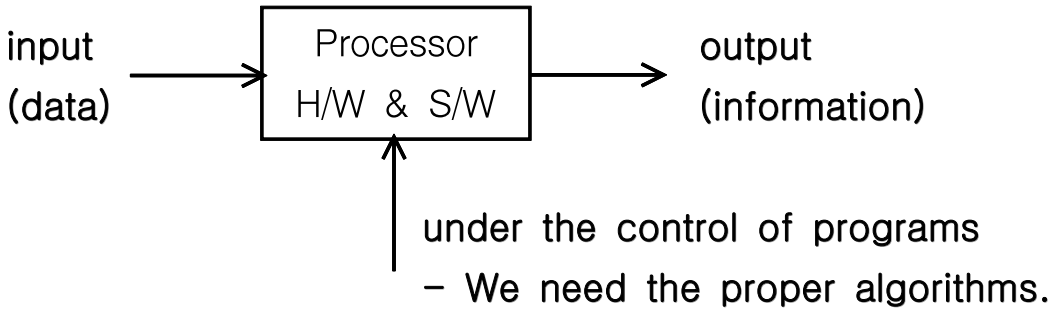
Introduction

- Information Processing System



1990s software crisis: hard to find out algorithms
(especially related to intelligence)

- Human Information Processing System

. The human information processing system consists of the biological brain.

. The basic building block of the nervous system is the neuron, the cell that communicates information to and from the various parts of the body.

. The neuron consists of a cell body called a soma, several spine-like extensions of the cell body called dendrites, and a single nerve fiber called the axon that branches out from the soma and connects to many other neurons.

- Artificial Neural Networks (ANNs)

. An ANN has a parallel, distributed information processing structure consisting of processing elements (which can possess a local memory and carry out localized information processing operations) interconnected together with unidirectional signal channels called connections.
. Each processing element has a single output connection which branches ("fan out") into as many collateral connections as desired.
. All of the processing that goes on within each processing element must be completely local; that is, must depend only on the current values of the input signal.

- Classes of Artificial Neural Networks

. interconnection schemes:
1) Recurrent connections are connections that loop and connect back to the same processing element.
2) Feedforward connections are connections that allow information to flow amongst processing elements in one direction.

. supervised and unsupervised learning
1) Supervised learning implies the learning of ANNs for the given pairs of input and output patterns.
2) Unsupervised learning implies the learning of ANNs for the given patterns without supervised labels.

- Classes of Artificial Neural Networks

Learning / Recall	Feedback Recall	Feedforward Recall
Unsupervised Learning	Adaptive Resonance Theory (ART) Hopfield Networks Bidirectional Associative Memory (BAM) Boltzmann Machine (BM) Cauchy Machine (CM)	Fuzzy Associative Memory (FAM) Learning Vector Quantization (LVQ) Counterpropagation Networks (CPNs)
Supervised Learning	Brain-State-in-a-Box (BSB) Fuzzy Cognitive Map (FCM)	Perceptrons Multilayer Perceptrons (MLPs) Radial Basis Function Networks (RBFNs)

- Machine Learning

. programming computers to *learn* instead of finding out algorithms

. definition of learning:

a computer program is said to learn from experience E (data) with respect to some class of tasks T and performance P, if its performance at tasks in T, as measured by P, improves with experience E.

example: learning to recognize spoken words
  task T: recognizing spoken words
  performance measure P: recognition rate
  training experience E: speech data

example: learning to play checkers
  task T: playing checkers
  performance measure P: percentage of games won
  training experience E: playing practice games against itself

- Some disciplines related to machine learning:

artificial intelligence (AI),
computational learning theory (COLT),
information theory, probability theory, statistics,
differential equations, function analysis,
psychology and neurobiology,
philosophy, etc.

- Design procedure of machine learning system

Step 1. Choosing the training experience
Step 2. Choosing the target function
(what type of knowledge will be learned)
Step 3. Choosing a representation for the target function
(functional form of targets)
Step 4. Choosing a learning algorithm
Step 5. Measuring the performance and updating
the learning system

- Issues in machine learning

. What algorithms can approximate functions well?
. How does training examples influence accuracy?
(sample complexity)
. What is the upper bounds of general errors for
a learning system? (generalization bounds)
. How does noisy data influence accuracy?
. What is the optimal structure of learning models?
. How can prior knowledge of learner help?
. What clues can we get from biological learning system?

- Performance improvement of machine learning models

. mean square error (MSE):

MSE(Model) = Variance(Model) + Bias(Model)^2

. committee machines:

trying to minimize the variance terms of machine learning models by increasing the number of weak models.

. deep learning models:

trying to minimize the bias terms of machine learning models by increasing the number of layers.

- Topics for this course

. definition of machine learning
. concept learning, decision tree learning
. Perceptrons, linear models
. parameter estimation methods (learning algorithms)
. artificial neural networks (MLPs, RBFNs, etc.)
. evaluation methods
. computational learning theories, support vector machines
. Bayesian belief networks, class probability output networks
. committee machines (bagging, boosting)
. deep learning models (CNNs, DBNs, DAEs, DDNs, etc.)