Formulations of Learning Problem

- Improve performance P at task T from experience E
- Learn function f given $\{\langle x, f(x)\rangle\}$
  - Boolean-valued, discrete-valued, real-valued, vector-valued f
  - From discrete, real-valued inputs x
  - Noisy data, or noise-free (x or $f(x)$)
  - From randomly generated examples, or from active queries
- Learn from delayed reward (e.g., reinforcement learning)
- Learn from $\{\langle x, f(x)\rangle\}$ plus prior knowledge
- Learn from unclassified examples (clustering)

Formulation influences choice of hypothesis space, algorithm, and constrains its sample complexity

©Tom M. Mitchell, 1995

Hypothesis Representations

- Symbolic/discrete
  - Propositional (variable-free)
  - First order (variables)
- Decision trees
- Neural networks (and other parameterized expressions)
- Explicit probabilities (e.g., $P(a_i = v_j|h)$)
- Instances (e.g., nearest neighbor)

©Tom M. Mitchell, 1995

Search Methods/Hypothesis Space Structure

- Specific-to-general, general-to-specific (e.g., VS)
- Simple-to-complex (e.g., C4.5)
- Gradient descent (e.g., Backpropagation)
- Greedy divide and conquer (e.g., C4.5)
- Greedy covering (e.g., FOIL)
- Randomized search (e.g., GA's)

©Tom M. Mitchell, 1995

Perspectives: Learning as ...

- Search in hypothesis space
- Efficiently approximating Bayes optimal classifications
- An optimization problem (to optimize what?)
- Combining inductive and deductive reasoning
- Statistical estimation
- Futile without inductive bias

©Tom M. Mitchell, 1995

Frameworks for Characterizing Learners

- Inductive bias ($data + bias \vdash classifications$)
- Priors $P(h)$ and $P(D|h)$
- Ability to produce MAP or ML hypotheses
- Sample complexity (and computational complexity)
- VC dimension
- Mistake bounds

©Tom M. Mitchell, 1995

What is state of the art of ML?

Practical approaches:

- ANN Backpropagation, and decision tree learning quite useful
- So are Bayesian and other methods in some cases
- Cross validation important for choosing parameters
- Data character often determines learning algorithm

Some successful applications:

- Find database regularities: credit risk evaluation
- Classify text: NewsWeeder
- Image classification: face recognition
- Robot control: autonomous driving
- Game playing: Backgammon

©Tom M. Mitchell, 1995

Current Research Topics

Combine prior knowledge with induction

- KBANN, EBNN, ...
- learning Bayesian networks
- ...

Consider *learning agents* rather than *function approximation*

- Reinforcement learning, delayed rewards
- Active experimentation strategies
- Lifelong learning
- Learn new representations
- Strategies for agent to choose what/when to learn
- Long-term memory indexing/recall

©Tom M. Mitchell, 1995

Next Great Applications?

Data mining

- Learn to detect credit card fraud, medical patient outcomes, ...
- Great commercial interest (ATT, IBM, SiGraphics, SPR, ...)

World wide web?

- Newsreaders that learn your interests
- Web browsers that learn your interests and where to find things
- Mining the web as the world's (self-maintaining) knowledge base?

Learning to understand natural language?

- e.g., learn semantics of web hyperlinks

Intelligent self-adapting buildings?

- e.g., building learns to control heat, ventilation, light to minimize energy costs and maximize comfort to occupants

©Tom M. Mitchell, 1995