The questions of Artificial Intelligence

What do we mean by intelligence?

Is it theoretically possible
for a computer system to be intelligent?

Plato, Socrates

.
.
.
.
.
.
.
.
.
.
.
.
.

Galileo

Descartes

Leibniz

Kant

Boole

Cartesian approach

- Western culture
- Desire to explain *knowing how* in terms of *knowing that*
- Reduce skills to **facts & rules**
- To explain the *composite* in terms of its *constituents*

Plato, Socrates

.

.

.

Galileo

Descartes

Leibniz

Kant

Boole

Turing

Newell & Simon

McCarthy, Minsky

AI Manifesto 1976

"We are concerned with developing

a formalism, or *representation*,

with which to describe ... **knowledge**."

-- Terry Winograd

Plato, Socrates
.
.
.
Galileo
Descartes
Leibniz
Kant
Boole

Blaise Pascal

AI{

Turing
Newell & Simon
McCarthy, Minsky

Peter Naur
Dreyfus

Blaise Pascal 1640

“Mathematicians wish to treat matters of perception mathematically, and make themselves ridiculous … the mind … does it tacitly, naturally, and without technical rules.”

Logic Theorist 1956

- Newell, Simon, Shaw
- To prove theorems in ‘Principia Mathematica’ *(propositional logic)*
- Axioms $\wedge$ modus ponens

Logic Theorist 1956

- Newell, Simon, Shaw
- To prove theorems in ‘Principia Mathematica’ *(propositional logic)*
- Axioms $\wedge$ modus ponens

- General Problem Solver
- McCarthy & Minsky => Dartmouth ‘*artificial intelligence*’

Main aim of AI

- Not to build an artificial person/mind
- To build computer programs that, when run on a digital computer, exhibit intelligence

"Since computers are nothing more than symbol-manipulating devices, it is clear that if we want to program a computer to perform logical reasoning, we have to express all the knowledge required in a symbolic form and arrange matters so that logical reasoning corresponds to various of those symbols."

carry out the deduction the way we did. The problem is step 3. This uses the fact that if someone dies, then they are dead and they remain dead at all later times. You and I know that, but computers (and occasionally writers of long-running television series) do not. The computer has to be told this additional fact of life; thus, the following fact has to be added to the program's list of basic data:

8. $\forall x \forall d \forall t[\text{Mortal}(x) \wedge \text{Die}(x,d) \wedge (t > d) \rightarrow \text{Dead}(x,t)]$

This last fact is particularly convoluted and indicates the kinds of linguistic contortions that are sometimes required in order to express seemingly simple facts in the highly restricted language of predicate logic.

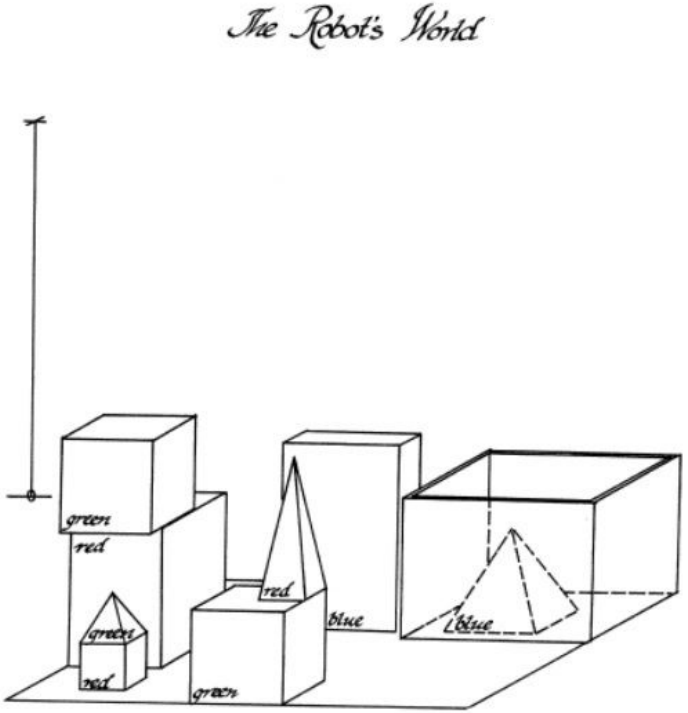
=> combinatorial explosion (Lighthall Debate)

“The reason the computer has such difficulty is that it is simply a symbol manipulator. Though sentences in predicate logic can be hard to understand (for an untrained person), it is nevertheless *possible* for us to read them & understand. ”

Terry Winograd

Deliberate intention that the computer should *understand* the input sentences.

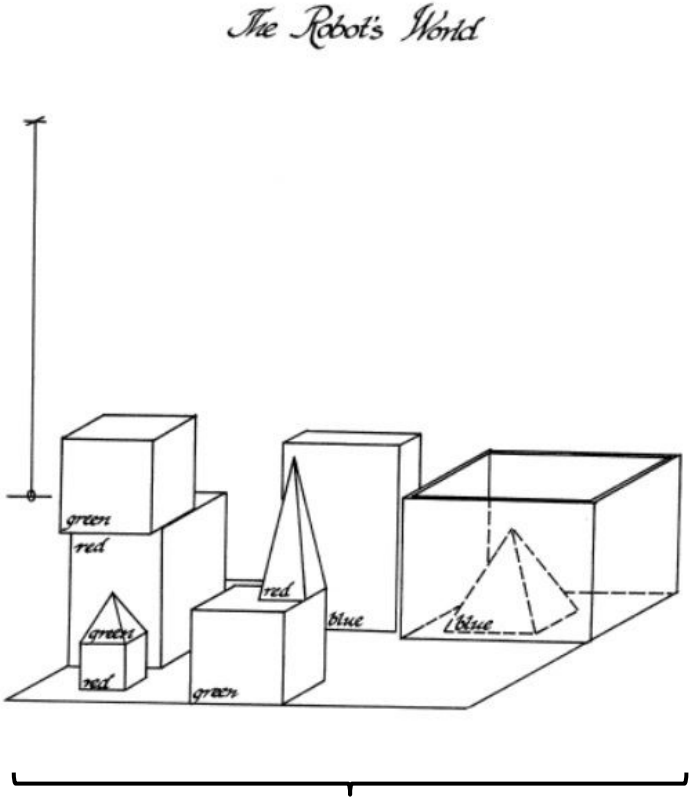
> Microworld



Terry Winograd

Deliberate intention that the computer should *understand* the input sentences.

> Microworld



Does not scale up

according to Dreyfus and Winograd himself

Crisis in AI

Minsky 1967

“AI solved within a generation.”

Minsky 1982

“The AI problem is one of the hardest science has ever undertaken.”

Common-sense knowledge problem

late 70's

e.g. "if you are dead, you stay dead"

Minsky

"We can program 100.000 common-sense knowledge facts

into a computer for it to work as an intelligent agent."

Patrick Hayes
Naive Physics Manifesto 1979

Attempt to write down, *in the language of predicate logic*,

all of our common-sense knowledge of the physical world.

- ➔ how solids & liquids behave
- ➔ the effects of gravity on people
- ➔ ...

Patrick Hayes
Naive Physics Manifesto 1979

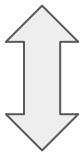
Attempt to write down, *in the language of predicate logic*,

all of our common-sense knowledge of the physical world.

But even if you are able to list enough of such basic facts...

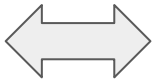
- *fridge* example
- Minsky: *frames*

Open-ended question whether it is *even theoretically* possible to achieve *intelligence* by the manipulation of symbols.



Expert Systems =/= Databases

- Knowledge base
- Inference engine
- User interface



Dreyfus & Dreyfus

5-stage model of human performance

(similar to Naur, Pascal, ...)

5-stage model of human performance

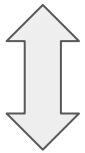
1. Novice
2. Advanced Beginner
3. Competence
4. Proficiency
5. Expert

5-stage model of human performance

1. Novice → **Follow rules**
2. Advanced Beginner → **Modifies some rules according to context**
3. Competence
4. Proficiency → **Trained reflex (*most of the time:* no rules)**
5. Expert → **Intuition, Instinct, following *no rules***

Pascal - Naur - Dreyfus

Human activity is not based on rules.

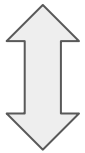


Logician’s rule-based view of human thought.

M. Davis - J. McCarthy - M. Minsky

Pascal - Naur - Dreyfus

Human activity is not based on rules.



K. Gödel?

Logician’s rule-based view of human thought.

M. Davis - J. McCarthy - M. Minsky