CMSC 671
Fall 2015

Tim Finin, finin@umbc.edu

What is AI?

What is AI?

- 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.

http://www-formal.stanford.edu/jmc/whatisai/

Ok, so what is intelligence?

- 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.

http://www-formal.stanford.edu/jmc/whatisai/

A little bit of AI history

Ada Lovelace

- Babbage thought his machine was just a number cruncher
- Lovelace saw that numbers can represent other entities, enabling machines to reason about anything
- But: *"The Analytical Engine has no pretentions whatever to originate anything. It can do whatever we know how to order it to perform."*

AI prehistory and early years

- George Boole invented propositional logic (1847)
- Karel Capek coined the term “robot” (1921)
- John von Neumann: minimax (1928)
- Norbert Wiener founded field of cybernetics (1940s)
- Neural networks, 1940s and 1950s, among the earliest theories of how we might reproduce intelligence
- Isaac Asimov *I, Robot* (1950) Laws of Robotics
- Turing test, proposed in 1950 and debated ever since
- Early work on Chess By Turing

AI prehistory and early years

- Logic Theorist and GPS, 1950s, early symbolic AI
- Early years: focus on search, learning, knowledge representation
- Marvin Minsky: neural nets (1951), AI founder, blocks world, Society of Mind
- John McCarthy invented Lisp (1958) and coined the term AI (1957)
- Allen Newell, Herbert Simon: GPS (1957), AI founders
- Noam Chomsky: analytical approach to language (1950s)
- Dartmouth University summer conference, 1956, established AI as a discipline

1956 Dartmouth AI Project

Five of the attendees of the 1956 Dartmouth Summer Research Project on AI reunited in 2006: Trenchard More, John McCarthy, Marvin Minsky, Oliver Selfridge, and Ray Solomonoff. Missing were: Arthur Samuel, Herbert Simon, Allen Newell, Nathaniel Rochester and Claude Shannon.

1956 Dartmouth AI Project

“We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.”

http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

1956 Dartmouth AI Project

"We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."

http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Recent AI History

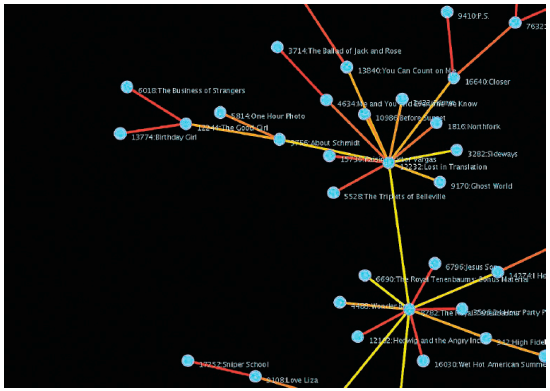
- AI has had it's ups and downs
  - 50-60 up, 70s down, 80s up, 90s down, 00s up, 10s up
- Hot topics today?
  - Machine learning, datamining, etc.
  - Neural networks again: deep learning
  - Exploiting big data
  - Autonomous vehicles, robotics
  - Text mining, natural language technology, speech
  - Computer vision

DARPA Grand Challenge

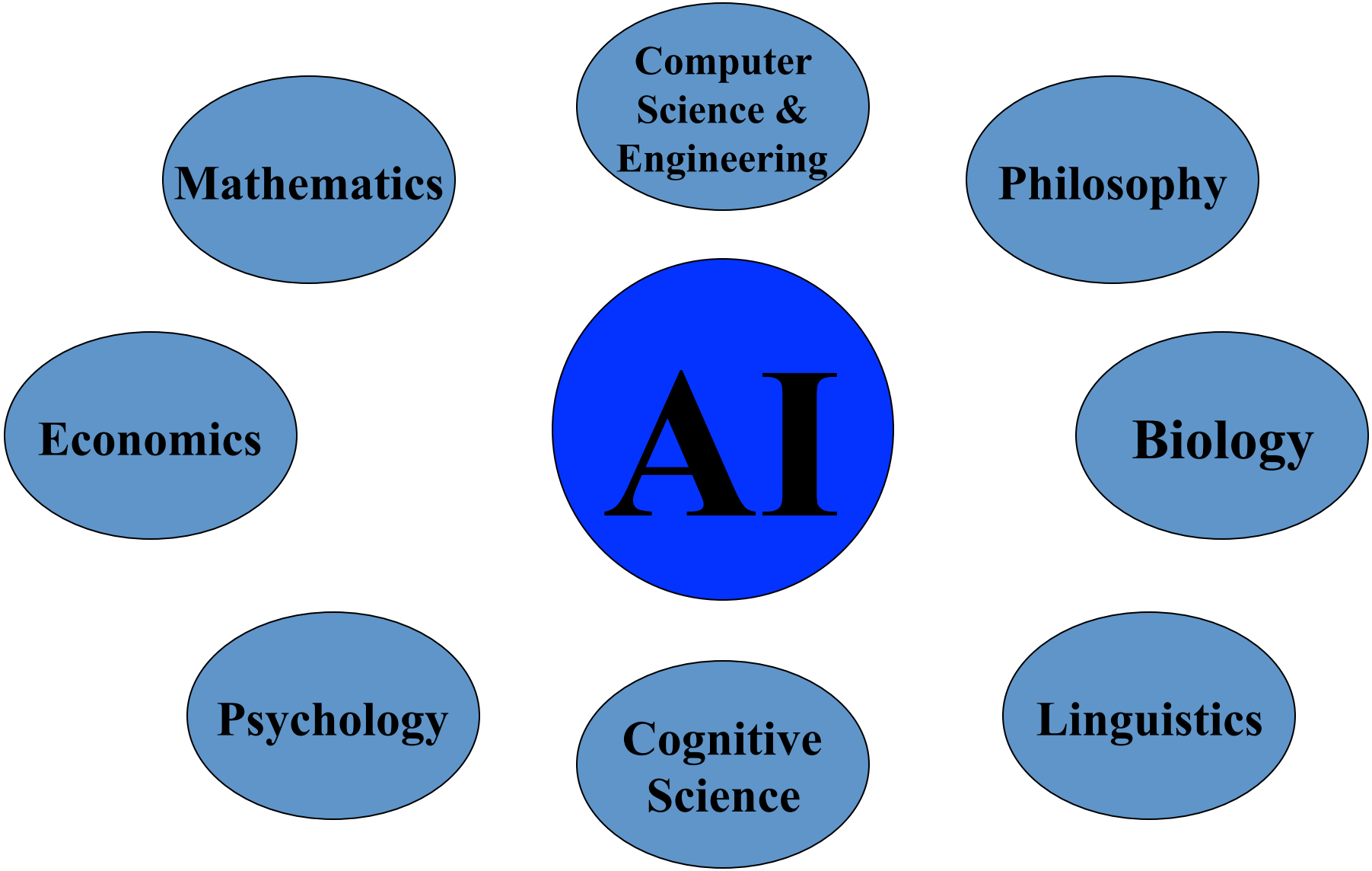
- A series of DARPA sponsored prize competition for autonomous vehicles
- Run in 2004, 2005, 2007
- $1M prize in each year
- See Wikipedia article and Wired video

Netflix Prize

- **Recommendation systems** use machine learning to predict preferences
- In 2006 Netflix offered a $1M prize for a 10% improvement in their recommendation system
- In 2009 the prize was won as two ‘ensemble’ teams exceeded the threshold in the same month



Foundations of AI



Big questions

- Can machines think?
- If so, how?
- If not, why not?
- What does this say about human beings?
- What does this say about the mind?

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.

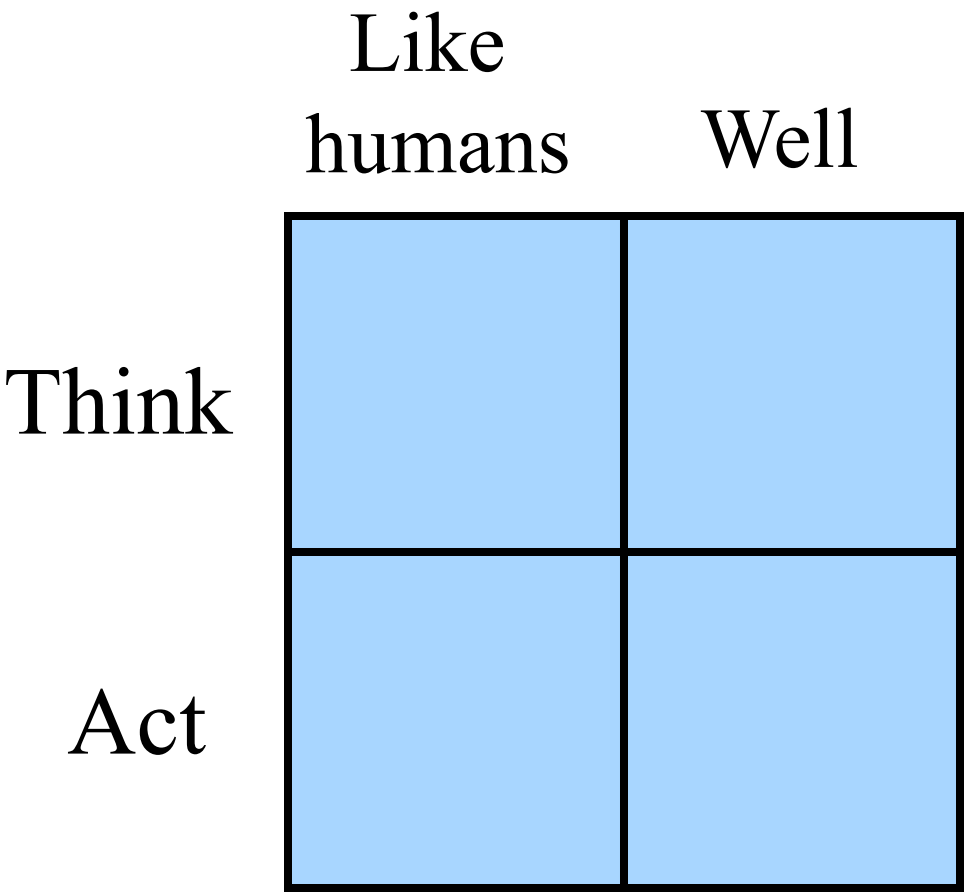
- **Cognitive Science:** As a way to understand how natural minds and mental phenomena work

  e.g., visual perception, memory, learning, language, decision making, etc.

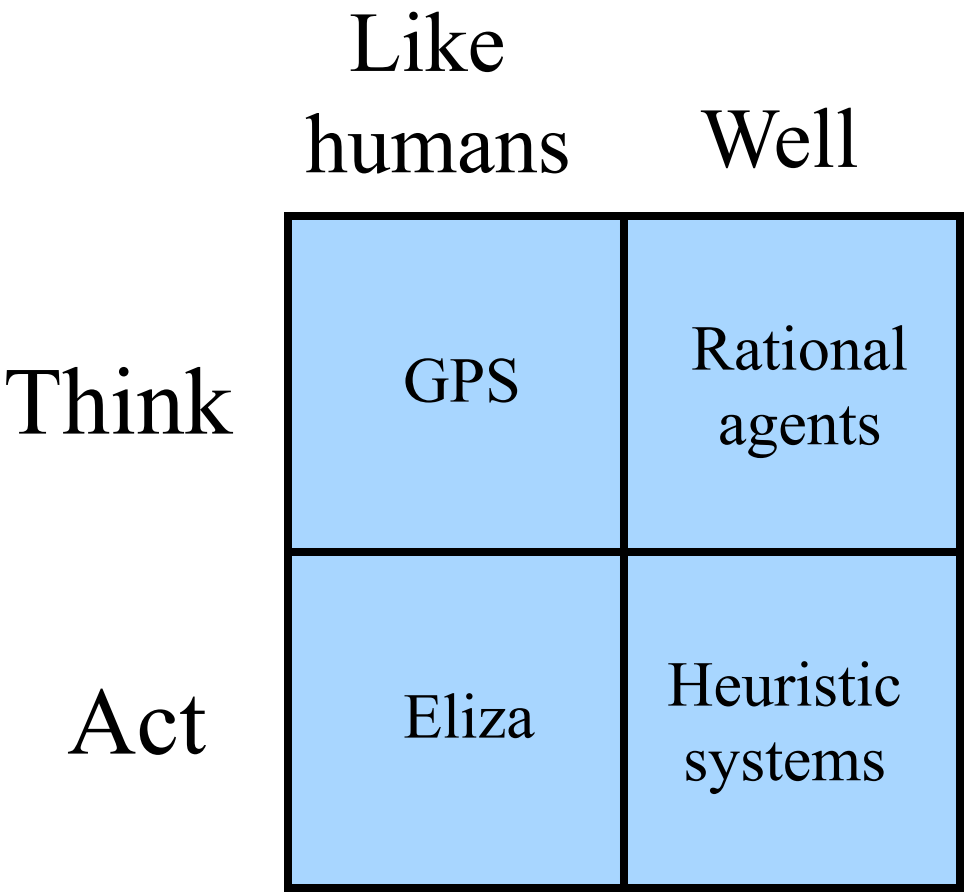
- **Philosophy:** As a way to explore some basic and interesting (and important) philosophical questions

  e.g., the mind body problem, what is consciousness, etc.

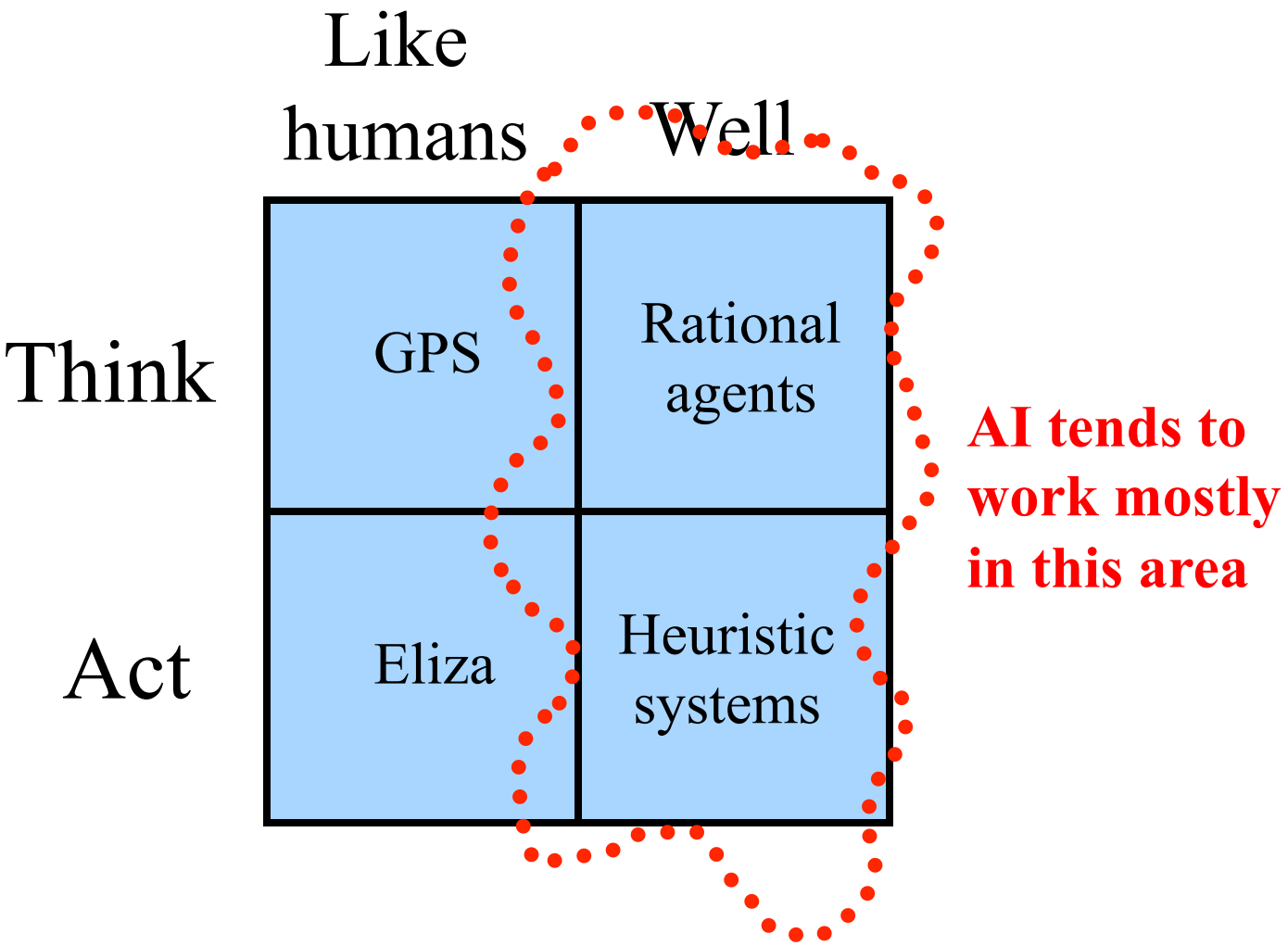
Possible approaches



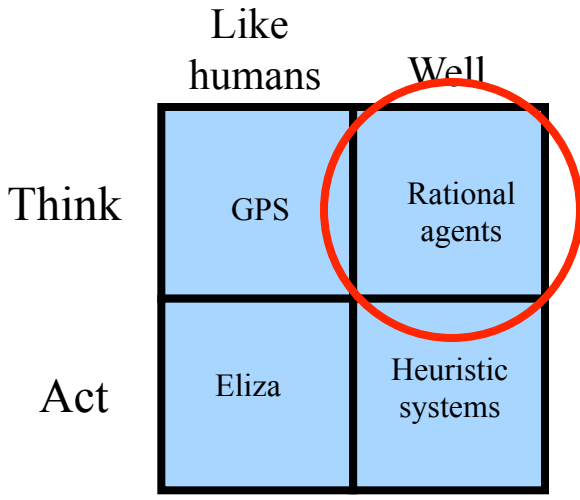
Possible approaches



Possible approaches

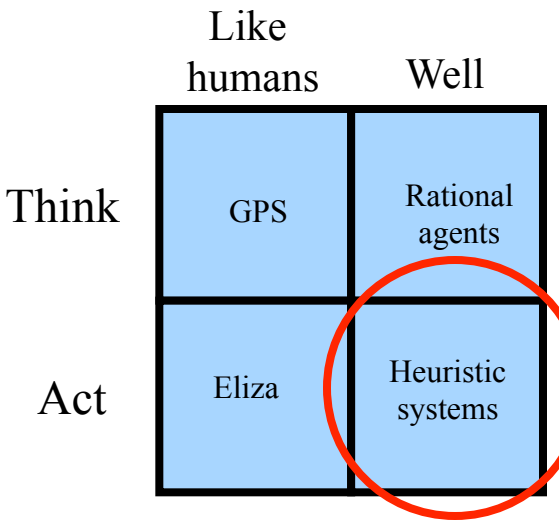


Think well



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

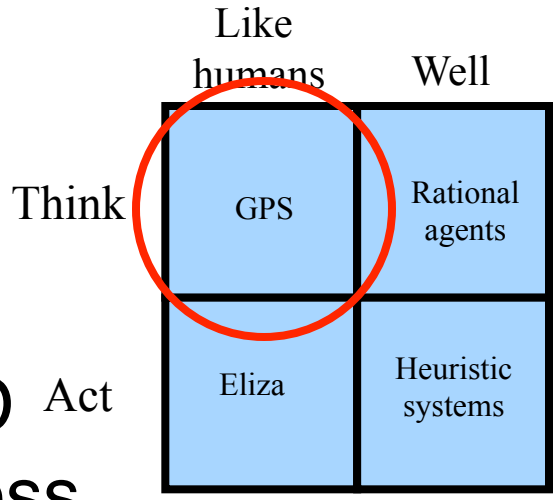
Act well



- For a given set of inputs, generate output that's not necessarily correct but gets job done
- A **heuristic** (heuristic rule, heuristic method) is a rule of thumb, strategy, trick or simplification which drastically limits search for solutions in large problem spaces
- Heuristics don't guarantee optimal solutions or even 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**

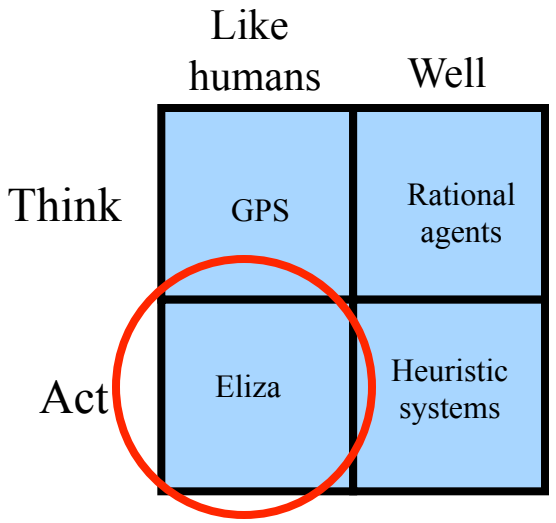
–Feigenbaum and Feldman, 1963, p. 6

Think like humans



- Cognitive science approach
- Focus not just on behavior & I/O but also look at reasoning process
- Computational model should reflect "how" results were obtained
- Provides new language for expressing cognitive theories & new mechanisms for evaluating them
- GPS (**General Problem Solver**): Goal not just to produce humanlike behavior, but to produce a sequence of steps of reasoning process that was similar to those followed by a person

Act like humans



- Behaviorist approach
- Not interested in how you get results, just similarity to what human results are
- Exemplified by the Turing Test (Alan Turing, 1950)
- Has applications in interactive entertainment (e.g., computer games, CGI), virtual worlds and in modeling human intentions

What's easy and what's hard?

- Easy: many high-level tasks usually associated with "intelligence" in people
  - e.g., symbolic integration, proving theorems, playing chess, medical diagnosis
- Hard: tasks many animals can do
  - walking around without running into things
  - catching prey and avoiding predators
  - Interpreting sensory info. (e.g., visual, aural, …)
  - modeling internal states of other from behavior
  - working as a team (e.g., with pack animals)
- Is there a fundamental difference between these?

Turing Test

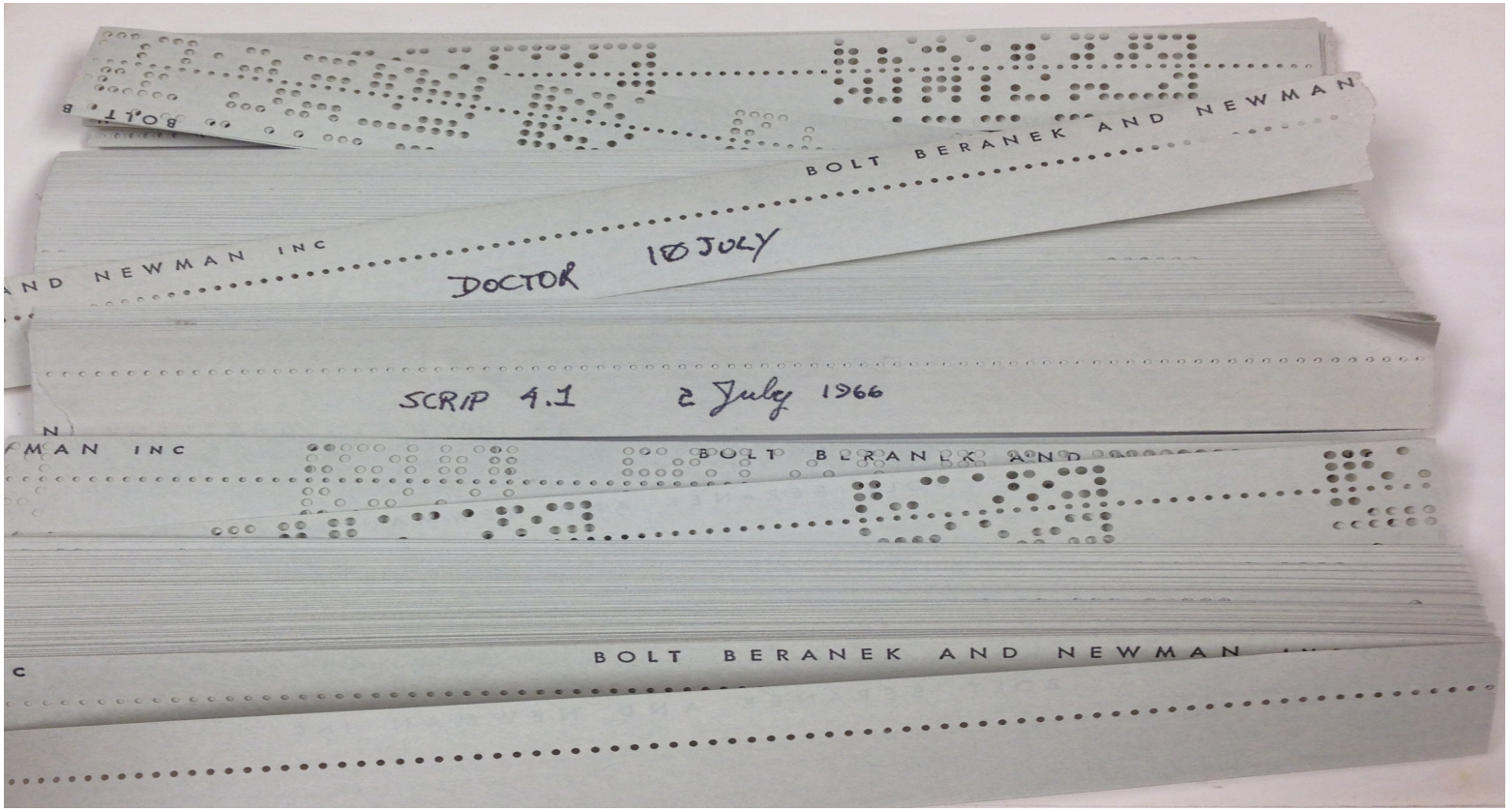
- Three rooms contain a person, a computer, and an interrogator
- The interrogator can communicate with the other two by ‘teleprinter’
- Interrogator tries to determine which is the person and which is the machine
- The machine tries to fool the interrogator into believing that it is the person
- If the machine succeeds, then we conclude that the machine can think

Eliza

- ELIZA: program simulating a psychotherapist interacting with a patient
  - Joe Weizenbaum (MIT), ~1966, FORTRAN
- First script was DOCTOR
  - A simple collection of syntactic patterns not unlike regular expressions
  - Patterns had associated replies that might include bits of the input after simple transforms (my → your)
- Weizenbaum shocked at reactions:
  - Psychiatrists thought it had potential
  - People unequivocally anthropomorphized
  - Many thought it solved the NL problem

It lives!

Amazingly, a 1966 version was found and restored, see http://elizagen.org/



Eliza

This is a transcript of a conversation with an Eliza-like program found in Emacs

Try typing

Meta-x doctor

in your favorite Emacs

- I am the psychotherapist. Please, describe your problems. Each time you are finished talking, type RET twice.
- *i'm having a hard time with my students*
- Why do you say that?
- well, they keep asking me for money
- Maybe your plans have something to do with this.
- *what plans?*
- Is it because of your life that you say what plans?
- *No, it's because I didn't know what you were talking about*
- Is it because you didn't know what I was talking about that *you came to me?*
- *no, it's because i wanted to see if you could pass the turing test*
- Does it bother you that you wanted to see if I could really pass the turing test?
- …..
- How do you reconcile problems at school?
- *This is a waste of time. Good bye*
- My secretary will send you a bill.

Colby's PARRY

- Kenneth Colby modeled a paranoid using the same techniques in 1968
- PARRY has basic emotions. If it gets angry, its 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

Please tell me *this* isn't AI!

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 Turing Test, held annually, with a $100,000 cash prize.
- Hugh Loebner was once director of UMBC's Academic Computing Services (née UCS)
- http://www.loebner.net/Prizef/loebner-prize.html
- Restricted topic (removed in '95) 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 (Nobody yet…)

What can AI systems do?

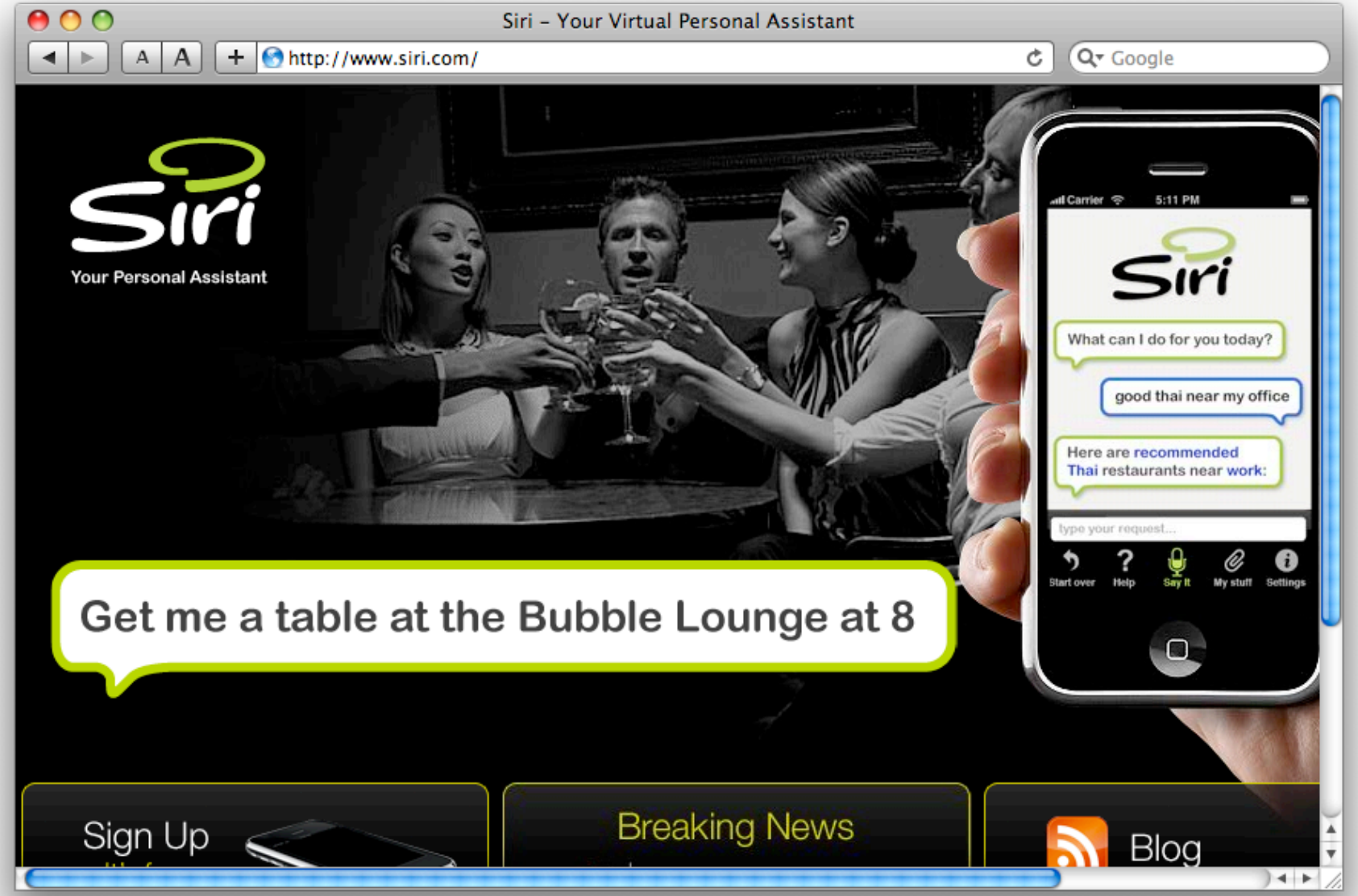
- **Computer vision:** face recognition from a large set
- **Robotics:** autonomous (mostly) automobile
- **Natural language processing:** simple machine translation
- **Expert systems:** medical diagnosis in narrow domains
- **Spoken language systems:** e.g., Google Voice, Siri
- **Planning and scheduling:** Hubble Telescope experiments
- **Learning:** text categorization into ~1000 topics
- **User modeling:** Bayesian reasoning in Windows help (the infamous paper clip…)
- **Games:** Grand Master level in chess (world champion), checkers, etc.

What can't AI systems do yet?

- Understand natural language robustly (e.g., read and understand articles in a newspaper)
- Surf the web and find interesting knowledge
- Interpret an arbitrary visual scene
- Learn a natural language
- Play Go well
- Construct plans in dynamic real-time domains
- Refocus attention in complex environments
- Perform life-long learning

Exhibit true autonomy and intelligence!

http://siri.com/



Are we there yet?

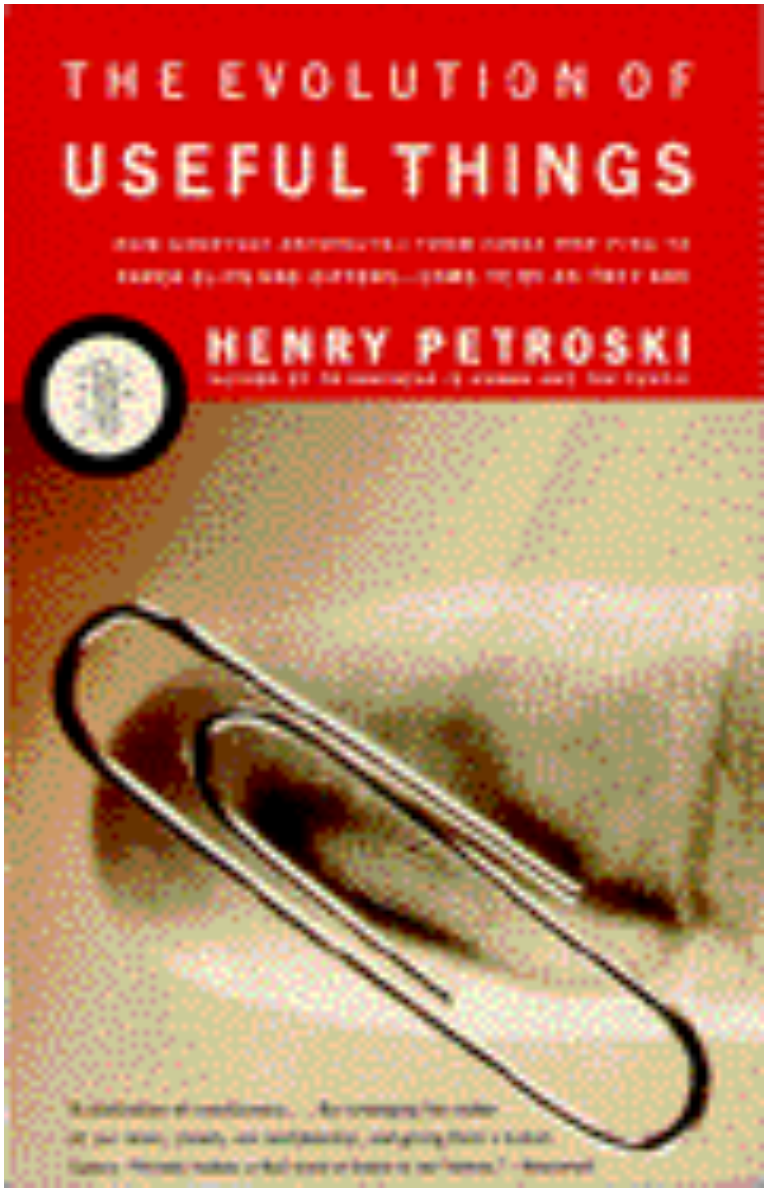
- Great strides have been made in knowledge representation and decision making
- Many successful applications have been deployed to (help) solve specific problems
- Key open areas remain:
  - Incorporating uncertain reasoning
  - Real-time deliberation and action
  - Perception (including language) and action (including speech)
  - Lifelong learning / knowledge acquisition
  - Common-sense knowledge
  - Methodologies for evaluating intelligent systems

T.T.T

Put up in a place
where it's easy to see
the cryptic admonishment
T. T. T.
When you feel how depressingly
slowly you climb,
it's well to remember that
Things Take Time.

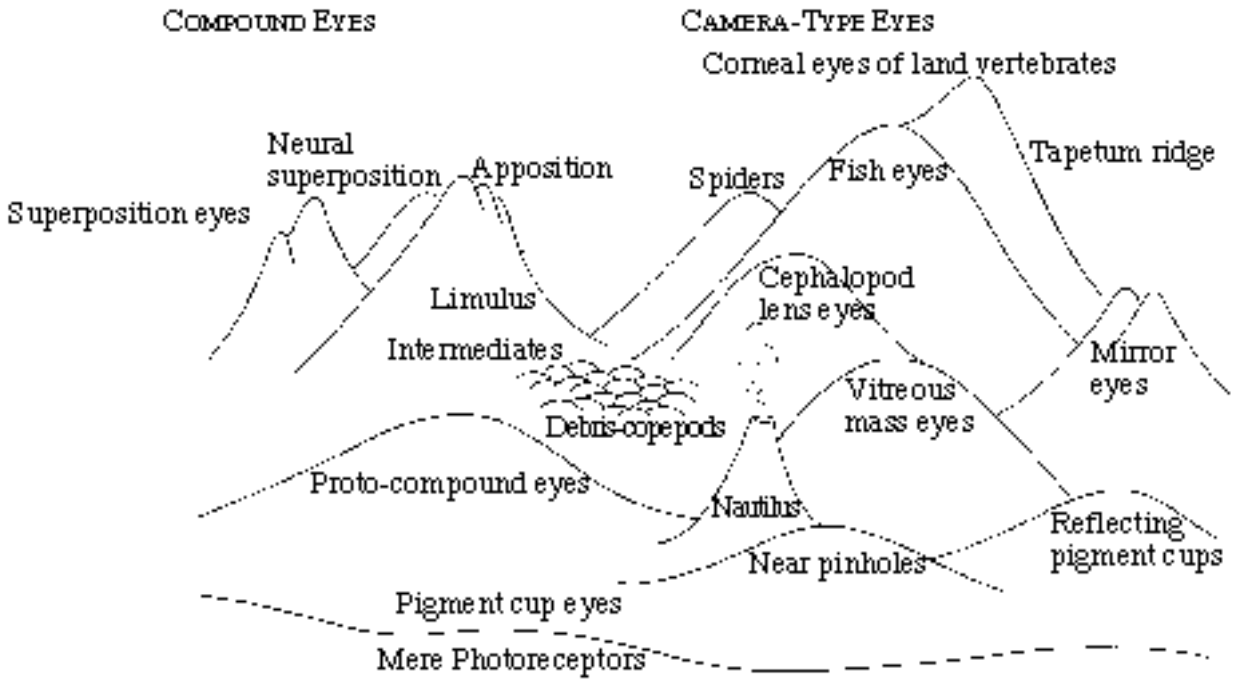
-- Piet Hein

T.T.T: things take time



- Prior to the 1890's, papers were held together with straight pens.
- The development of "spring steel" allowed the invention of the paper clip in 1899.
- It took about 25 years (!) for the evolution of the modern "gem paperclip", considered to be optimal for general use.

Climbing Mount Improbable



"The sheer height of the peak doesn't matter, so long as you don't try to scale it in a single bound. Locate the mildly sloping path and, if you have unlimited time, the ascent is only as formidable as the next step."

-- Richard Dawkins, *Climbing Mount Improbable*, Penguin Books, 1996.