Chapter 1
Introduction
(Summary)

Goals of AI

- Replicate human intelligence

  "AI is the study of complex information processing problems that often have their roots in some aspect of biological information processing. The goal of the subject is to identify solvable and interesting information processing problems, and solve them." -- David Marr

- Solve knowledge-intensive tasks

  "AI is the design, study and construction of computer programs that behave intelligently." -- Tom Dean

- Intelligent connection of perception and action
- Enhance human-human, human-computer and computer-computer interaction/communication

  Computer can sense and recognize its users, see and recognize its environment, respond visually and audibly to stimuli. New paradigms for interacting productively with computers using speech, vision, natural language, 3D virtual reality, 3D displays, more natural and powerful user interfaces, etc. (See, for example, projects in Microsoft's "Advanced Interactivity and Intelligence" group.)

Design Methodology and Goals

- **Engineering Goal**: Develop concepts, theory and practice of building intelligent machines
- **Science Goal**: Develop concepts, mechanisms and vocabulary to understand biological intelligent behavior

Alternatively, methodologies can be defined by choosing (1) the goals of the computational model, and (2) the basis for evaluating performance of the system:

1 Think like humans "cognitive science" Ex. GPS	**2** Think rationally => formalize inference process "laws of thought"
3 Act like humans Ex. ELIZA Turing Test	**4** Act rationally "satisficing" methods

- **Box 1**
  "Cognitive science" approach - Focus not just on behavior and I/O, look at reasoning process. Computational model should reflect "how" results were obtained. GPS (General Problem Solver): Goal not just to produce humanlike behavior (like ELIZA), but to produce a sequence of steps of the reasoning process that was similar to the steps followed by a person in solving the same task.

- **Box 2**
  The goal is to formalize the reasoning process, producing a system that contains logical inference mechanisms that are provably correct, and guarantee finding an optimal solution. This brings up the question: How do we represent information that will allow us to do inferences like the following one? "Socrates is a man. All men are mortal. Therefore Socrates is mortal." -- Aristotle

  Problem : Not all problems can be solved just by reasoning and inferences.

- **Box 3**
  Behaviorist approach. Not interested in how you get results, just the similarity to what human results are. Exemplified by the Turing Test (Alan Turing, 1950): 3 rooms contain a person, a computer, and an interrogator. The interrogator can communicate with the other 2 by teleprinter. The interrogator tries to determine which is the person and which is the machine. The machine tries to fool the

interrogator into believing that it is the person. If the machine succeeds than we conclude that the machine can think.

ELIZA: A program that simulated a psychotherapist interacting with a patient and successfully passed the Turing Test. Some extracts from the test:

Patient: Men are all alike.
ELIZA: In what way?
Patient They're always bugging us about something or other.
ELIZA: Can you think of a specific example?
Patient: Well, my boyfriend made me come here.
ELIZA: Your boyfriend made you come here?
Patient: He says I'm depressed much of the time.
ELIZA: I am sorry you are depressed.
...

Problems: ELIZA and other similar programs stressed simple syntactic analysis and generation of sentences. They used pattern matching with known sentences. They used pattern matching with known vocabulary and key words with templates of sentences to generate. For example,

if sentence = "* mother *" then respond with "Tell me about your family."

Note that even with simple syntactic style, ELIZA managed to fool people. Purely behavioral-based approach can be simulated without a deeper understanding or true "intelligence."

- **Box 4**
  For a given set of inputs, tries to generate an appropriate output that is not necessarily correct but gets the job done. Rational and sufficient ("satisficing" methods, not "optimal").

Most of AI work falls into Boxes 2 and 4. These don't rely on tests that correspond to human performance.

Symbols versus Signals

Most of AI built on an information processing model called a "Physical-Symbol System" (PSS) (Newell and Simon). Symbols usually correspond to objects in the environment. Symbols are physical patterns that can occur as components of an expression or symbol structure. A PSS is a collection of symbol structures plus processes that operate (i.e., create, modify, reproduce) expressions to produce other expressions. Hence, a PSS produces over time an evolving collection of symbol structures.

=> AI is the enterprise of constructing physical-symbol system that can reliably pass the Turing Test [or whatever your performance test is].

Physical-Symbol System Hypothesis (Newell and Simon, 1976): A physical-symbol system has the necessary and sufficient means for general intelligent action.

==> Intelligence is a functional property and is completely independent of any physical embodiment.

An alternative, less-symbolic paradigm: Neural Networks