

# Introduction to Artificial Intelligence

## Chapter 1

### Big Questions

- Can machines think?
- And if so, how?
- And if not, why not?
- And what does this say about human beings?
- And what does this say about the mind?

### What is AI?

- There are no crisp definitions
  - Here's one from John McCarthy, (He coined the phrase AI in 1956) - see <http://www.formal.Stanford.EDU/jmc/whatisai/>
- Q. What is artificial intelligence?*
- A. It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.
- Q. Yes, but what is intelligence?*
- A. Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.

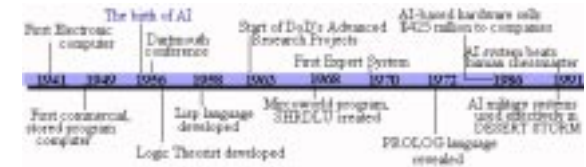
### Other possible AI definitions

- AI is a collection of hard problems which can be solved by humans and other living things, but for which we don't have good algorithms for solving.
  - E.g., understanding spoken natural language, medical diagnosis, circuit design, etc.
- AI Problem + Sound theory = Engineering problem
- Some problems used to be thought of as AI but are now considered not
  - e.g., compiling Fortran in 1955, symbolic mathematics in 1965
- AI is thus, by nature pre-scientific in Kuhn's terms

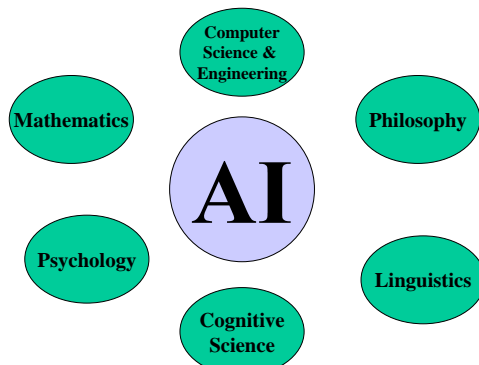
### What's easy and what's hard?

- It's been easier to mechanize many of the high level tasks we usually associate with "intelligence" in people
  - e.g., Symbolic integration, proving theorems, playing chess, medical diagnosis, etc.
- It's been very hard to mechanize tasks that lots of animals can do
  - walking around without running into things
  - catching prey and avoiding predators
  - interpreting complex sensory information (e.g., visual, aural, ...)
  - modeling the internal states of other animals from their behavior
  - working as a team (e.g. with pack animals)
- Is there a fundamental difference between the two categories?

### History



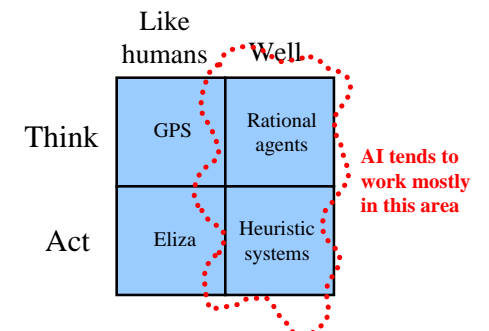
### Foundations of AI



### Why AI?

- **Engineering:** To get machines to do a wider variety of useful things
  - e.g., understand spoken natural language, recognize individual people in visual scenes, find the best travel plan for your vacation, etc.
- **Cognitive Science:** As a way to understand how natural minds and mental phenomena work
  - e.g., visual perception, memory, learning, language, etc.
- **Philosophy:** As a way to explore some basic and interesting (and important) philosophical questions
  - e.g., the mind body problem, what is consciousness, etc.

### Possible Approaches



## Think well

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

- Develop formal models of knowledge representation, reasoning, learning, memory, problem solving, that can be rendered in algorithms.
- There is often an emphasis on a systems that are provably correct, and guarantee finding an optimal solution.

## Act well

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

- For a given set of inputs generate an appropriate output that is not necessarily correct but gets the job done.
- A **heuristic (heuristic rule, heuristic method)** is a rule of thumb, strategy, trick, simplification, or any other kind of device which drastically limits search for solutions in large problem spaces. Heuristics do not guarantee optimal solutions; in fact, they do not guarantee any solution at all: **all that can be said for a useful heuristic is that it offers solutions which are good enough most of the time.**  
-- Feigenbaum and Feldman, 1963, p. 6

## Think like humans

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

- Cognitive science approach
- Focus not just on behavior and I/O but also look at reasoning process.
- Computational model should reflect "how" results were obtained.
- Provide a new language for expressing cognitive theories and new mechanisms for evaluating them
- 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.

## Act like humans

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

- 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).

## Turing Test

- 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 then we conclude that the machine can think.

## Eliza

- ELIZA: A program that simulated a psychotherapist interacting with a patient and successfully passed the Turing Test.
- Coded at MIT during 1964-1966 by Joel Weizenbaum.
- First script was DOCTOR.
  - The script was a simple collection of syntactic patterns not unlike regular expressions
  - each pattern had an associated reply which might include bits of the input (after simple transformations (my -> your))
- Weizenbaum was shocked at reactions:
  - Psychiatrists thought it had potential.
  - People unequivocally anthropomorphised.
  - Many thought it solved the NL problem.

## Eliza

- This is a transcript of a conversation with an Eliza like program found in Emacs (Gnu 20.2.2)
- Try typing `$x doctor` in your favorite Emacs.

```
E: I am the psychotherapist. Please, describe your problems.
Each time you are finished talking, type RET twice.
P: All men are alike
E: Why do you say all men are alike?
P: They're always bugging us about something or other.
E: Is it because of your plans that you say they're always
bugging us about something or other?
P: My boyfriend made me come here
E: Maybe your life have something to do with this.
P: He thinks I am depressed
E: Are you depressed often?
P: You are asking me a lot of questions
E: What makes you believe I am asking you a lot of
questions?
P: Because you seem very nosy
E: Is it because I seem very nosy that you came to me?
P: I am leaving.
E: Earlier you said you felt depressed?
P: Good bye.
E: My secretary will send you a bill.
```

## Colby's PARRY

- Kenneth Colby modeled a paranoid using the same techniques circa 1968.
- PARRY has basic emotions. If it gets angry, it's replies become more hostile.
- In a Turing test, expert psychiatrists were unable to distinguish PARRY's ramblings from those of real paranoids.

```
H: Good afternoon. My name is
Doctor Jones.
P: Pleased to meet you.
H: How are you?
P: All right I guess.
H: Could you tell me your name?
P: Pat Smith.
H: I though you might have some
problems.
P: I've been upset lately.
H: By what?
P: People get on my nerves
sometimes.
H: Tell me more.
P: Do you know anything about
bookies?
...
```

## Parry meets Eliza

```
P: I don't understand your motives.
E: Why don't you understand my motives?
P: Shouldn't I be?
E: Why do you say that just now?
P: That's just the way it is.
E: Suppose that's the way it were not.
P: So?
E: I am not sure I understand you fully.
P: You harp on the same questions too much.
E: What does that similarity suggest to you?
P: I don't get you.
E: Do you wish that you did get I?
```

## The Loebner Contest

- A modern version of the turning test with \$100,000 cash prize.
- Hugh Loebner was once director of UMBC's Academic Computing Services (nee UCS)
- <http://www.loebner.net/Prizef/loebner-prize.html>
- Restricted topic (removed in 1995) and limited time.
- Participants include a set of humans and a set of computers and a set of judges.
- Scoring
  - Rank from least human to most human.
  - Highest median rank wins \$2000.
  - If better than a human, win \$100,000.

## What can AI systems do

### Here are some example applications

- **Computer vision:** face recognition from a large set
- **Robotics:** autonomous car
- **Natural Language Processing:** simple machine translation
- **Expert Systems:** medical diagnosis in a narrow domain
- **Spoken language systems:** ~1000 word continuous speech
- **Planning and scheduling:** scheduling Hubbell experiments
- **Time series predications:**
- **Learning:** text categorization into ~1000 topics
- **User modeling:** Baysean reasoning in windows help
- **Games:** grandmaster level in chess, checkers, etc.

## What can't AI systems do yet?

- Robust natural language understanding (e.g. read and understand articles in a newspaper)
- Arbitrary visual scene understanding
- Learning a natural language
- Play Go well