х	Х	Х	Х	Х								
Sg↑	Х	Х	Х	Х	Х	Х	Х							
Na↓				Х	Х		Х	Х	Х					
pH↓		Х		Х						Х	Х	Х	Х	Х

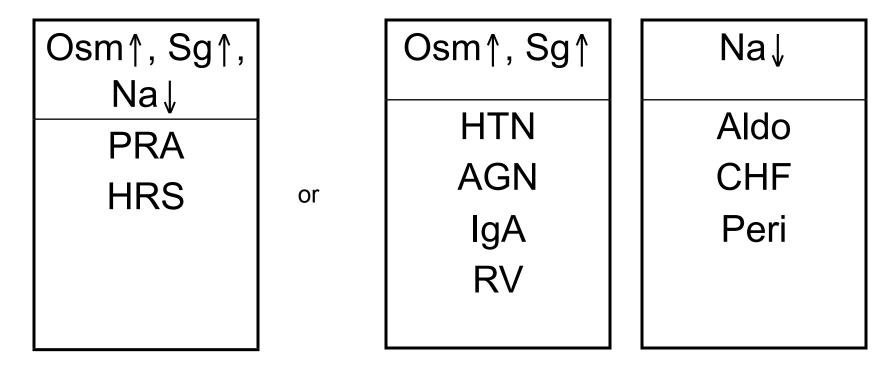


	HTN	AGN	lgA	PRA	HRS	RV	CHF	Aldo	Peri	DKA	AN	HKN	CRF	RTA
Osm	Х	Х	Х	Х	Х	Х								
Sg↑	Х	Х	Х	Х	Х	Х	X							
Na↓				Х	Х		X	Х	Х					
pH↓		Х		Х						Х	Х	Х	Х	Х

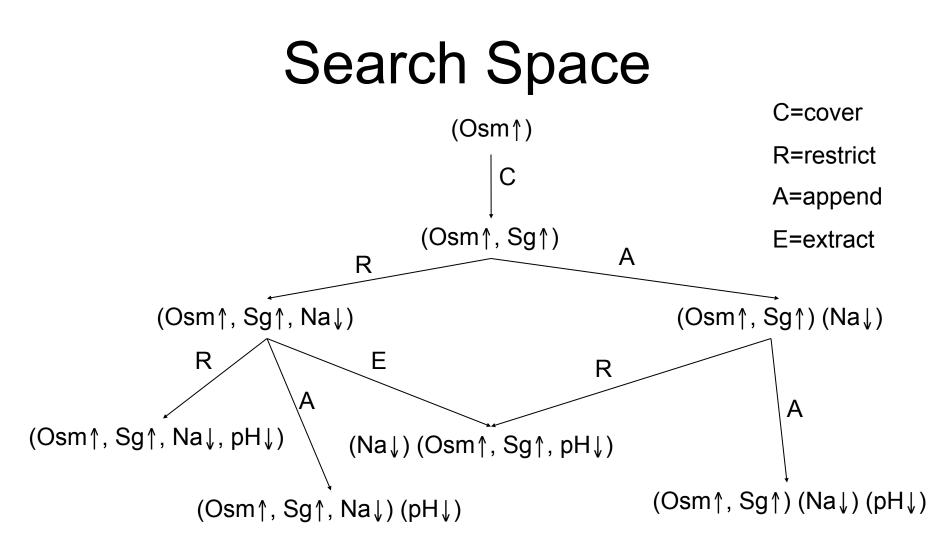
Add Na↓

Restrict

Append



	HTN	AGN	lgA	PRA	HRS	RV	CHF	Aldo	Peri	DKA	AN	HKN	CRF	RTA
Osm	Х	Х	Х	Х	Х	Х								
Sg↑	Х	Х	Х	Х	Х	Х	Х							
Na↓				Х	Х		Х	Х	Х					
pH↓		Х		Х						Х	Х	Х	Х	Х



	HTN	AGN	lgA	PRA	HRS	RV	CHF	Aldo	Peri	DKA	AN	HKN	CRF	RTA
Osm	Х	Х	Х	Х	Х	Х								
Sg↑	Х	Х	Х	Х	Х	Х	X							
Na↓				Х	Х		Х	Х	Х					
pH↓		Х		Х						Х	Х	Х	Х	Х

Symptom Clustering is Efficient

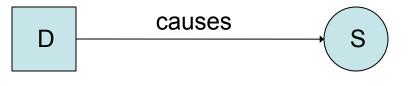
- Like in any "planning island" approach, reducing an exponential problem to several smaller exponential problems vastly improves efficiency, *if it captures some insight into the problem.*
- Wu's algorithm (SYNOPSIS) will keep a compact encoding even if it overgenerates slightly.
 - E.g., suppose that 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(!)

More Expert Systems

- Causality?
- What's in a Link?
- Temporal reasoning
- Quantitative reasoning
- Model-based reasoning
- Workflow

Meaning of Representation?



- Always? →probability
- Magnitude? \rightarrow severity; bad cold \rightarrow worse fever?
- Delay? →temporality
- Where? → spatial dependency
- Under what conditions? → context
- Interaction of multiple causes \rightarrow physical laws
- Cross-terms → high-dimensional descriptions

Temporal Reasoning

Keeping track of multiple forms of temporal relations (Kahn '75)
The time line

➤ "On Dec. 12 last year . . ."

> Special reference events

"Three days after I was hospitalized in 1965 . . ."

Temporal Ordering Chains

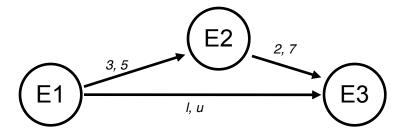
> "It must have been before I graduated from high school."

Constraint propagation (Kohane '87)

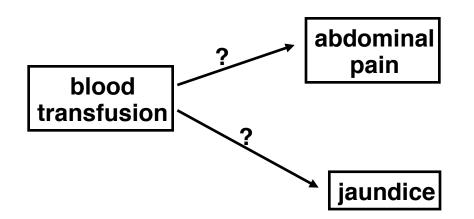
Primitive relation: e1, e2, lower, upper bounds

Heuristics for propagation based on semantic grouping

 $3 \le T(E2)-T(E1) \le 5$ $2 \le T(E3)-T(E2) \le 7$ Therefore $l=5 \le T(E3)-T(E1) \le 12=u$

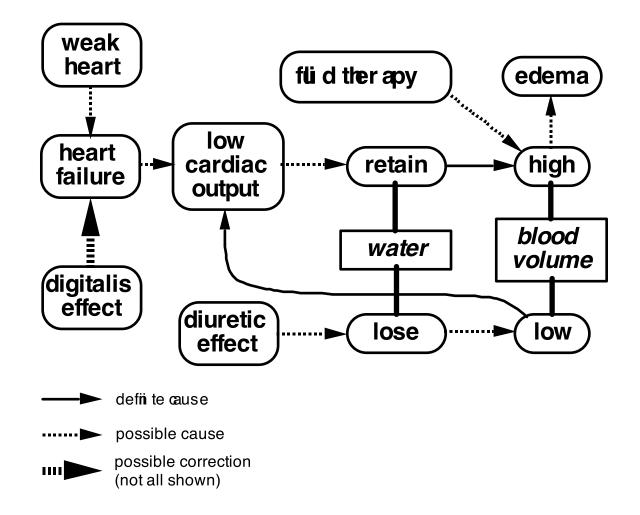


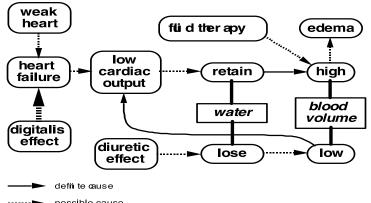
Exploiting Temporal Relations



- transfusion precedes both abdominal pain and jaundice implies transfusion-borne acute hepatitis B
- \succ as in 1, but only by one day
- jaundice occurred 20 years ago, transfusion and pain recent
- Can be very efficient at filtering out nonsense hypotheses.

Interpreting the Past with a Causal/Temporal Model



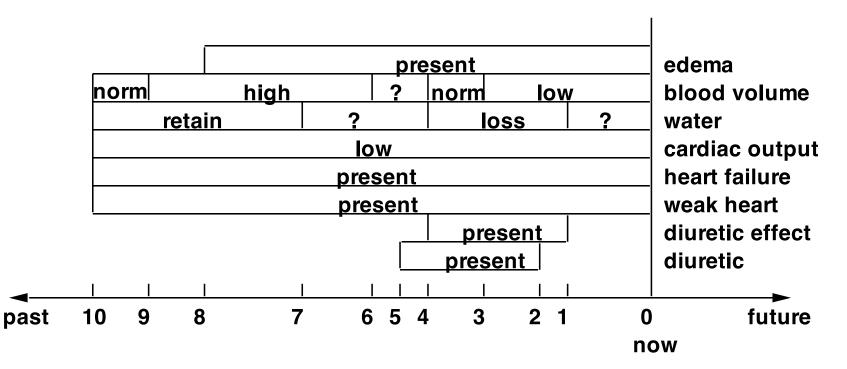


possible cause

(not all shown)

Postdiction

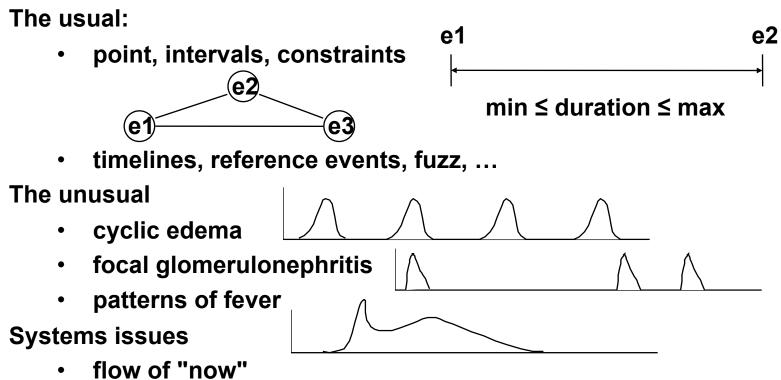
Long, Reasoning about State from Causation and Time in a Medical Domain, *AAAI 83*



Temporal Representation can be Complex

Figure of EKG data from heart removed due to copyright restrictions.

Time



- supporting the illusion of "instantaneous" decision-making within a temporal reasoner
 - correcting the past
 - reasoning by hindsight

The Surprisingly Normal pH

- Diarrhea causes bicarbonate (alkali) loss
- Vomiting causes acid loss
- Therefore, normal pH is a manifestation of {diarrhea + vomiting}!

Multi-Level Causal Model

Figure of three-level causal model removed due to copyright restrictions.

Reasoning from Models

- Model handles all possible interactions, without having explicitly to anticipate them all
- Reasoning: Fit parameters to a physiological model, then predict consequences to suggest
 - ➤ other expected findings
 - reasonable interventions
- Qualitative models
- Combining associational and model-based reasoning

Guyton's Model of Cardiovascular Dynamics

Figure removed due to copyright restrictions.

Long's Clinical Model of Heart Failure Predictions for Mitral Stenosis with Exercise

Figure removed due to copyright restrictions.

Physiological

"All variations in myocardial contractile activity can be expressed as displacements of the force-velocity curve. However, there are two fundamental ways in which the force-velocity curve can be shifted. Figure {left} shows a family of force-velocity curves obtained from an isolated cardiac muscle; each curve was obtained at a different preload, i.e., with a different degree of stretch on the muscle. Note that changing the preload has altered the intercept of the force-velocity curve on the horizontal axis; i.e., it has increased the isometric force developed by the muscle. However, these alterations in preload have not altered the intrinsic velocity of shortening, since all the curves extrapolate to the same intercept on the vertical axis. Thus, a change in initial length of heart muscle shifts the force-velocity curve by altering the total force which can be developed by the muscle.

This type of shift in the force-velocity curve may be contrasted with that obtained when a positive inotropic agent, such as norepinephrine or digitalis, is added to the muscle while the initial length is held constant (Fig. {right}). These agents not only increase the force which the muscle is capable of lifting, i.e., the intercept of the force-veolocity curve on the horizontal axis, but also increase the velocity of shortening of the unloaded muscle, i.e., the extrapolated intercept on the vertical axis."

— Harrison's (6th ed.)

Figures

Normal cat-muscle

Inotropic Agent

Figures removed due to copyright restrictions.

Clinical Knowledge

"... from the clinical point of view, heart failure may be considered to be a disease state in which an abnormality of myocardial function is responsible for the inability of the heart to pump blood at a rate commensurate with the requirements of the metabolizing tissues. Though a defect in myocardial contraction always exists in heart failure, this disorder may result from a *primary abnormality* in the heart muscle or it may be secondary to a *chronic excessive work load*. It is important to distinguish heart failure from (1) states of *circulatory insufficiency* in which myocardial function is not primarily impaired, such as cardiac tamponade, hemorrhagic shock, or tricuspid stenosis, (2) conditions in which there is *circulatory congestion* because of abnormal salt and water retention but in which there is no serious disturbance of myocardial function, and (3) conditions in which the normal heart is suddenly presented with a load which *exceeds its capacity*, e.g., accelerated hypertension." HST.950J / 6.872 Biomedical Computing Fall 2010

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.