

Plan for lecture: Artificial Intelligence, Turing, and Dreyfus

- The very idea of machines thinking. Can a computer/machine think or behave intelligently? Can anything think or behave intelligently simply by virtue of being a computer?
- Potted history of Artificial Intelligence and related cognitive sciences
- Alan Turing and Turing test
- Dreyfus' arguments against GOF AI (Good Old-Fashioned AI) in *What Computers Still Can't Do*

Turing test

Alan Turing, British, 'Computing Machinery and Intelligence', 1950

Instead of asking can machines think, we should ask whether they can pass a behavioral intelligence test.

Turing test: You converse via typed messages (e.g., a chat window) for 5 minutes. Then you guess whether you are talking to a person or a machine.

A possible criterion for AI = computer can fool 70% of the people for 5 minutes.

Objections to Turing Test

1. It is too hard (chimps can't pass it)
2. It is purely verbal and does not relate word to world (is verbal quizzing an adequate test of intelligence?)
3. Black box objection (we are looking only at outputs, behaviour processes, not inner workings, and perhaps this is insufficient for establishing existence of thought)

Three problems for AI?

Knowledge problem – How to organise vast quantities of diverse knowledge (e.g. millions of encyclopedias, and their background assumptions) such that computer can have our common sense?

Frame problem – What methods will ensure that a new bit of knowledge flows through into the vast array of connected information that now needs updating? Real-world knowledge is consistently changing.

Relevance problem – Which information and facts are relevant to solving a particular problem? How to retrieve relevant parts?

Recommended reading

Copeland, J., *The Philosophy of Artificial Intelligence* (on reserve)

Crane, T., *The Mechanical Mind*, chapter 3 and 4 (on reserve)

Stanford Encyclopedia of Philosophy
(entries on Turing Test, Cognitive Science, Computational Mind, AI, etc.)

Internet Encyclopedia of Philosophy (entries on Turing, Cognitive Science, AI, etc.)

Various links via LMS

Hubert Dreyfus

- *Mind over machine*, with Stuart Dreyfus, 1986 (on reserve)
- *What computers still can't do*, 1992, (on reserve)
- Various links via LMS

Some of his papers are online:

<http://socrates.berkeley.edu/~hdreyfus/html/papers.html>

Inspired by: Heidegger's, *Being and Time* and Merleau-Ponty's, *Phenomenology of Perception*



Dreyfus' Five Stages of Skill Acquisition

NOVICE

ADVANCED BEGINNER

COMPETENT

PROFICIENT

EXPERT

Novice

- * The novice follows rules
- * Specific rules for specific circumstances
- * No modifiers
- * “Context free”
- * Don't feel responsible for other than following the rule

Advanced Beginner

- New “situational” elements are identified
- Rules begin to be applied to related conditions
- Decisions still are made by rule application
- Does not experience personal responsibility

Competence

- Numbers of rules becomes excessive
- Learn organizing principles or “perspectives”
- Perspectives permit assorting information by relevance
- The experience of responsibility arises from active decision-making

Proficiency

- Intuitive diagnosis
- Approach to problem molded by perspective arising from multiple real world experiences
- “Holistic similarity recognition”
- Learner uses intuition to realize “what” is happening
- Conscious decision-making and rules used to formulate plan

Expertise

- Don't make decisions
- Don't solve problems
- Do what works
- No decomposition of situation into discrete elements
- Pattern recognition extends to plan as well as diagnosis
- know-how (skills and abilities) rather than know-that (facts, rules)

- “One has to abandon the traditional view that a beginner starts with specific cases and, as he becomes more proficient, abstracts and interiorises more and more sophisticated rules. It turned out that skill acquisition moves in just the opposite direction – from abstract rules to particular cases” (Dreyfus and Dreyfus, *Mind Over Machine*, p108).

- “Current AI is based on the idea, prominent in philosophy since Descartes, that all understanding consists in forming and using appropriate representations. Given the nature of inference engines, AI’s representations must be formal ones, and so common-sense understanding must be understood as some vast body of precise propositions, beliefs, rules, and procedures. Thus formulated, the problem has so far resisted solution. We predict it will continue to do so” (Dreyfus and Dreyfus, *Mind Over Machine*, p99).

Dreyfus identified several basic assumptions about the nature of human knowledge which grounded cognitive science.

- A belief that the mind functions like a digital computer using symbolic manipulations – **psychological assumption** (163ff)
- A belief that computer programs could be understood as formalising human thought – **epistemological assumption** (189)
- An assumption about the data about the human world that we employ in thought, namely that it consists of discrete, determinate, and explicit pieces that can be processed heuristically – **ontological assumption** (206). <http://ejap.louisiana.edu/EJAP/1996.spring/wrathall.kelly.1996.spring.html>

Varieties of the letter 'A'



Some urls

- <http://www.theage.com.au/world/science/can-machines--201think20623-20urj.html>
- <http://www.abc.net.au/tv/newinventors/txt/s3163754.htm>
- <http://www.latrobe.edu.au/reccsi>
- <http://www.youtube.com/watch?v=e8vZy8a9lSc>
- http://socrates.berkeley.edu/%7Ehdreyfus/html/paper_socrates.html
- <http://www.youtube.com/watch?v=99iTDUcBuRQ>
- <http://www.youtube.com/watch?v=jhrGTrj4DOI&feature=relmfu>