CSE 473
Artificial Intelligence

Review

Logistics

- Project due tonight
- Exam next Mon 2:30—4:20
  - Regular classroom
  - Closed book
  - Cover all quarter's material
  - Emphasis on material not covered on midterm

Defining AI

human-like *vs.* rational

	human-like	rational
thought *vs.* behavior	**Systems that think like humans**	**Systems that think rationally**
	Systems that act like humans	**Systems that act rationally**

Goals of this Course

- To introduce you to a set of key:
  - Paradigms &
  - Techniques
- Teach you to identify when & how to use
  - Heuristic search
  - Constraint satisfaction
  - Planning
  - Logical inference
  - Bayesian inference
  - Policy construction
  - Machine learning

Theme I

- Problem Spaces & Search

How to specify PS?

Two kinds of search?

Learning as Search

- Decision trees
- Structure learning in Bayesian networks
- Unsupervised clustering
- Boosting

Theme II

- In the knowledge lies the power
- Adding knowledge to search

Heuristics

- How to generate?
- Admissibility?

Propositional Logic *vs.* **First Order**

Ontology	Facts (P, Q)	Objects, Properties, Relations
Syntax	Atomic sentences Connectives	Variables & quantification Sentences have structure: terms father-of(mother-of(X)))
Semantics	Truth Tables	Interpretations (Much more complicated)
Inference Algorithm	DPLL, WalkSAT Fast in practice	Unification Forward, Backward chaining Resolution, theorem proving
Complexity	NP-Complete	Semi-decidable

Planning

- Problem solving algorithms that operate on explicit propositional representations of states and actions.
- Make use of specific heuristics.
- State-space search: forward (progression) / backward (regression) search
- Partial order planners search space of plans from goal to start, adding actions to achieve goals
- GraphPlan: Generates planning graph to guide backwards search for plan
- SATplan: Converts planning problem into propositional axioms. Uses SAT solver to find plan.

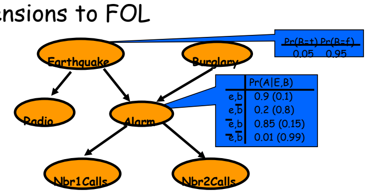
Probabilistic Representations

- How encode knowledge here?

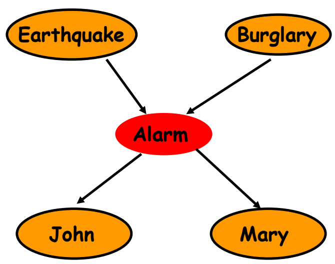
In the knowledge lies the power

Uncertainity

- Joint Distribution
- Prior & Conditional Probability
- Bayes Rule
- [Conditional] Independence
- Bayes Net
  - Propositional
  - Hot topic: extensions to FOL

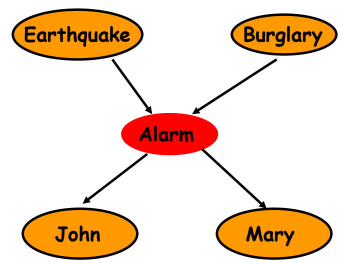


P(B | J=true, M=true)



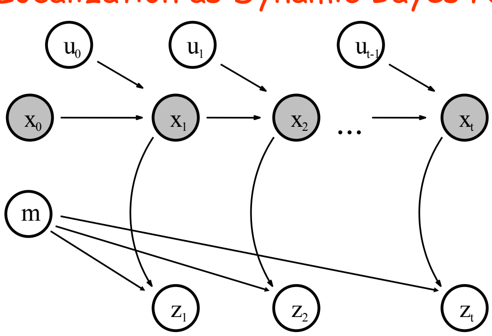
$$P(b|j,m) = \alpha \sum_{e,a} P(b,j,m,e,a)$$

P(B | J=true, M=true)

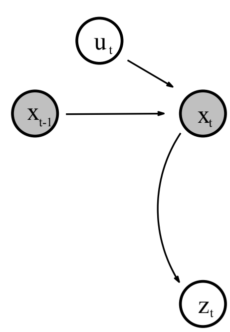


$$P(b|j,m) = \alpha P(b) \sum_{e} P(e) \sum_{a} P(a|b,e)P(j|a)P(m,a)$$

Localization as Dynamic Bayes Net



Markov Assumption Helps!



Representations for Bayesian Robot Localization

Discrete approaches ('95)
- Topological representation ('95)
  - uncertainty handling (POMDPs)
  - occas. global localization, recovery
- Grid-based, metric representation ('96)
  - global localization, recovery

Particle filters ('99)
- sample-based representation
- global localization, recovery

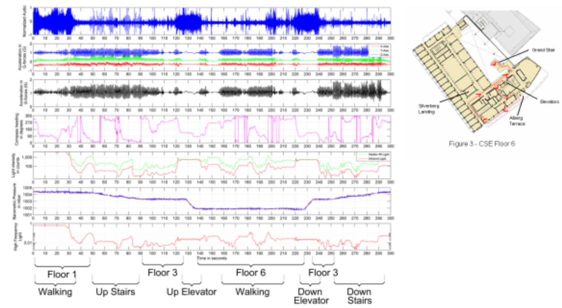
AI

Kalman filters (late-80s?)
- Gaussians
- approximately linear models
- position tracking

Robotics

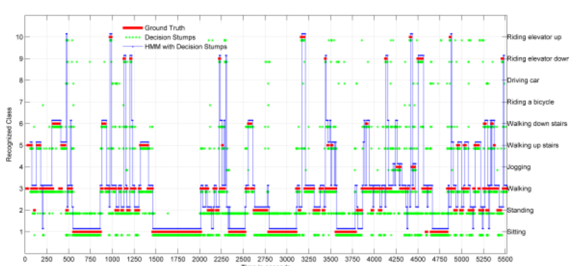
Multi-hypothesis ('00)
- multiple Kalman filters
- global localization, recovery

Sensor board: Data Stream



Courtesy G. Borriello

Example Evaluation Run



Decision stumps classifiers (at 4Hz)
HMM with probabilities as inputs (using a 15 second sliding window with 5 second overlap)
Ground truth for a continuous hour and half segment of data.

Specifying an MDP

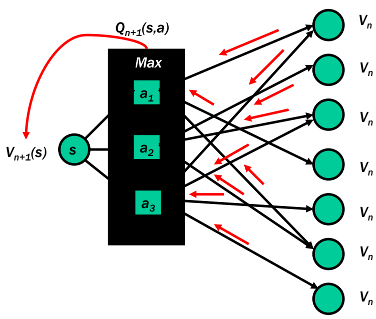
S = set of states set (|S| = n)

A = set of actions (|A| = m)

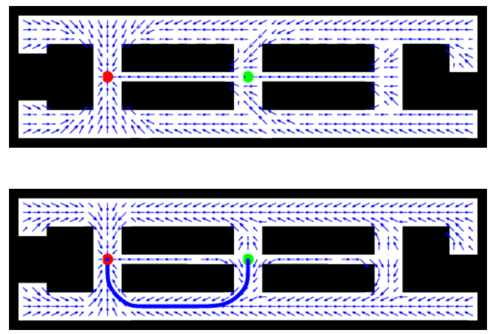
Pr = transition function Pr(s,a,s')
represented by set of m n x n stochastic matrices
each defines a distribution over SxS

R(s) = bounded, real-valued reward fun
represented by an n-vector

Bellman Backup, Value Iteration



Stochastic, Fully Observable



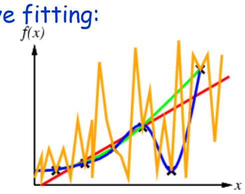
Why is Learning Possible?

Experience alone never justifies any conclusion about any unseen instance.

Learning occurs when PREJUDICE meets DATA!

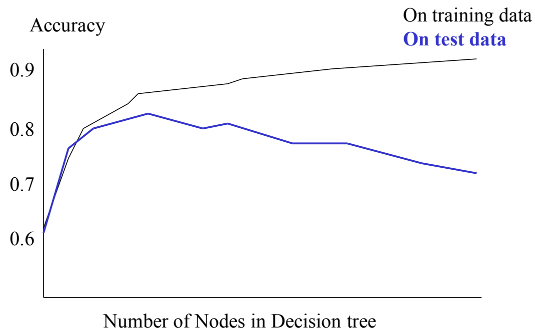
Inductive learning method

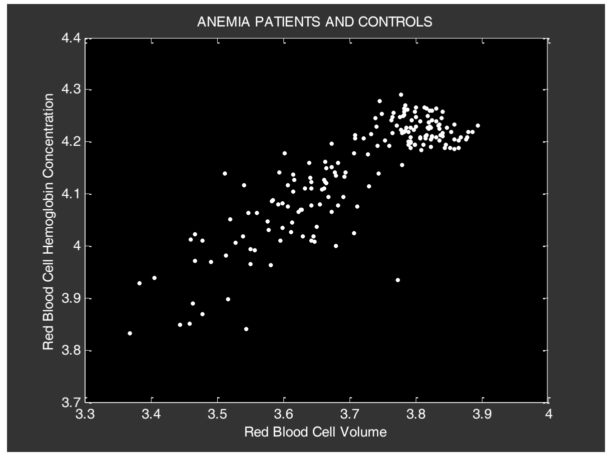
- Construct/adjust *h* to agree with *f* on training set (*h* is consistent if it agrees with *f* on all examples)
-
- E.g., curve fitting:

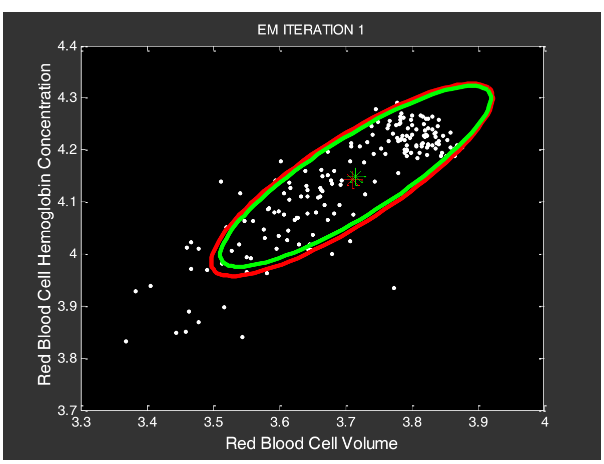


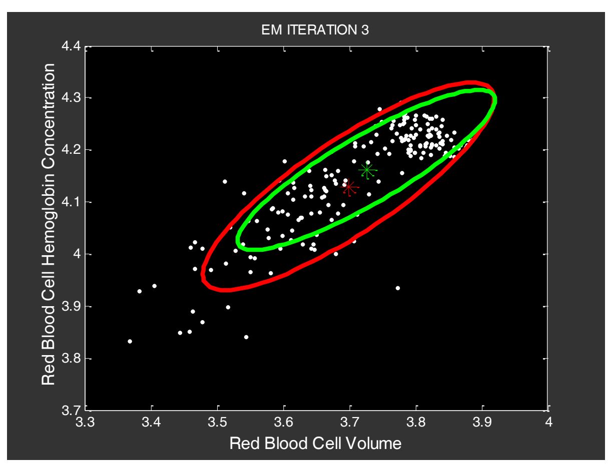
- Ockham's razor: prefer the simplest hypothesis consistent with data

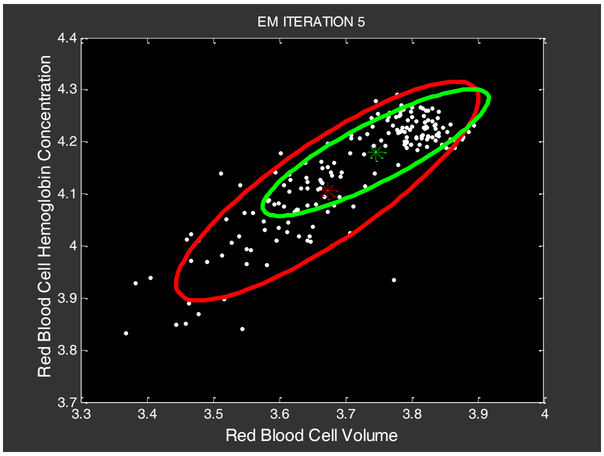
Decision Tree Overfitting

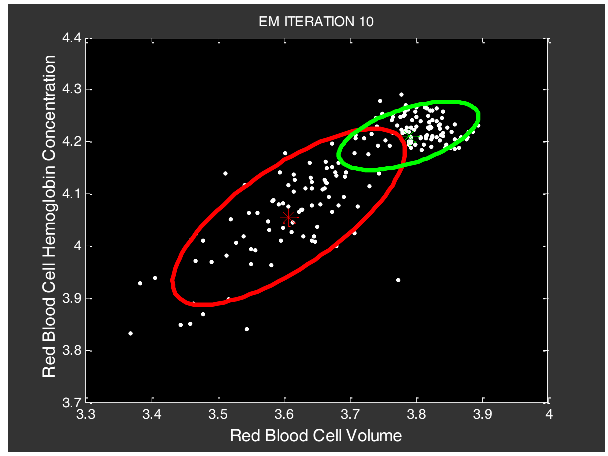


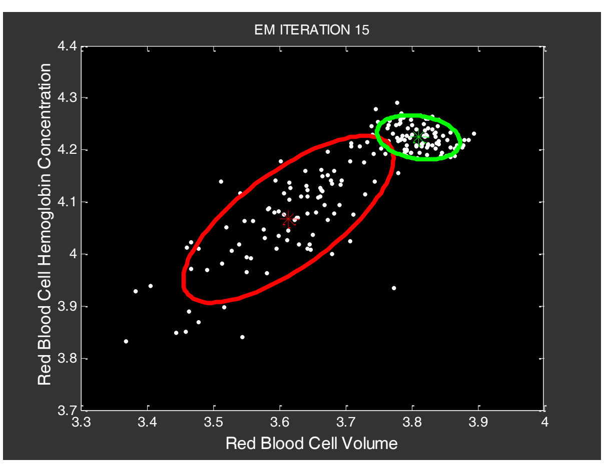


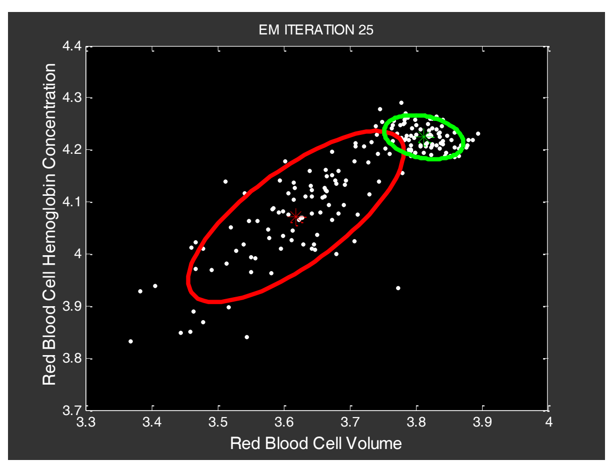


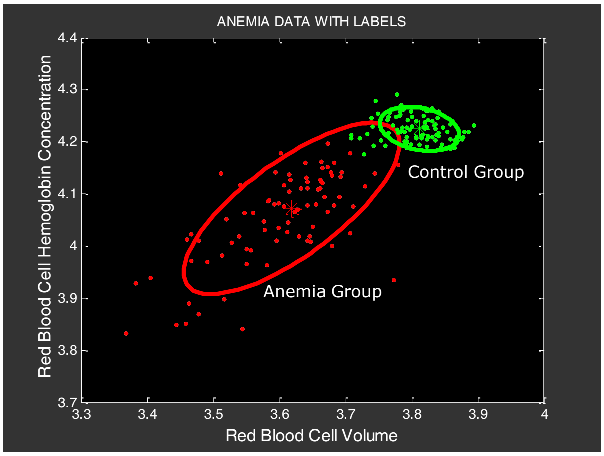



EM ITERATION 5
Red Blood Cell Hemoglobin Concentration
Red Blood Cell Volume
3.3
3.4
3.5
3.6
3.7
3.8
3.9
4
3.7
3.8
3.9
4
4.1
4.2
4.3
4.4


EM ITERATION 10
Red Blood Cell Hemoglobin Concentration
Red Blood Cell Volume
3.3
3.4
3.5
3.6
3.7
3.8
3.9
4
3.7
3.8
3.9
4
4.1
4.2
4.3
4.4


EM ITERATION 15
Red Blood Cell Hemoglobin Concentration
Red Blood Cell Volume
3.3
3.4
3.5
3.6
3.7
3.8
3.9
4
3.7
3.8
3.9
4
4.1
4.2
4.3
4.4


EM ITERATION 25
Red Blood Cell Hemoglobin Concentration
Red Blood Cell Volume
3.3
3.4
3.5
3.6
3.7
3.8
3.9
4
3.7
3.8
3.9
4
4.1
4.2
4.3
4.4



And More

- Specific search & CSP algorithms
- Adversary Search
- Inference in Propositional & FO Logic
- Learning: decision trees, boosting, EM, RL
- Lots of details