The Future of the Application of Artificial Intelligence Methods to Medical Decision Making and Design of Information Systems for Health Care Institutions and Patients

Peter Szolovits
MIT Computer Science and Artificial Intelligence Lab
Prof. of Electrical Engineering & Computer Science
Prof. of Health Sciences & Technology
July 20, 2006

http://medg.csail.mit.edu/

~~The Future of the Application of Artificial Intelligence Methods to Medical Decision Making and Design of Information Systems for Health Care Institutions and Patients~~

Making Health Information Systems

Smarter

and

More Usable

"Medicine and the Computer: The Promise and Problems of Change"

—W.B. Schwartz, 1970

- Perceived problems
  - Physician shortage and maldistribution
  - Ever-expanding body of knowledge, so that the physician cannot keep up
- Exploit the computer as an "intellectual", "deductive" instrument
  - Improve medical care
  - Separate practice from memorization
  - Allow time for human contact
  - Encourage different personalities in medicine — the "healing arts"

AI's early (1970's to mid-80's) successes

- Diagnostic Programs
  - *Mycin* -- rule-based systems
  - *Internist/QMR/Caduceus, PIP, DXPLAIN* -- frame-based matching
  - *Acid/Base & Electrolytes* -- multi-level causal (pathophysiologic) reasoning
- Therapy Planning and Management
  - *Digitalis* -- pharmacokinetic models, feedback on risk and utility
  - *Heart Failure* -- qualitative temporal and pathophysiologic modeling
  - *Radiation targeting* -- computational geometry, tissue models

Mycin—Rule-based Systems

- **Task: Diagnosis and prescription for bacterial infections of the blood (and later meningitis)**
- **Method:**
  Collection of modular rules (400-700)
  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 primary-bacteremia 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**

. . .

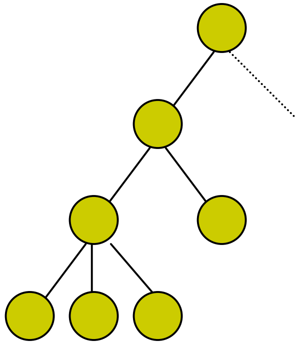
Davis, *et al.*, *Artificial Intelligence* **8**: 15-45 (1977)

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

- Goal-reduction (AND/OR) search

- *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).

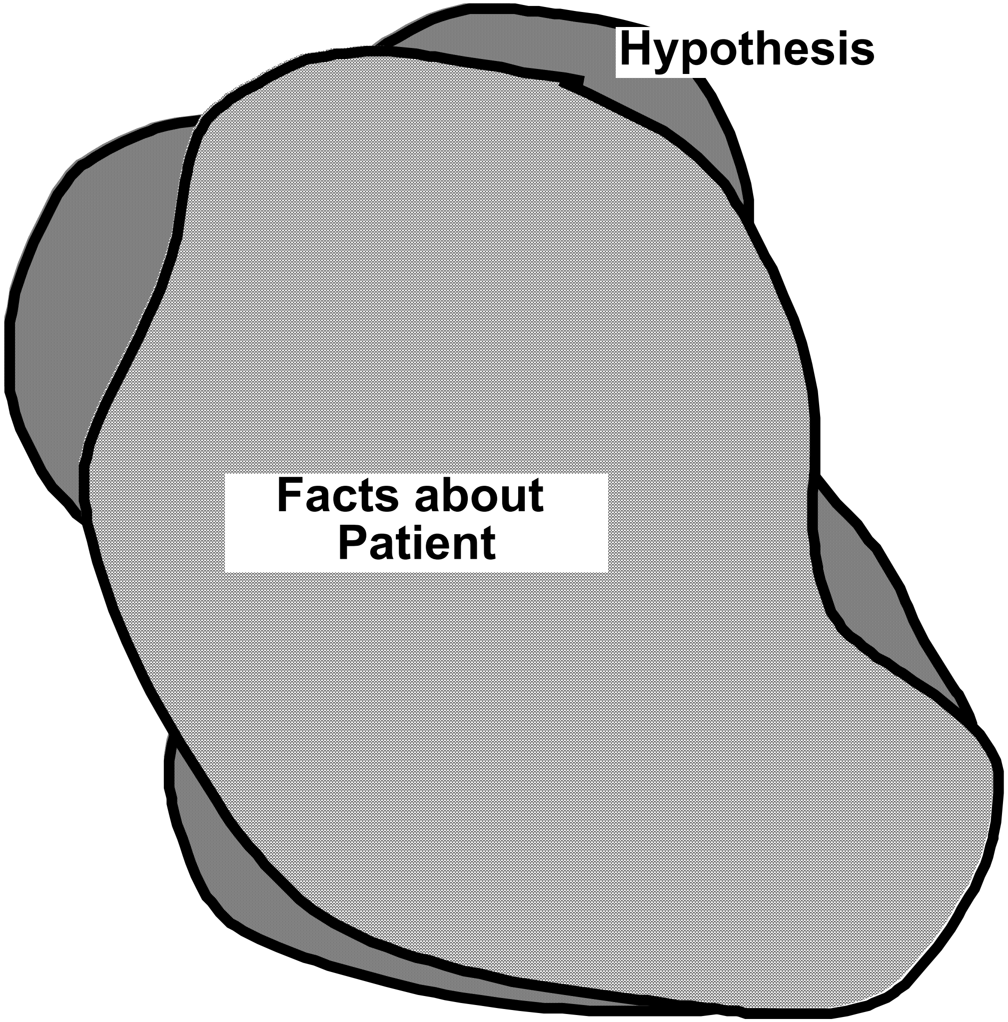
**** Why didn't you consider streptococcus as a possibility?**

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)

Diagnosis by Matching



Descriptions of Disease Support Diagnosis

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 ➲ constrictive pericarditis
ascites present ➲ cirrhosis
pulmonary emboli present ➲ renal vein thrombosis

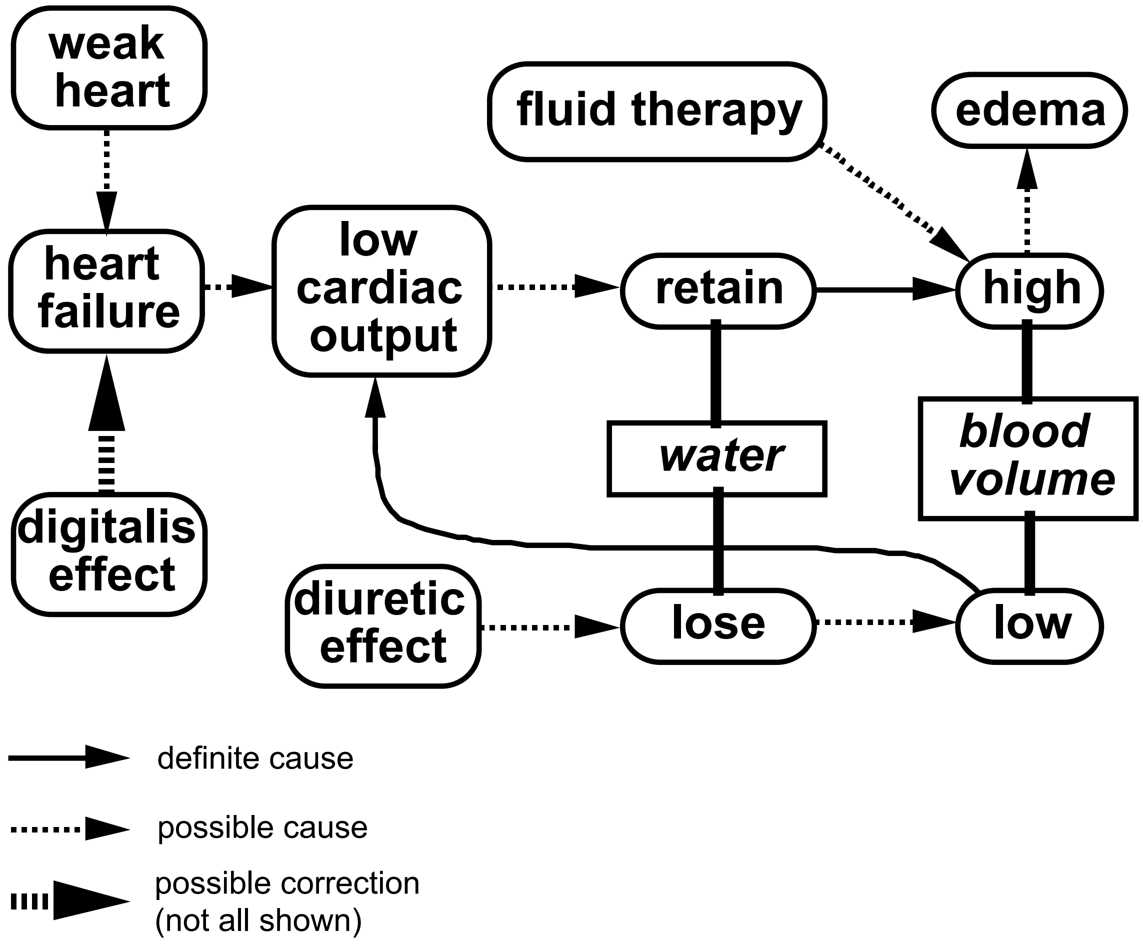
"Present Illness" Program's Theory of Diagnosis

- 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

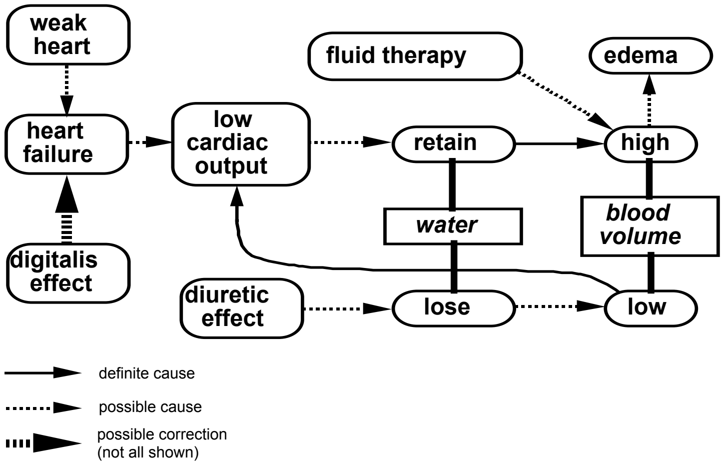
Causality

- Value of explicitly reasoning about “how it works:”
  - flexibility – ability to analyze unanticipated combinations
  - possibly richer descriptive language: relative timing and duration, severity, likelihood, nature of dependency
  - meaningful explanations and justifications
- Uses
  - Assessment of coherence of a complex hypothesis
  - Identification of components of an hypothesis or of the set of facts that don't fit properly
  - Prediction and postdiction
  - “Deeper” explanation

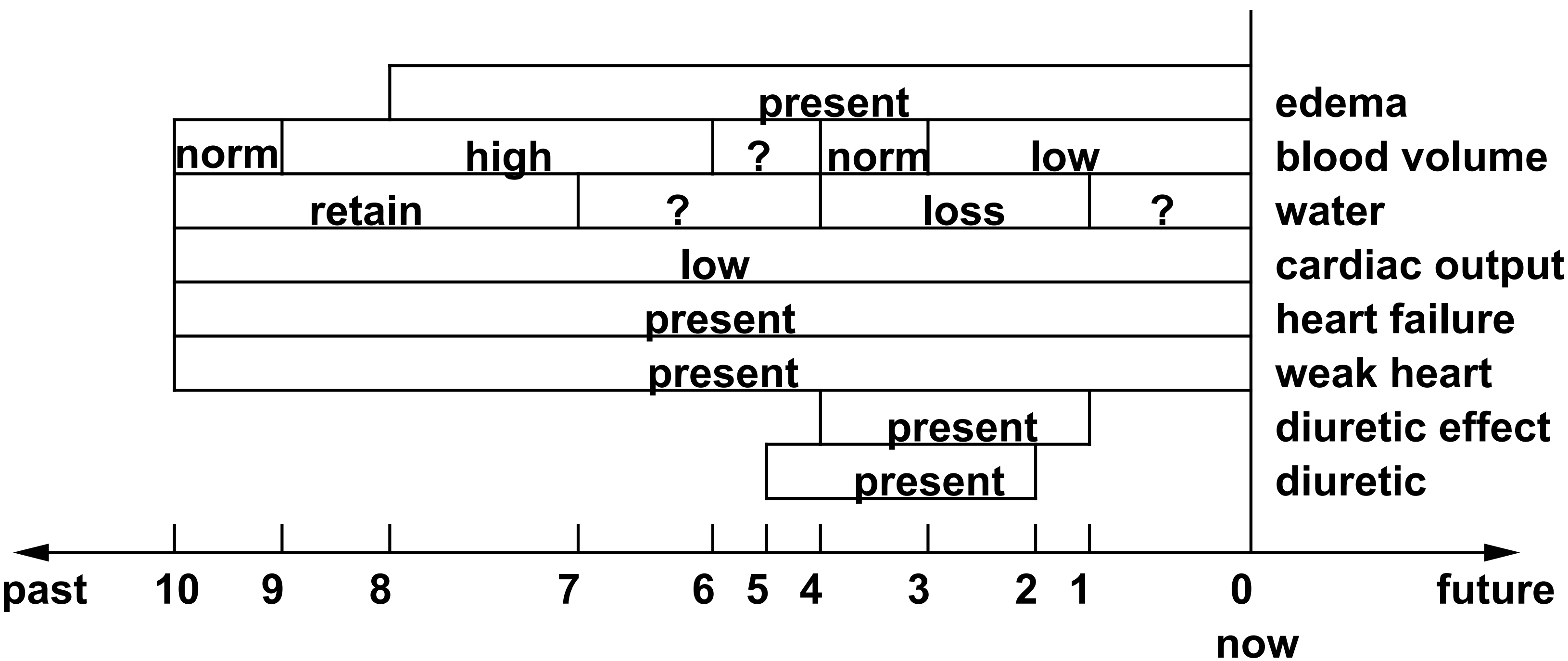
Interpreting the Past with a Causal/Temporal Model



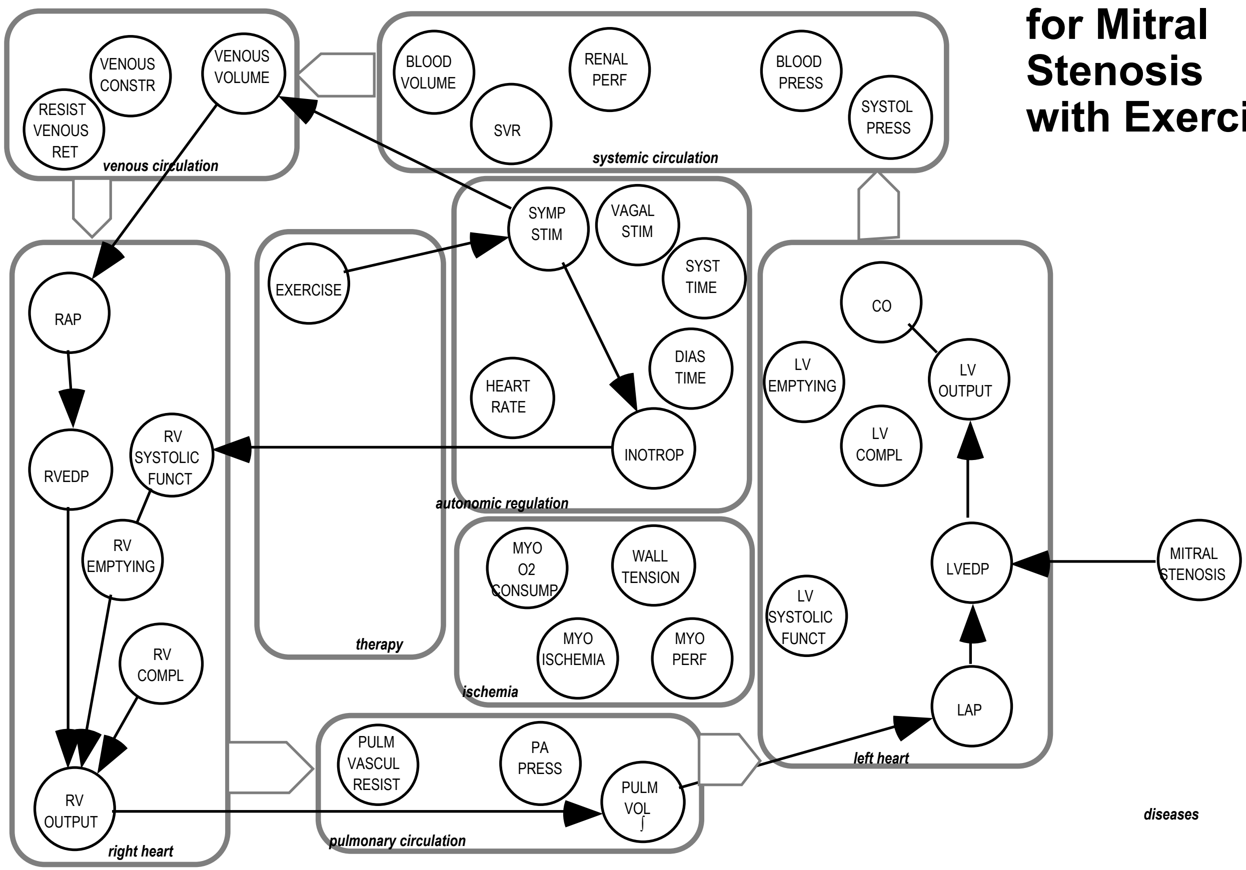
Postdiction



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



Causal Model for Heart Failure



Predictions for Mitral Stenosis with Exercise

“AI Summer” in 1980’s

- Generalized tools for Expert Systems
- Start-up companies, visions of $$$
- Thousands of applications
  - Campbell’s Soup, American Airlines, Digital Equipment, Aetna Insurance, …
- Today, these are just part of the infrastructure
  - Mail filters, configuration experts, program checkers, Amazon preferences, …

"AI Winter" by late 1980's

- Like .com bust of 2000
- Companies left in droves
  - Even successes were re-labeled
- Funding agencies turned to other approaches
- In medicine, lack of data made even exciting ideas virtually impossible to test
  - Students turned to other areas

Changes in Medicine in the Past 35 Years

- Attention to cost
- Fee for service ☑ Capitation
- Health maintenance organizations
- Outcomes research
- Economies of scale ☑ Integrated Delivery Networks
- Data collection and analysis
- Medical progress: Drugs, less-invasive surgery, genomics, …

A Flood of Data

- Lab systems
- Pharmacy
- Imaging (CAT, MRI, PET, …)
- Other clinical data (not exploding, exactly)
- Genetic tests: RFLP, SNP
- Full genomic sequence
- Expression under various circumstances

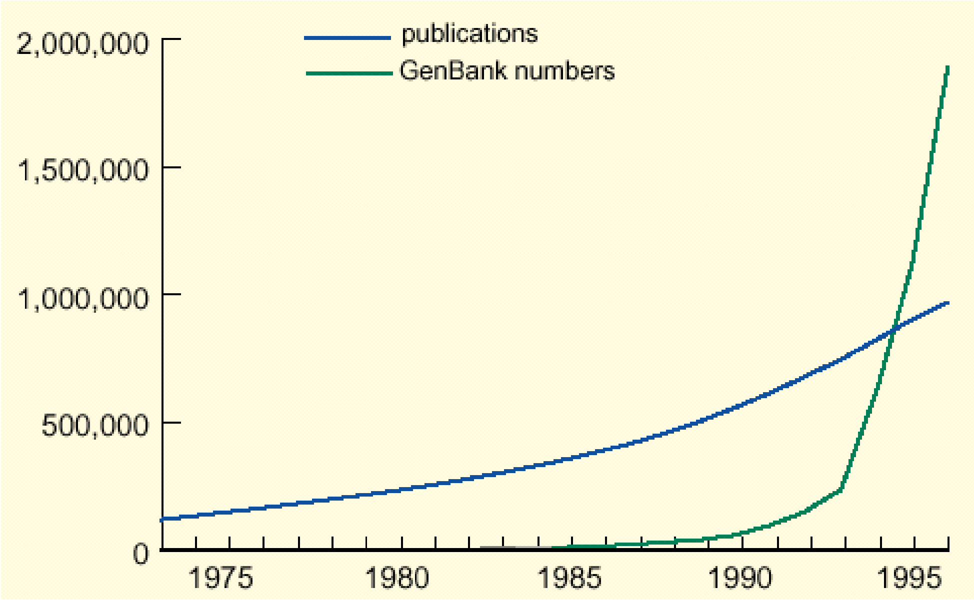
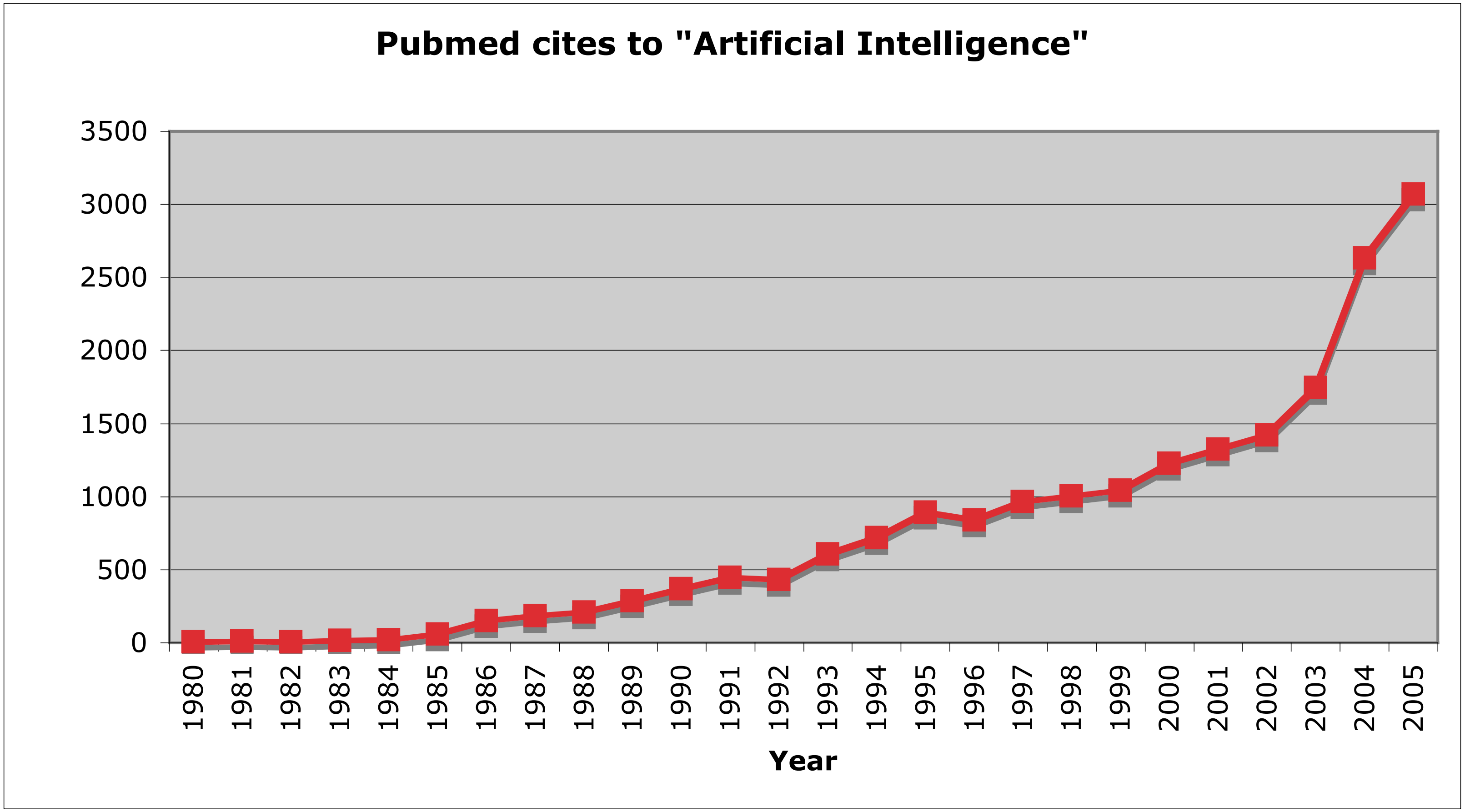


Fig. 1 Cumulative growth of molecular biology and genetics literature (blue) compared with DNA sequences (green). Articles in the 'G5' (molecular biology and genetics) subset of MEDLINE are plotted alongside DNA sequence records in GenBank over the same time period. The former data was obtained with the help of R.M. Woodsmall of NCBI and the latter data is available (ftp://ncbi.nlm.nih.gov/genbank/gbrel.txt). No attempt has been made to eliminate data redundancy among either the DNA sequence records or information contained in the literature.

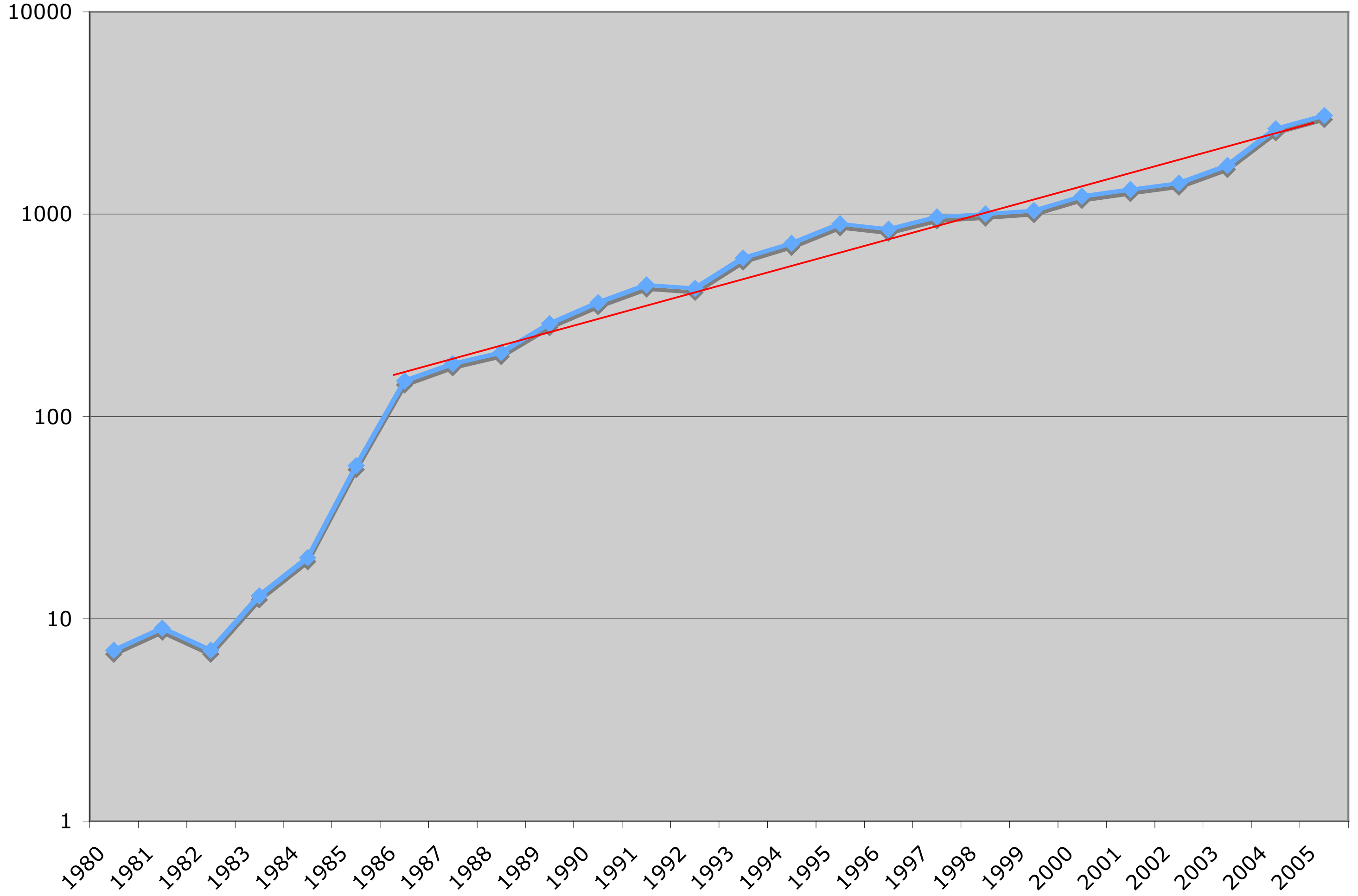
Knowledge ☑ Data

- Finally, usable data began to appear in 1990's
- Decline in "Expert Systems"
- Rise in machine learning, data mining

AI in Medicine grows rapidly



Exponential growth since 1986



Change in Medline topics

< 1990

Software; Diagnosis, Computer-Assisted; Computers; Expert Systems; Information Systems; Intelligence; Computer Simulation; Microcomputers; Models, Neurological; Algorithms; Decision Making, Computer-Assisted; Models, Psychological; Models, Theoretical; Computer-Assisted Instruction; Comparative Study; Decision Making; Models, Biological; Intelligence Tests; Medical Records; Cognition; Problem Solving; …

≥ 1995

Algorithms; Reproducibility of Results; Sensitivity and Specificity; Comparative Study; Computer Simulation; Pattern Recognition, Automated; Image Interpretation, Computer-Assisted; Neural Networks (Computer); Signal Processing, Computer-Assisted; Models, Statistical; Information Storage and Retrieval; Software; Cluster Analysis; Image Enhancement; Models, Biological; User-Computer Interface; Numerical Analysis; …

In the Era of Plentiful Data…

- Exploit simple relations first:
  - Don't give toxic doses of meds (mg vs. μg)
  - Don't give medications to which a patient is allergic
  - Don't give medications that have been ineffective for this patient
  - Generic drugs work well, cost less; studies can show this
  - Don't give renally excreted meds to patients with impaired renal function
  - …

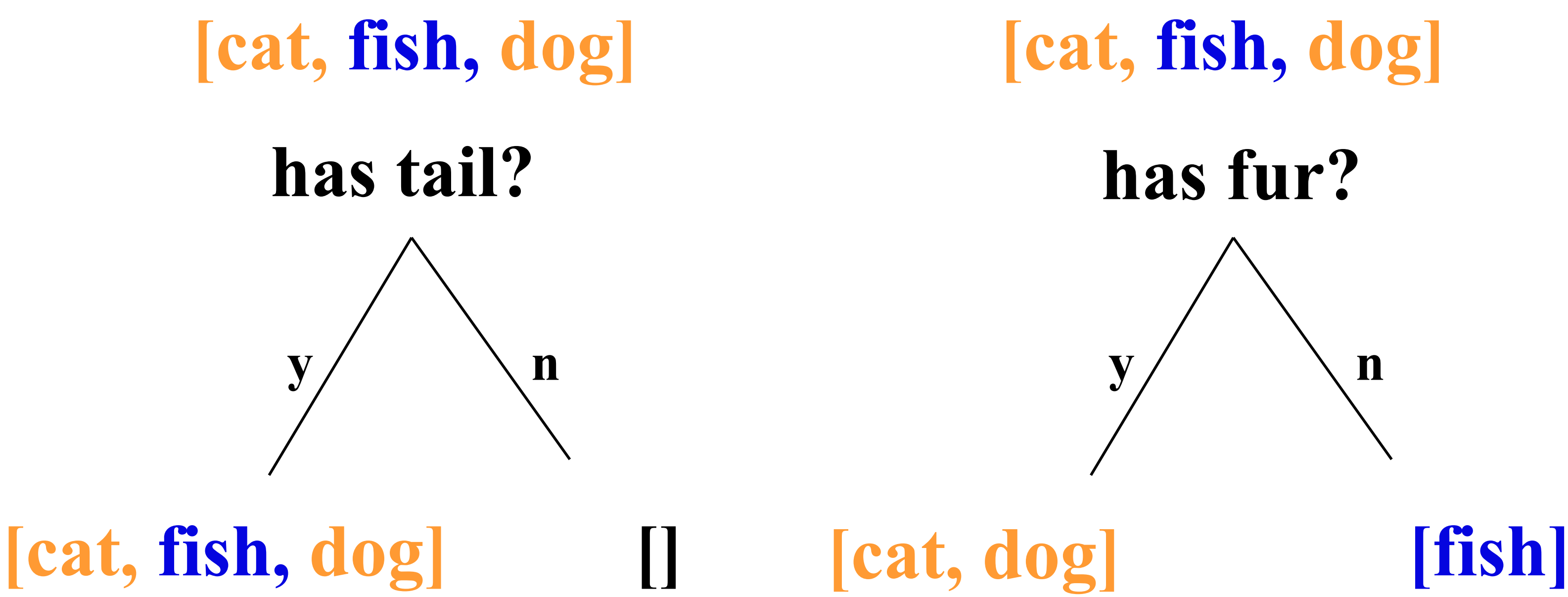
Learn New Useful Associations

- Who has an MI?
- 1752 ER patients with non-traumatic chest pain
  - Edinburgh Royal Infirmary, Scotland (1252 cases)
  - Sheffield Northern General Hospital, England (500 cases)
  - Kennedy *et al.*
- 45 attributes for each case

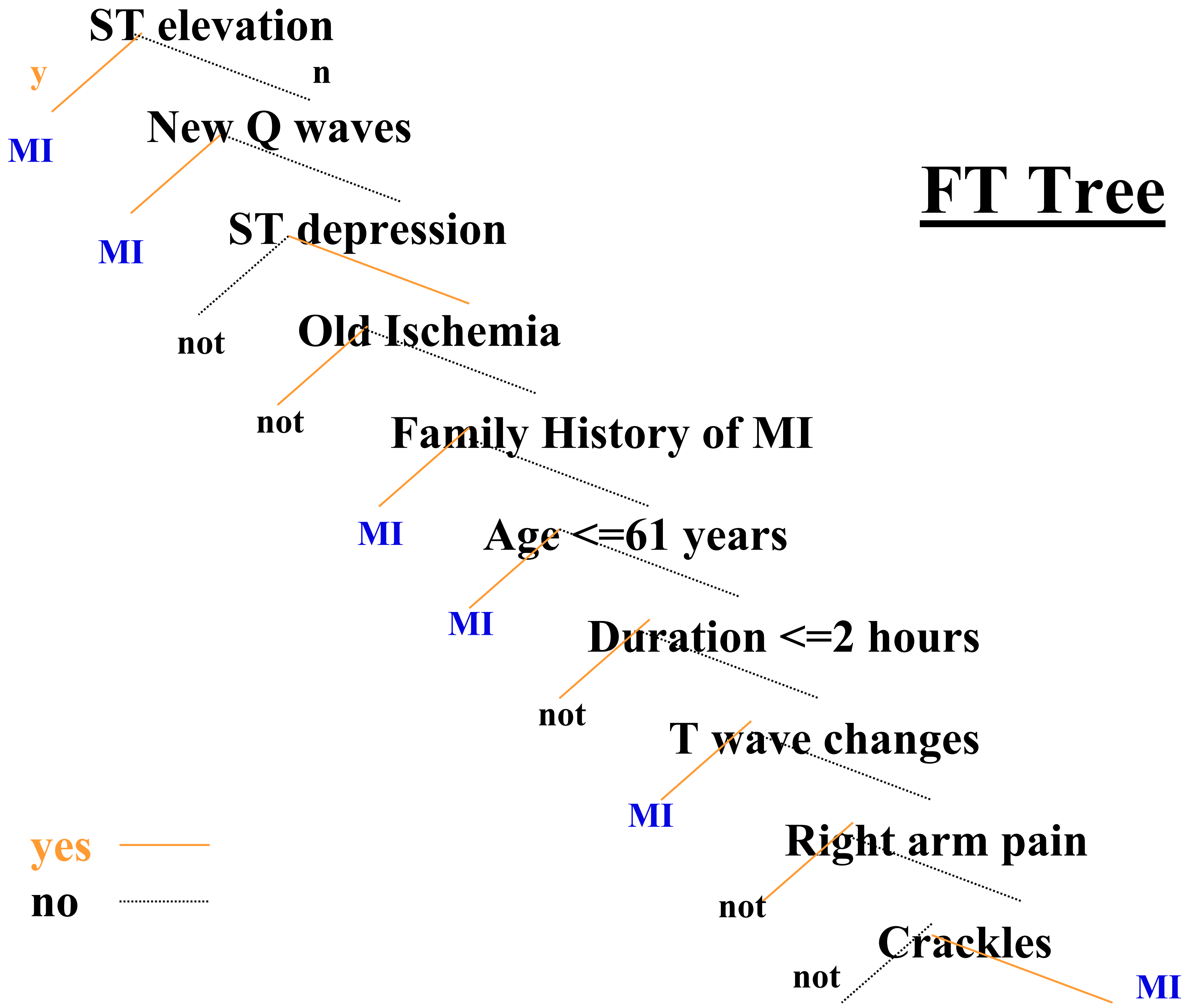
- Demographics
  - age, sex
- Coronary artery disease risk factors
  - smoker, diabetes, high BP
- History of pain
  - duration, “sharp,” radiating to the back
- Physical examination
  - crackles, chest wall tenderness
- ECG data
  - ST elevation, ST depression

--Chris Tsien, *et al.*

Example Tree: Trying to classify "Mammal" vs. "Not Mammal"

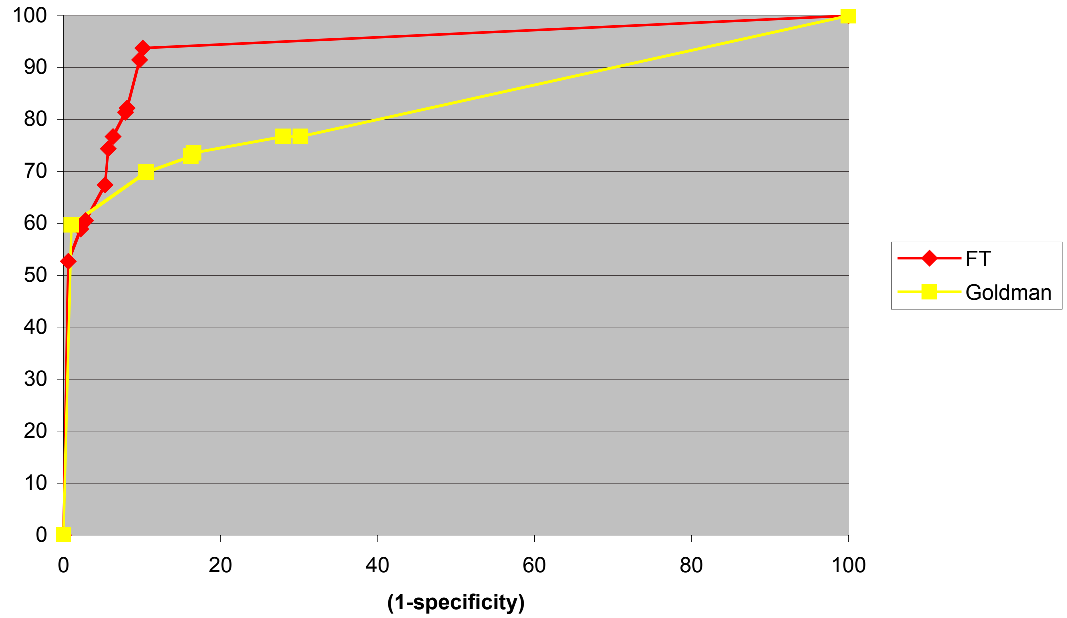


FT Tree



ST elevation
y
n
MI
New Q waves
MI
ST depression
not
Old Ischemia
not
Family History of MI
MI
Age <=61 years
MI
Duration <=2 hours
not
T wave changes
MI
Right arm pain
not
Crackles
not
MI
yes
no

Goldman Tree vs. FT Tree on Edinburgh test data, $p < 0.0001$



ROC Area:

FT Tree: 94%

Goldman: 84%

(Long: 86%)

Other Ways to Exploit Data

- Nearest Neighbor methods
- Clustering, Self-Organizing Maps, etc.
- Neural Networks, Radial Basis Functions
- Logistic Regression
- Rough Sets, Support Vector Machines
- Induce Bayes Networks
- Markov Decision Processes (Semi-Markov, Partially-Observable, …)

Whence New Knowledge?

- Traditionally:
  - Form hypothesis
  - Design, conduct experiment
  - Evaluate/revise hypothesis
- In time of Data Glut
  - Design experiment to collect data possibly relevant to countless hypotheses
  - Search data for interesting relations
  - Form hypotheses

Dogma

Phenotype	= Genotype	+ Environment
Traits	Gene sequence	Diet, smoking, drugs, …
Diseases	SNP's	Insults and injuries
Behaviors	Expression data	Exposures
…	…	…

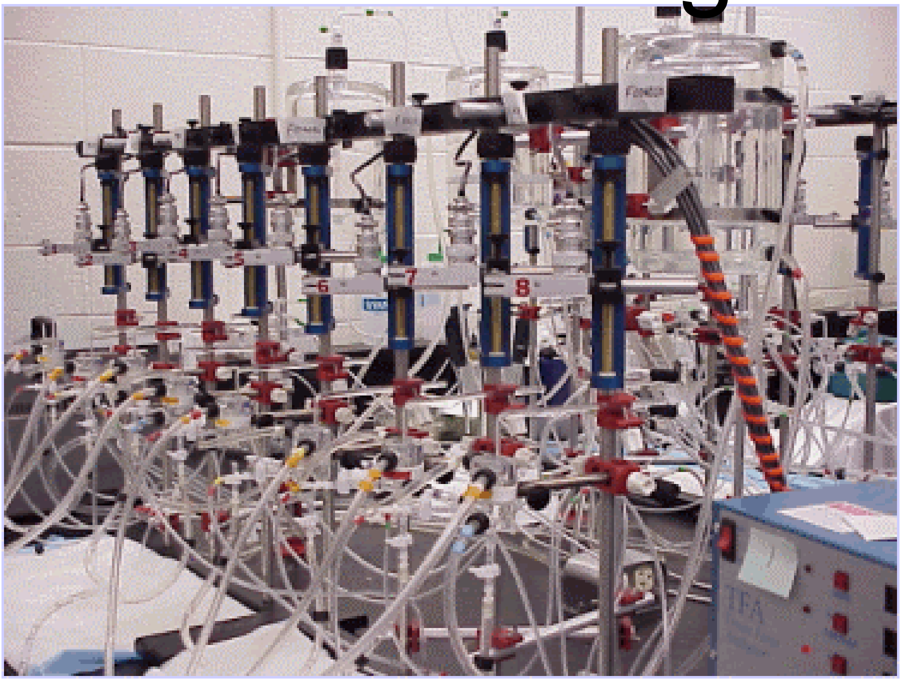
- What is the functional form?
- How do we investigate these relationships?
- Can we take advantage of the exponential growth of genomic data?

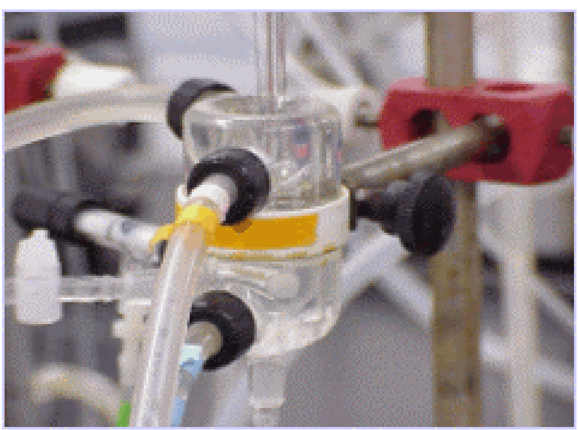
Where are the Phenotype and Environment-related Data?

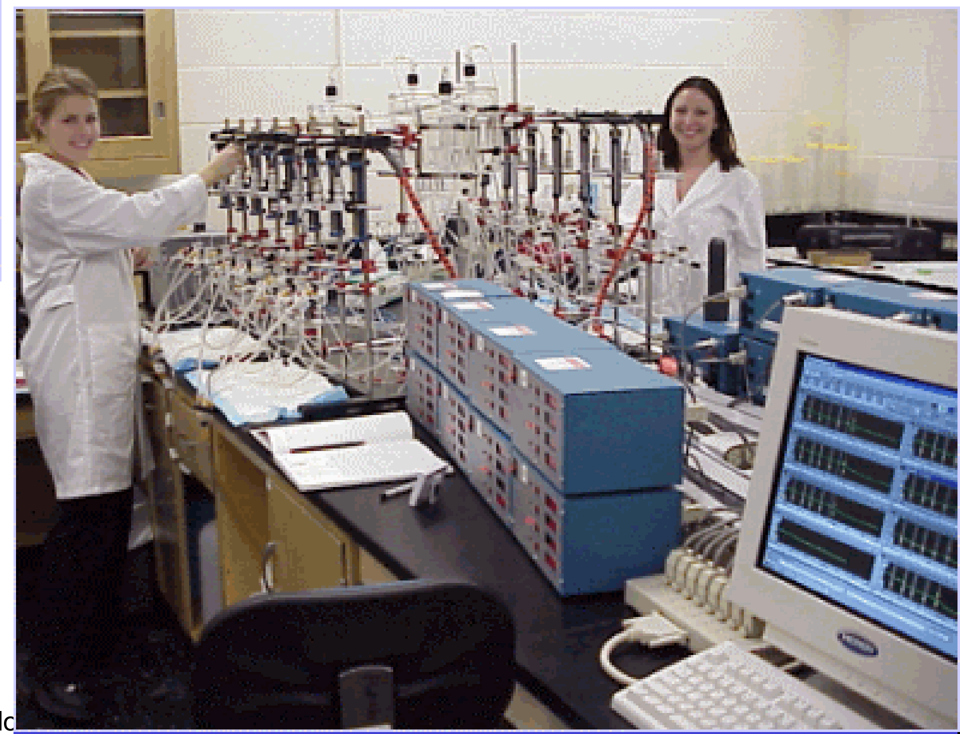
Phenotype	= Genotype	+ Environment
Traits	Gene sequence	Diet, smoking, drugs, …
Diseases	SNP’s	Insults and injuries
Behaviors	Expression data	Exposures
…	…	…

- Perform Controlled Experiments?
  - Unethical using human subjects!!!
  - OK on rats.

High-throughput phenotyping at Medical College of Wisconsin



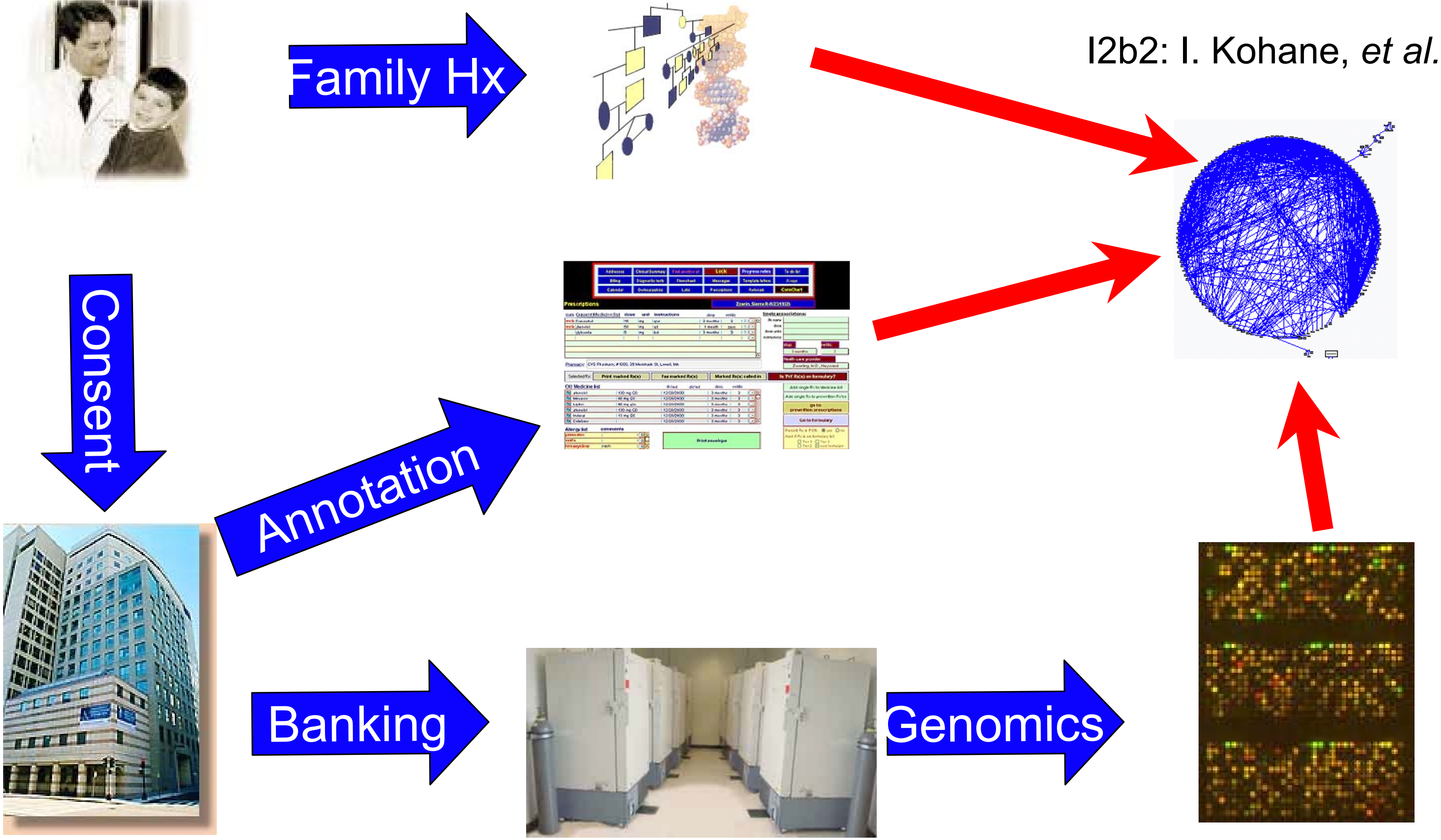




Where are the Phenotype and Environment-related Data?

- Environment
  - (Hardest to get)
  - Questionnaires,
    - e.g., Nurses' Health Study, Framingham Heart Study
  - Monitoring
    - e.g., LDS hospital infectious disease monitors
- Phenotype
  - "Natural Experiments"
  - ∴ Clinical Data

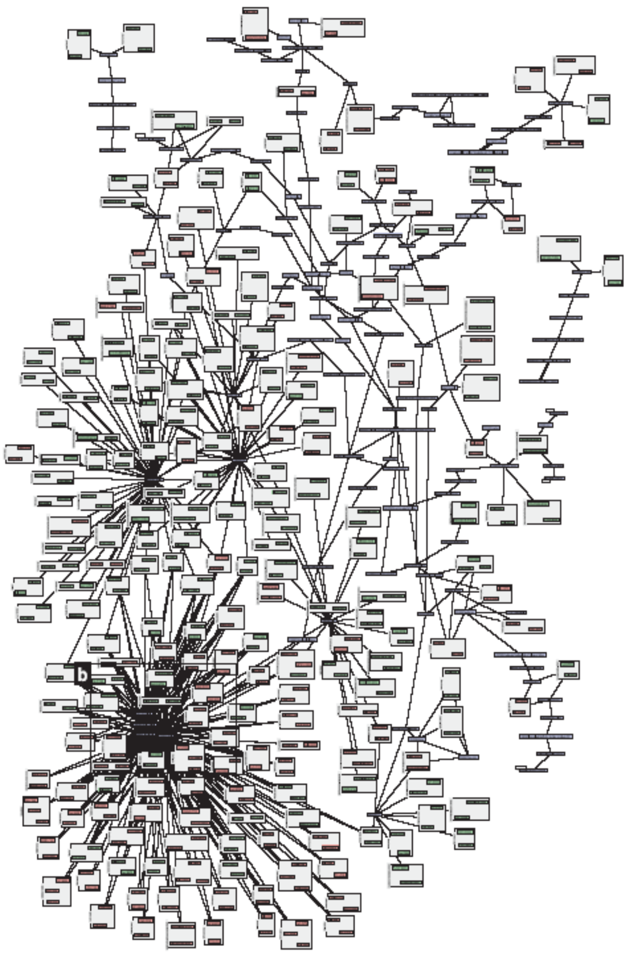
The fantasy: Informatics for Integrating Biology & Bedside



Plausibility

Butte & Kohane, *Nature Biotech* 2006

- Phenome-Genome Network
  - Gene Expression Omnibus
    - expression data
    - annotations: tissue, disease, experimental conditions, …
  - Interpret annotations to UMLS
  - Differential expression vs. condition
  - Interesting relations:
    - 11 genes & aging
    - DDX24 and leukemia
    - 2 genes & injury



Clinical Data are Mostly Text

- Need Text Understanding
  - Discharge summaries
  - Clinical notes (admitting, doctors', nurses', …)
  - Reports (radiology, pathology, …)
  - Letters
- Exceptions:
  - Lab data
  - Pharmacy orders
  - Billing codes
  - Images (but, need image understanding)

Octo Barnett's objection to "free text"

Text is critical, even in ICU

- Data not otherwise captured:
  - Procedures
  - Patient state
  - Medications
  - Episodic measurements
  - …

Current Needs & Opportunities

- Extraction of codified data from text
- Machine learning from vast data collections
  - For diagnosis, prognosis and therapy
- Revisiting symbolic, knowledge and model-based methods once the low-hanging fruit are picked
- Understanding, modeling and integrating with workflows

Medical Record Challenges

- Paper → Electronic
- Unstructured → Coded
  - SNOMED, ICD, HL7 structured documents, …
- Institutional → Patient-centered and controlled
  - Collection of all relevant data from all providers
  - Basis for monitoring, feedback, education, decision making
  - http://ga.org & http://ping.chip.org
- Integration of patient care, public health & research
  - Supporting workflow

Acknowledgements

- Colleagues
  - MIT: Bill Long, Ozlem Uzuner, Regina Barzilay, Ramesh Patil, Randy Davis
  - Harvard: Isaac S. Kohane, Marco Ramoni, Scott Weiss, Shawn Murphy, Henry Chueh, Ken Mandl, Susanne Churchill, …
  - Tufts: Stephen Pauker, Jerome Kassirer, William Schwartz
- Students
  - Christine Tsien, Atul Butte, Neeha Bhooshan, Jennifer Shu, Tawanda Sibanda, Tian He, Steve Kannan, Phil Bramsen, …
- $$$: NIH: NLM & NIBIB

The End

http://medg.csail.mit.edu