

Cognitive Science & Psychology: Mind, Brain and Behavior

Philosophical and historical roots I (Week 5)



Paul Verschure
csim.upf.edu
specs.upf.edu

Outline

Lecture 1 Introduction - robot future

Lecture 2 The Mind, Brain, behaviour Cycle

Lecture 3 The Knowledge Problem in the Science of Mind and Brain

Lecture 4 The Five Revolutions defining Current Reality

Lecture 5 Conceptual Revolutions in Philosophy of Mind

Lecture 6 (1850-1915) Structuralism and Functionalism

Lecture 7 (1915-1950) Behaviorism, Cognitive Behaviorism

Lecture 8 (1950-1960) The Demise of Behaviorism

Lecture 9 (1945-1960) Mind as Computation: AI. Cybernetics and the Cognitive Revolution

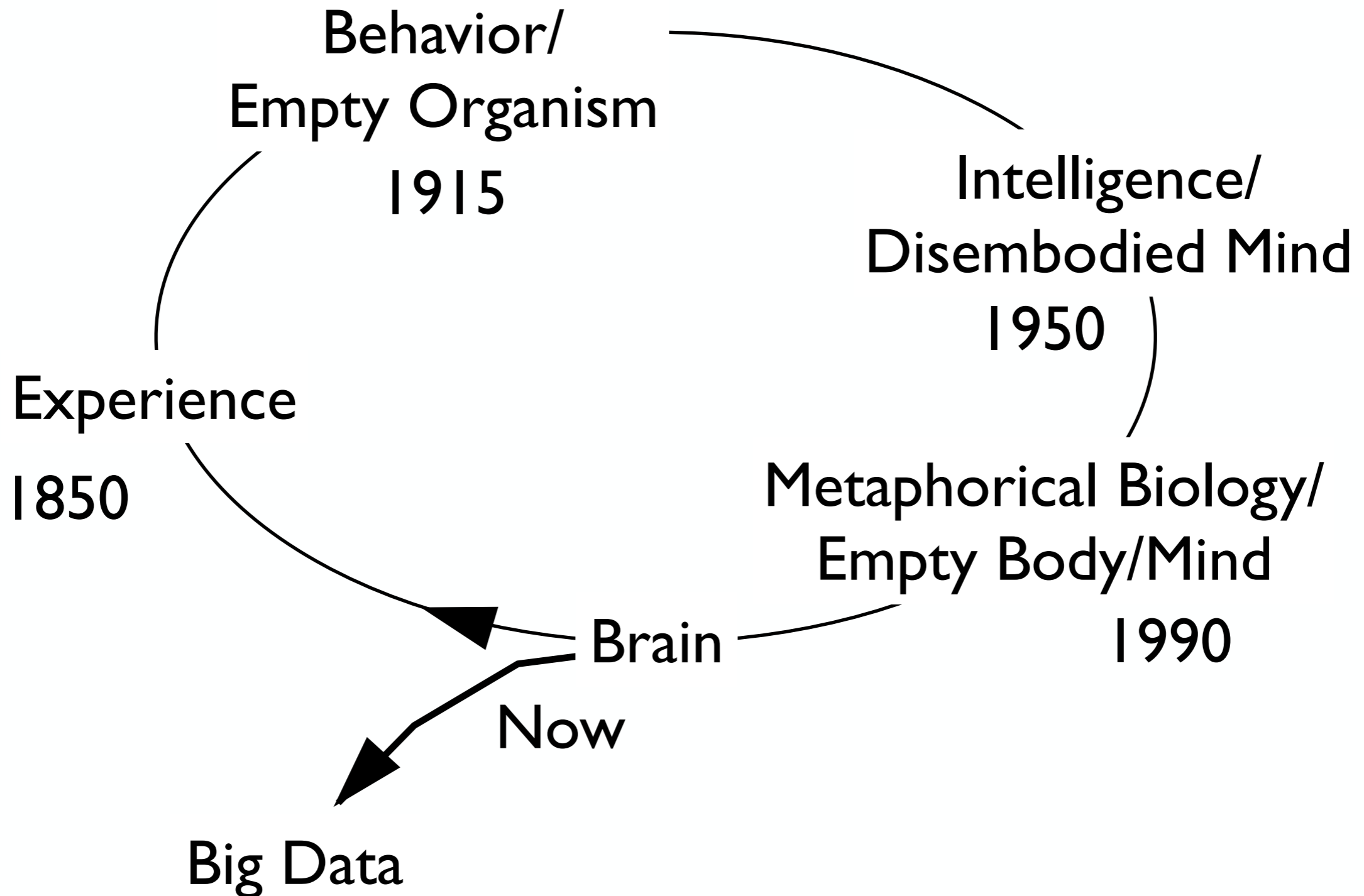
Lecture 10 (1985-Now) Biology as a Metaphor and Beyond

Lecture 11 (Now-Future) Flux and Synthesis

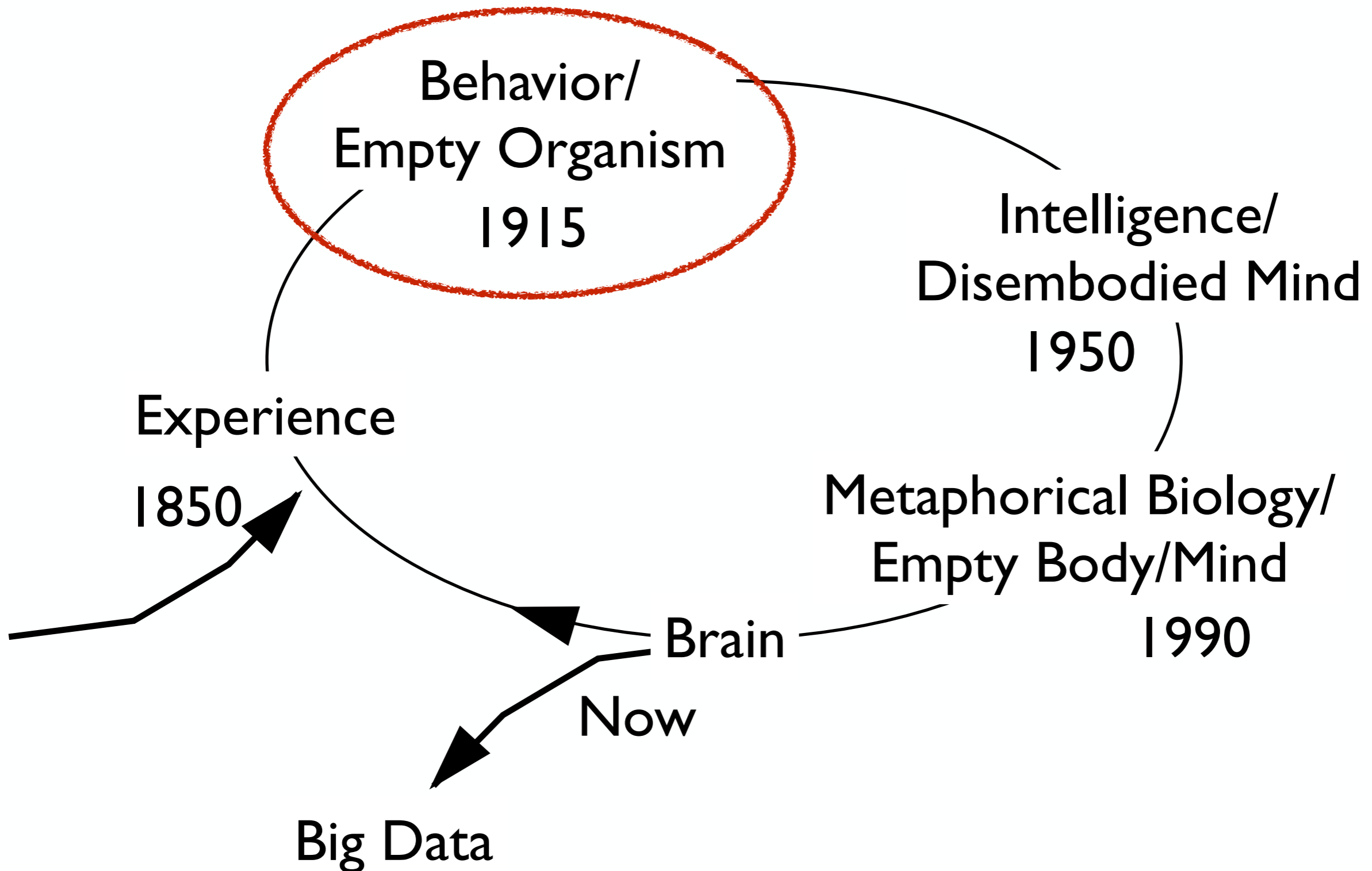


The Cognitive Revolution

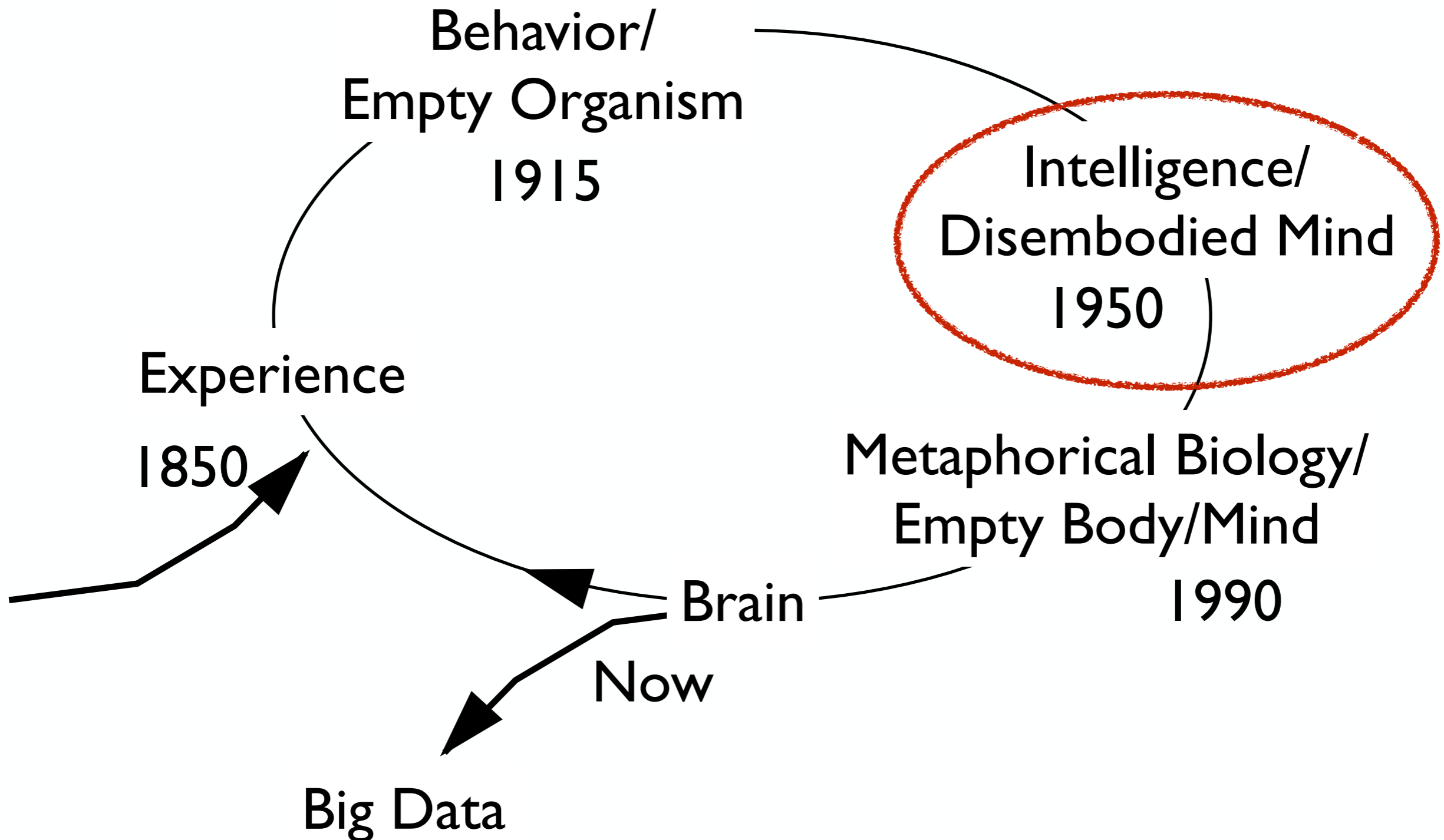
The mind/behavior/brain cycle



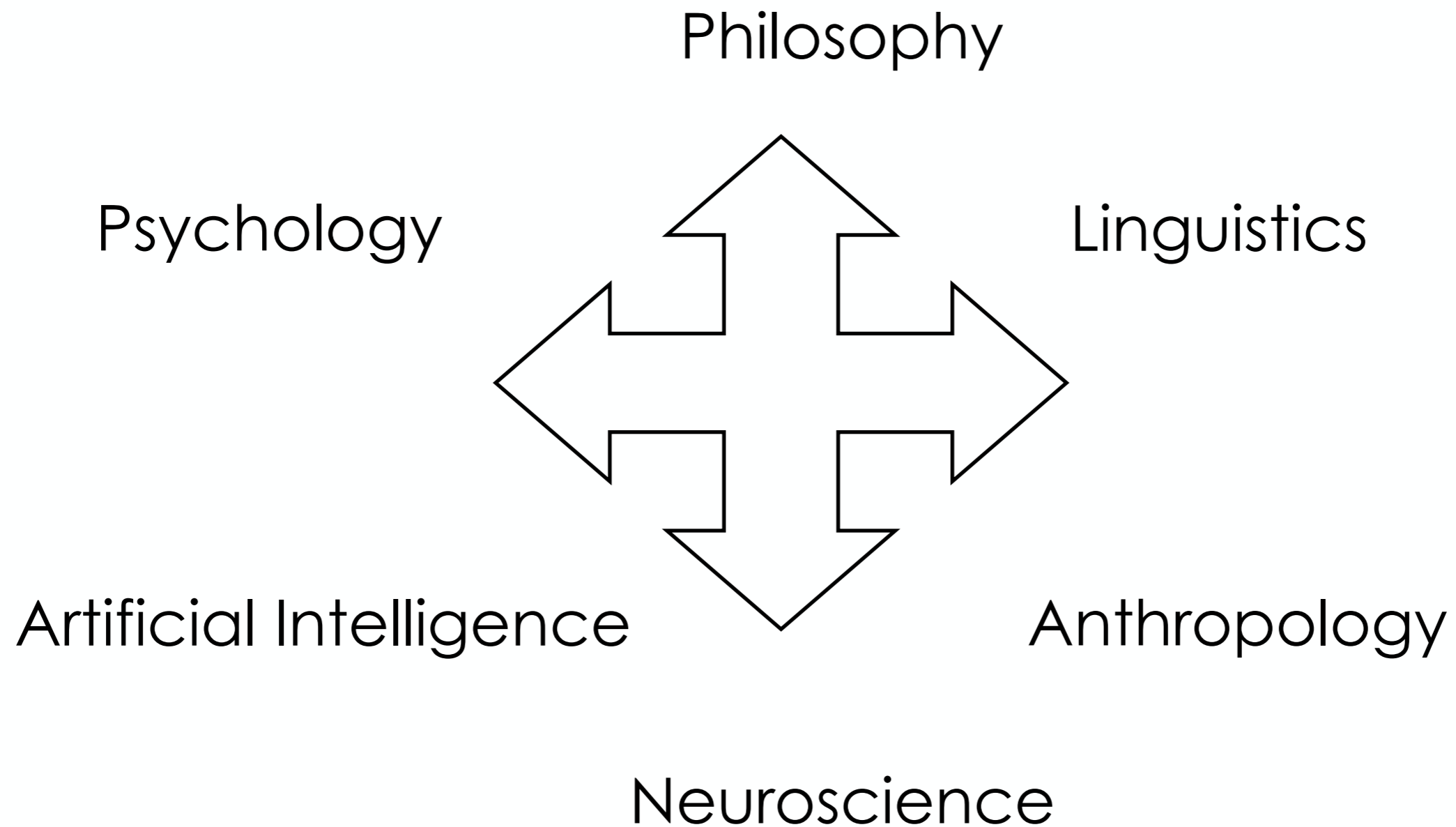
The mind/behavior/brain cycle



The mind/behavior/brain cycle



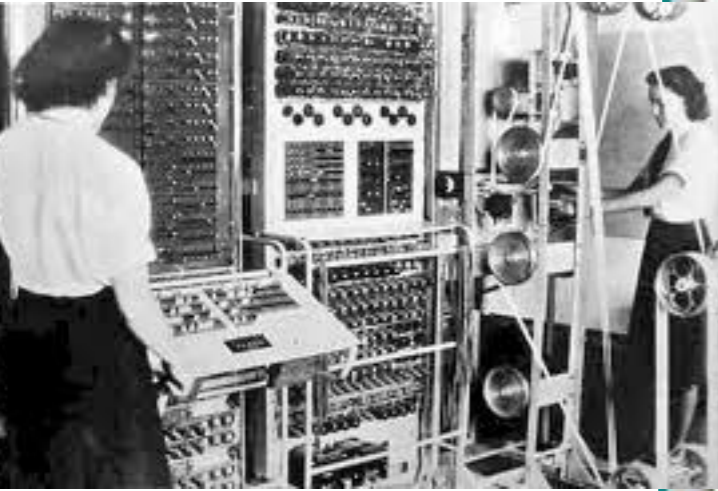
Cognitive Science



After Gardner (1985) The mind's new Science

THE WAR THAT CHANGED THE WORLD

How the Science and Technology of World War II Influences Your Life Today



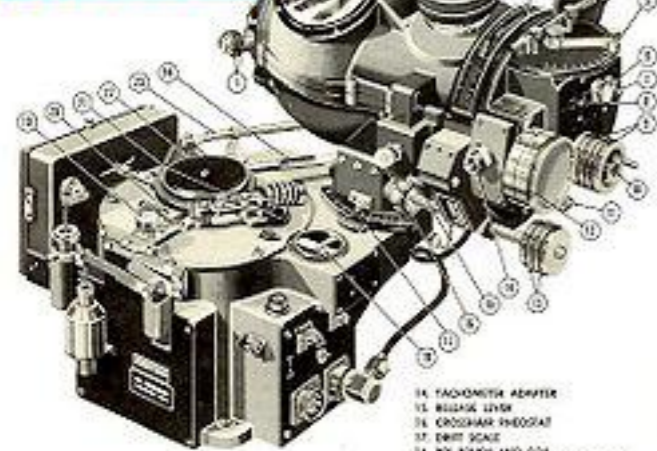
THE BOMBIGHT

The Bombight has 2 main parts, right-hand and left-hand. The right-hand part is the vibrator and is fed by the central feedings. The right-hand is connected to the vibrator by the vibrator through the bombight connecting rod and the bombight chain.

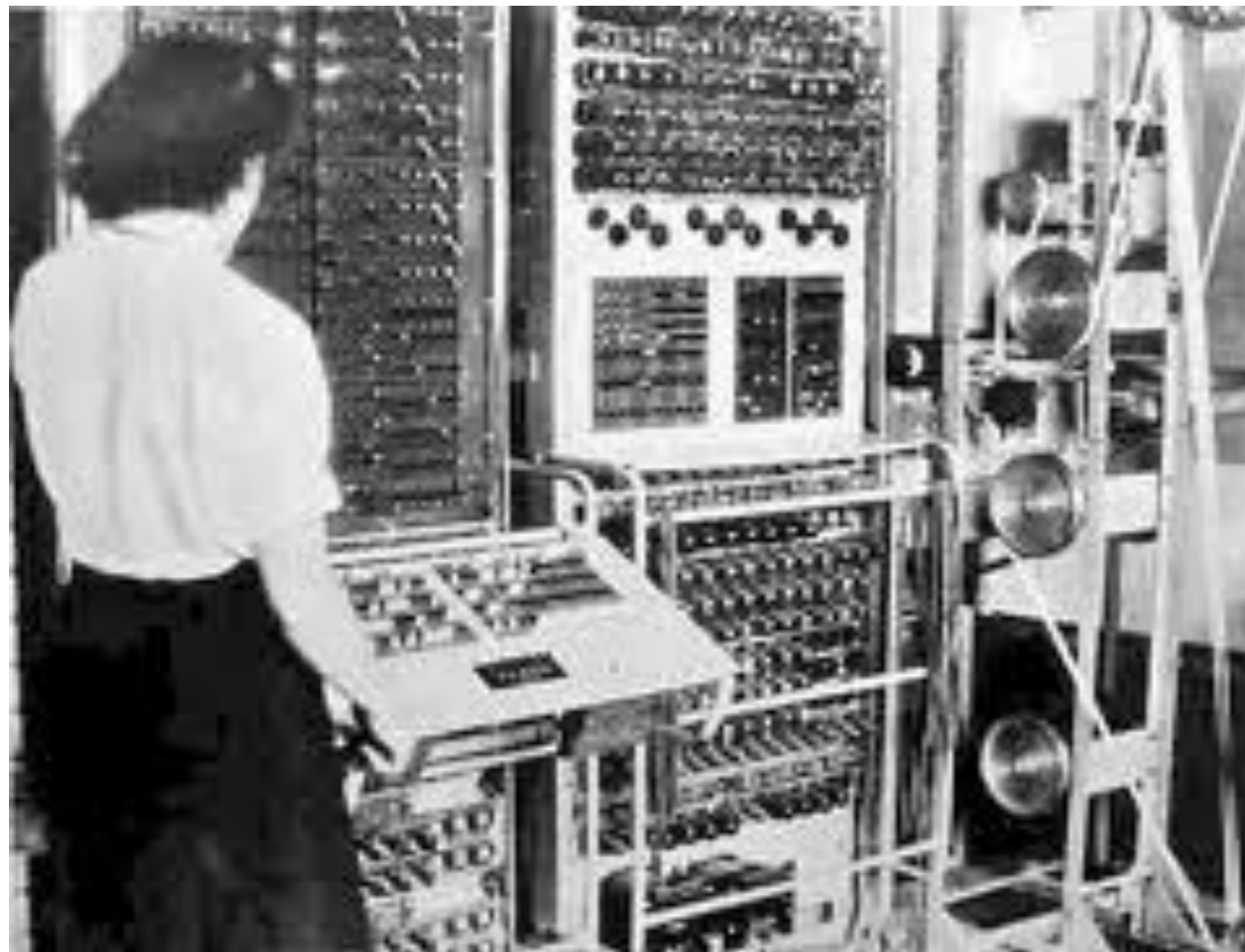
THE ORIGINAL WWII MODEL

MARCH, 1943 SEP 1943

NOMENCLATURE AND OPERATION



- 1. DRIVING KNOB
- 2. CASING KNOB
- 3. STOPPER
- 4. INDEX WINDOW
- 5. TRAIL ARM AND TRAIL PLATE
- 6. EXTENDED VISION KNOB
- 7. RATE MOTOR SWITCH
- 8. DISC SPEED GEAR SHIFTER
- 9. RATE AND REPLACEMENT KNOBS
- 10. MARCH DRIVE CLUTCH
- 11. MARCH KNOB
- 12. DISC SPEED GEAR
- 13. TURN AND EXIT KNOBS
- 14. TACHOMETER ADAPTER
- 15. RELEASE LEVER
- 16. CROSSHAIR PHOTOGRAPH
- 17. DRIFT SCALE
- 18. FOX BRASS AND COIL
- 19. AUTOMATIC CLUTCH ENGAGING LEVER
- 20. FORWARD CLUTCH
- 21. FORWARD CLUTCH ENGAGING LEVER
- 22. FORWARD CLUTCH
- 23. FORWARD CONNECTING ROD
- 24. AUTOMATIC CONNECTING ROD



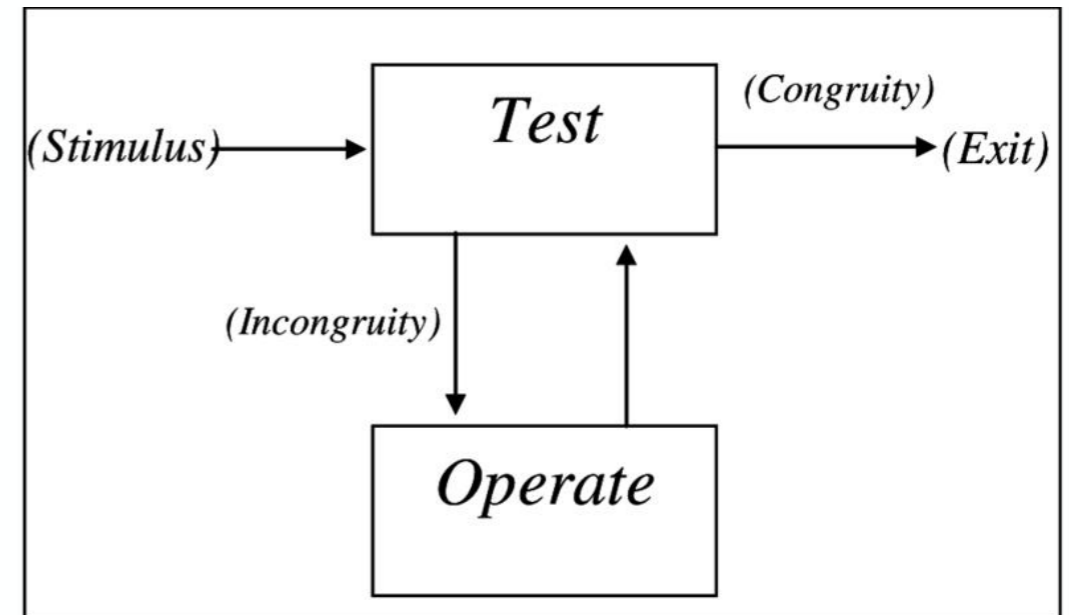
Artificial Intelligence



Cybernetics

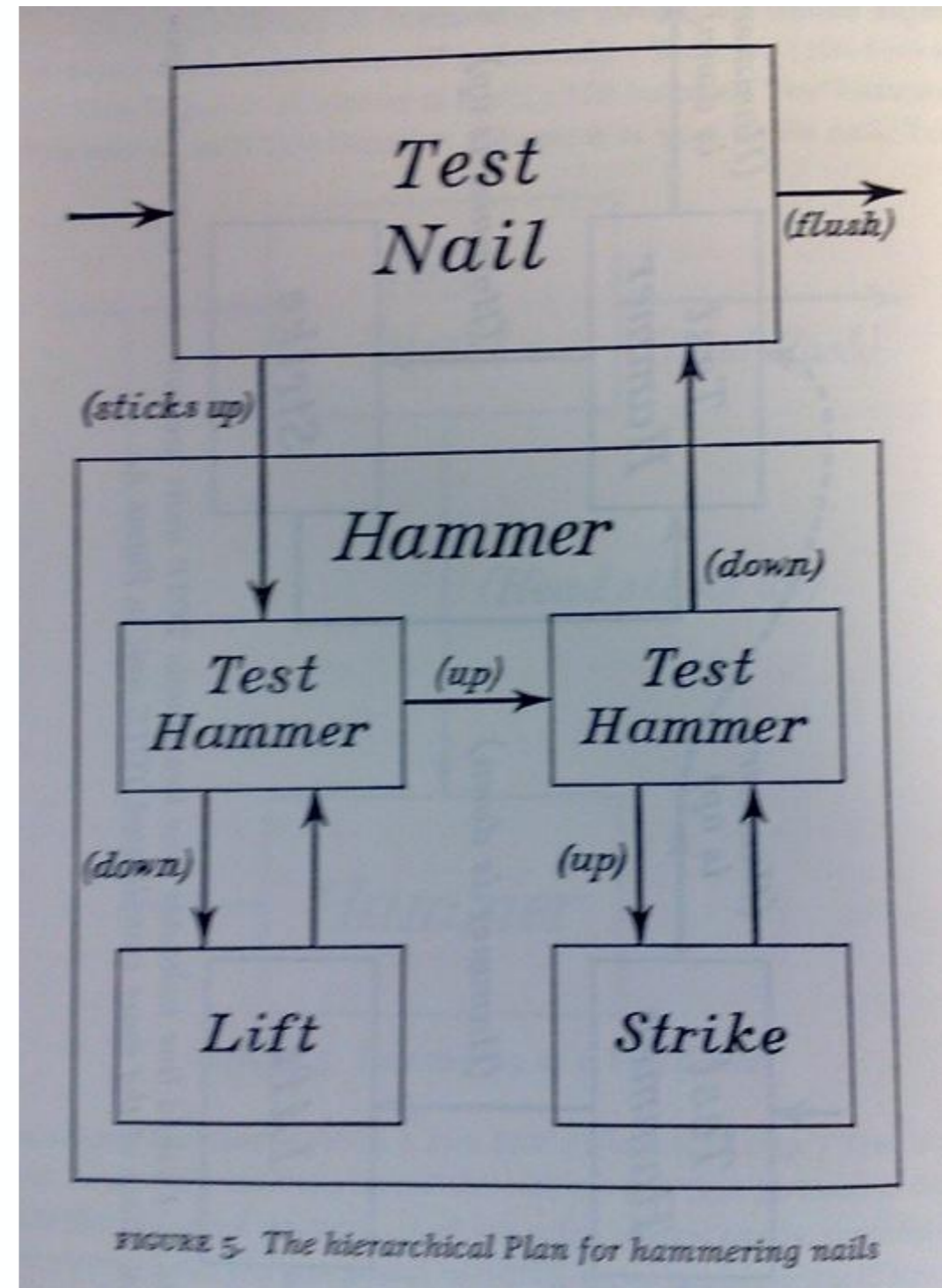
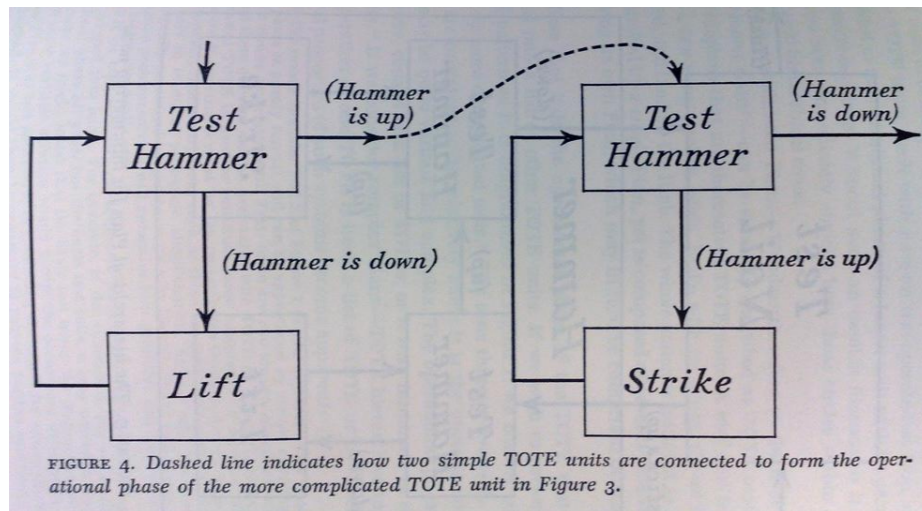
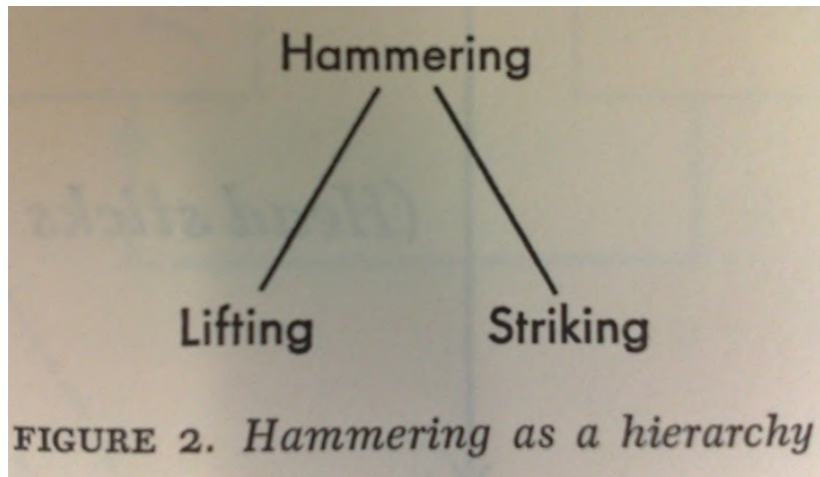
Transformation: The TOTE Unit

- Reflex is dead but what is the unit of behavior?
- A **plan** is any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed
- **Molar-Molecular** units: Strategy and tactics of behavior
- **Execution**: the plan is controlling the sequence of behavior
- **Image**: all organized knowledge about the world and the self



Plans and the Structure of Behavior
Miller, Gallanter & Pribram (1960)

A simple behavioral plan



Cybernetics

- Postwar Cybernetics movement
- The power of feedback
- Formal analysis of real-world systems



GW fixing a turtle



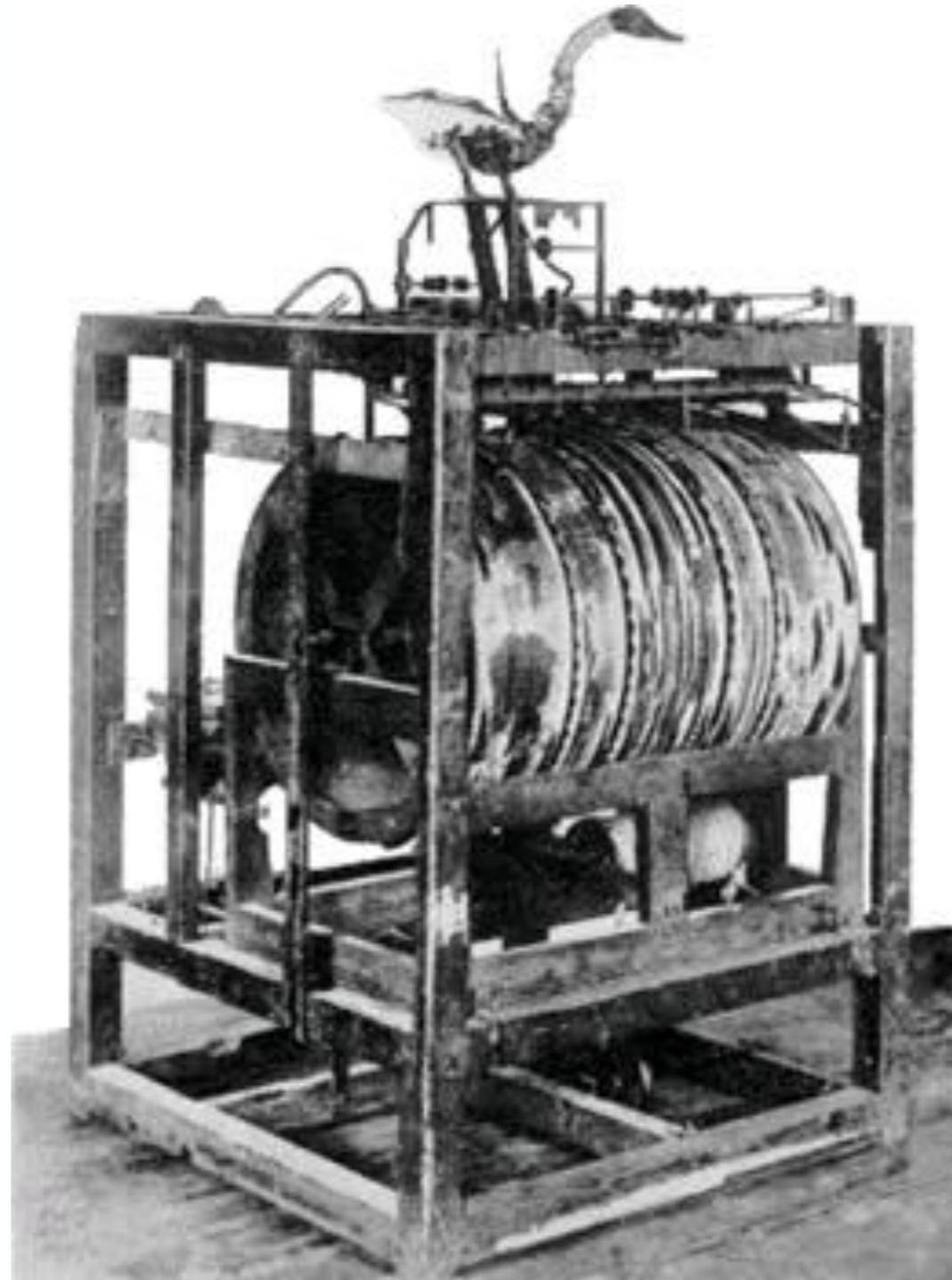
fltr: W.Ross Ashby, Warren McCulloch, Grey Walter, and Norbert Wiener (from Latil, P de: Thinking By Machine, 1956)

Leonardo's Robot



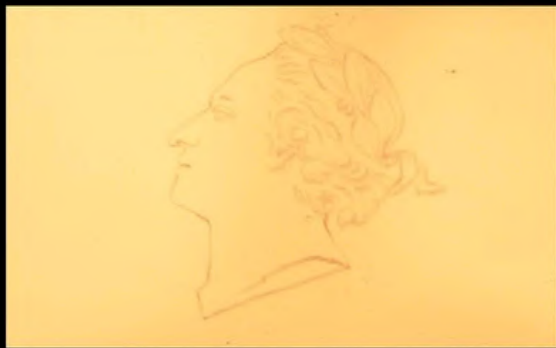
Leonardo's robot 1495: The robot is a knight, clad in German-Italian medieval armor, that is apparently able to make several human-like motions. These motions included sitting up, moving its arms, neck, and an anatomically correct jaw

Automata of Vaucanson



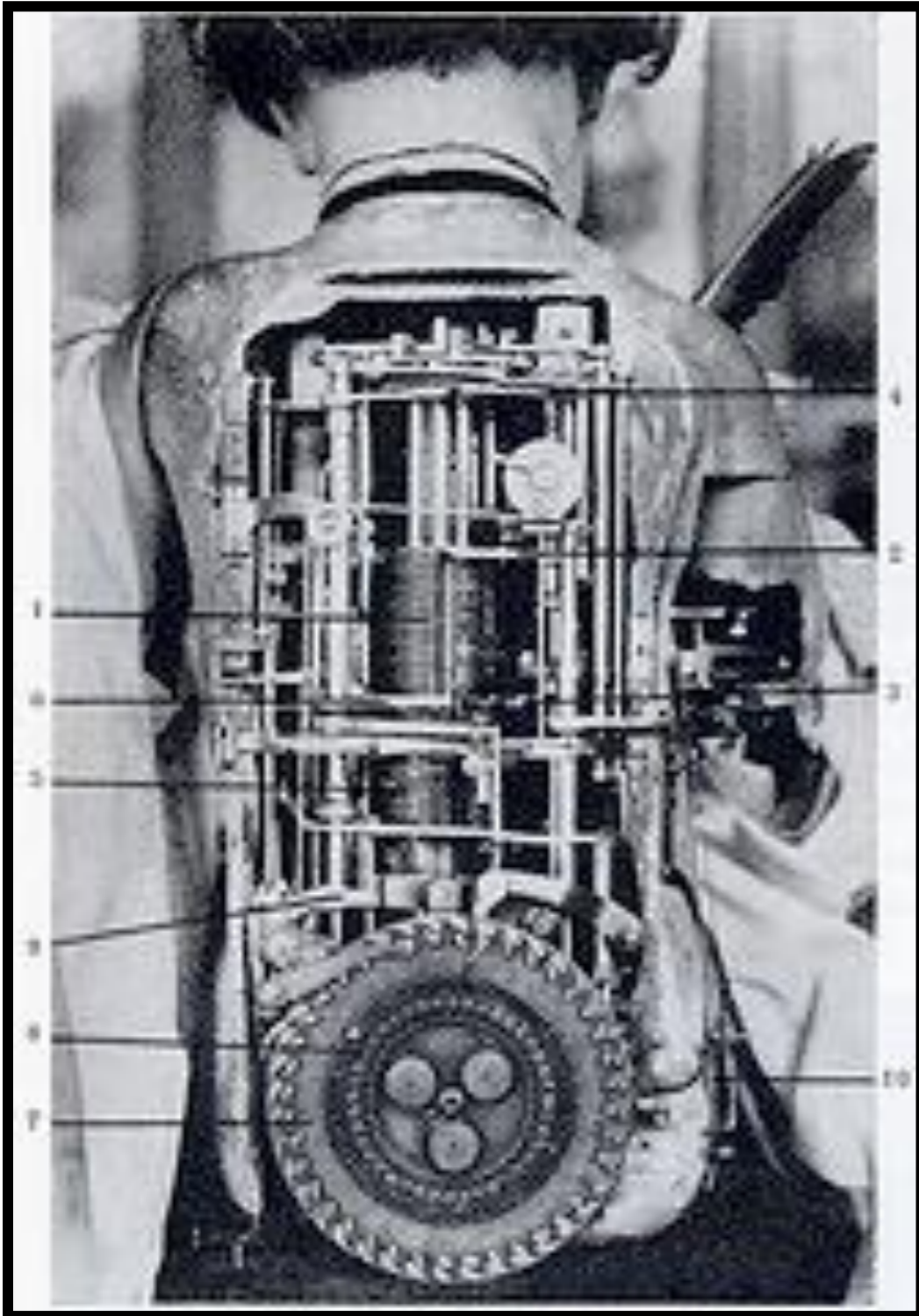
Duck, 1798

Brothers Droz-Neuchâtel



Constructed between 1768 and 1774 by Pierre Jaquet-Droz, his son Henri-Louis (1752-1791), and Jean-Frederic Leschot (1746-1824) were The Writer (made of 6000 pieces), The Musician (2500 pieces) and The Draughtsman (2000 pieces). By some considered as the first computers

Automata



Automata, such as the little moving figures of people or animals that emerge from cuckoo clocks and music boxes, were popular in the 1700's and machines capable of thinking were a subject for speculation long before the electronic computer was invented.

L'Homme Machine (1747)

Humans are machines

mental states depend on bodily actions

The organization of matter at a high and complex level resulted in human thought.

Psycho/Neuropathology depends on the body

(Introspective/Physician)

Determinism

Hedonism drives behavior

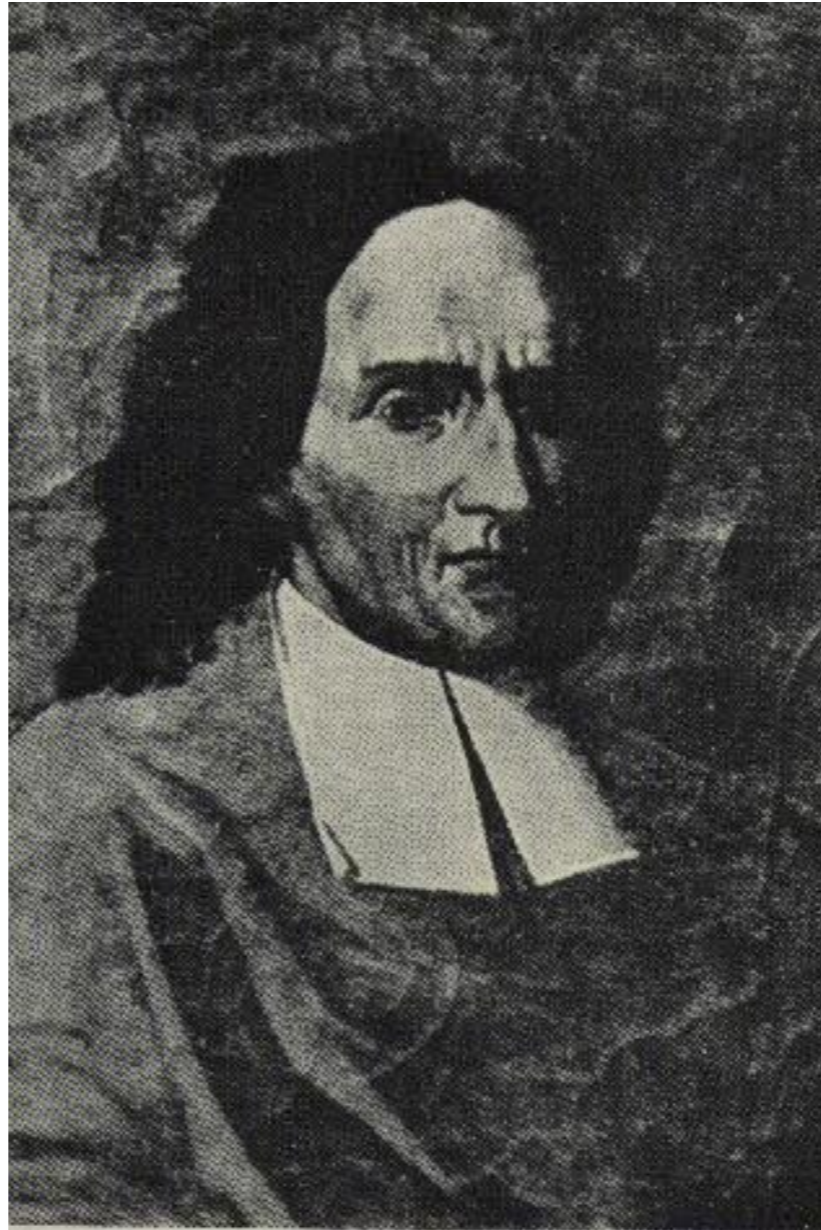
Humans have no “higher” morality than animals (e.g. humans torture)

Inspires behaviorism and a reductionist approach towards mind.



Julien Offray de La Mettrie
1709-1751

Factum et Verum Convertit



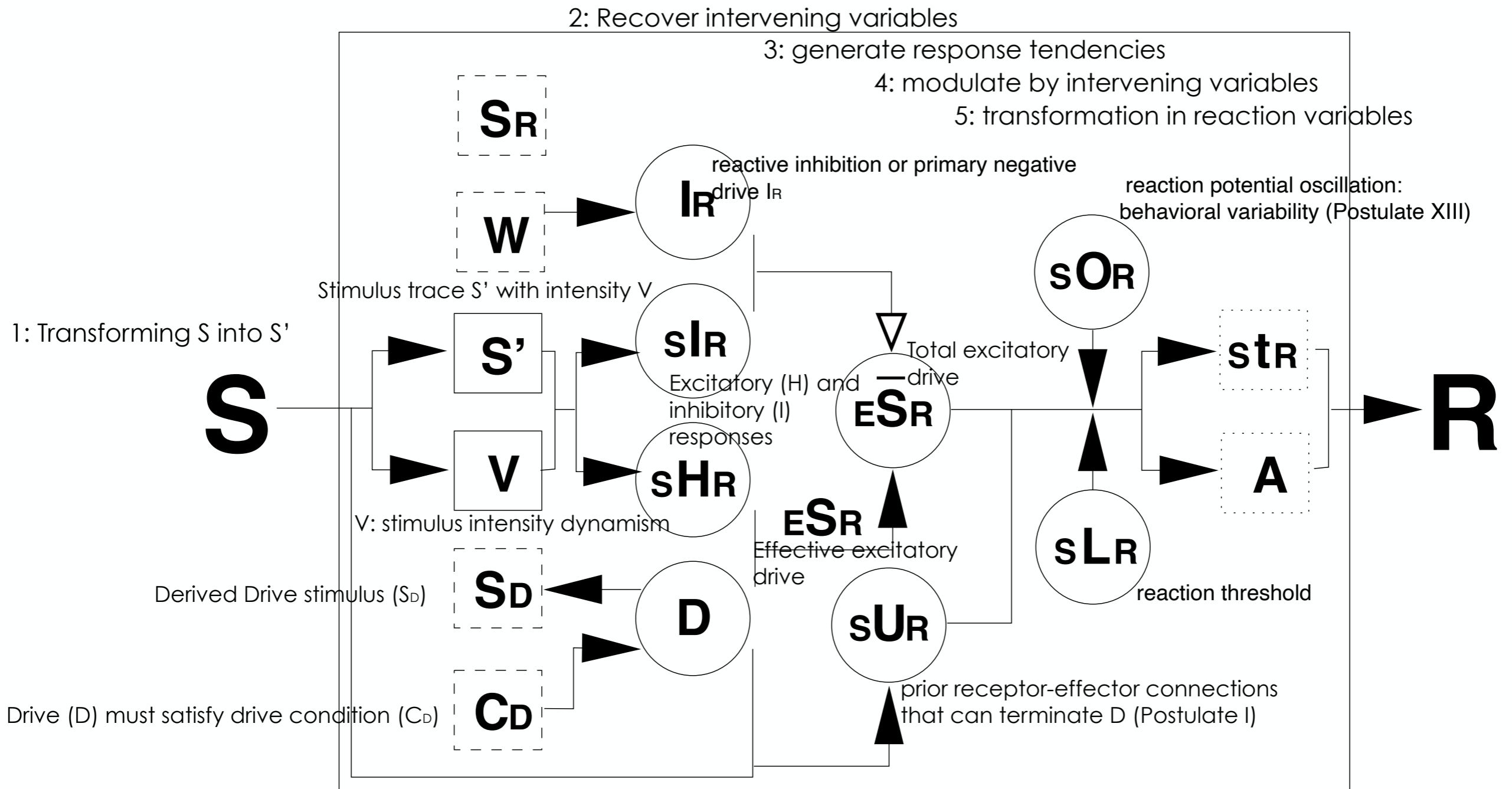
Giambattista Vico (1668-1744)

Truth and fact are exchangeable

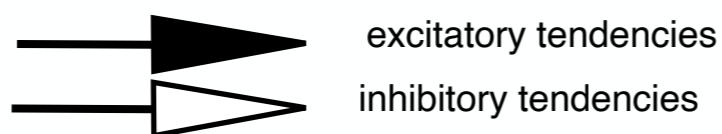
“The criterion and rule of the true is to have made it. Accordingly, our clear and distinct idea of the mind cannot be a criterion of the mind itself, still less of other truths. For while the mind perceives itself, it does not make itself.” (1710) *De Italorum Sapientia*.

Cybernetics and
alternative movement:
The power of feedback
and construction

Transforming a stimulus into a response



- traces of external stimuli
- traces of internal stimuli
- Intervening variables
- response variables



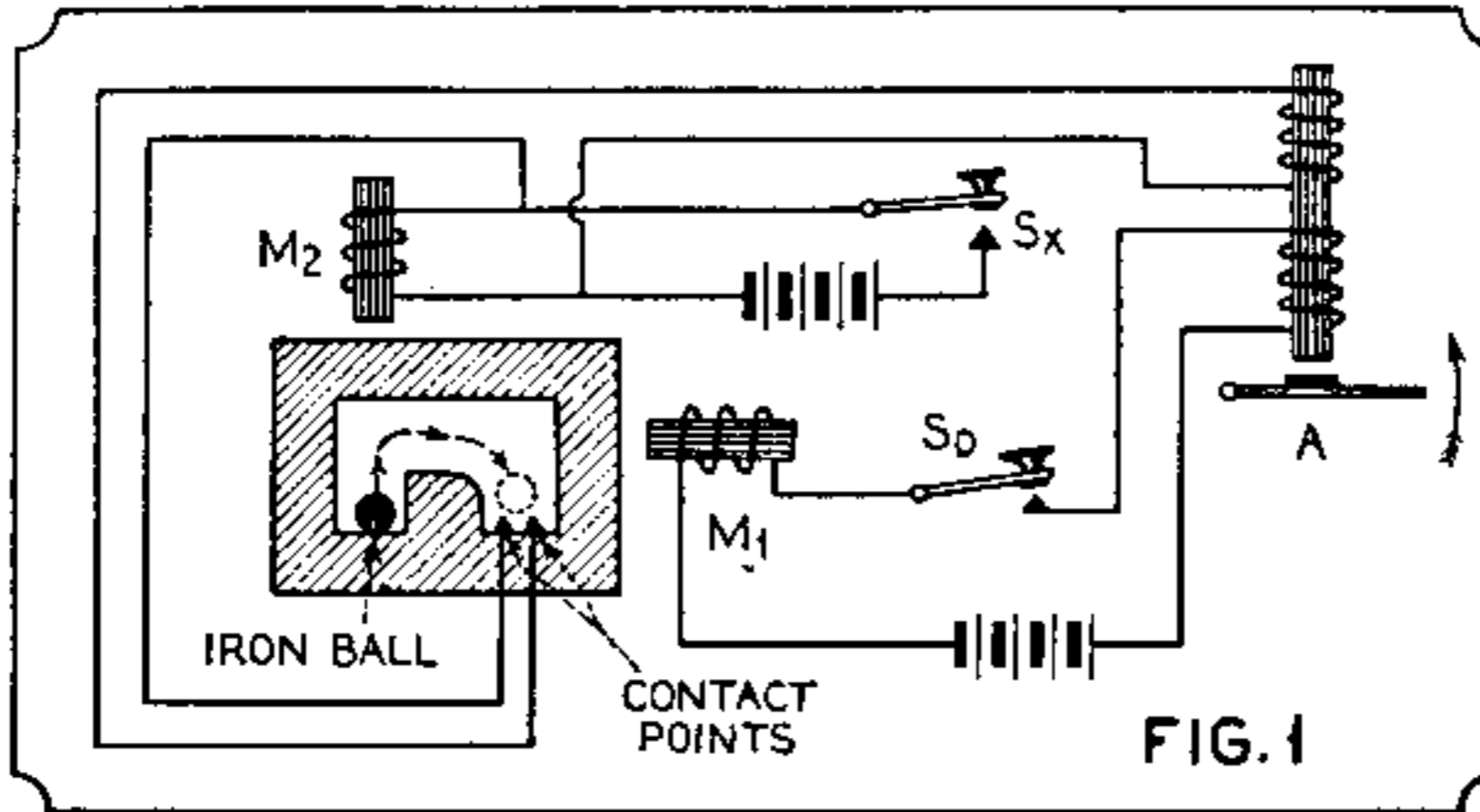
$$SER = (S_{HR} \times D \times K \times V) - (S_{IR} + I_R)$$

$$S_{HR} = 1 - 10^{-.0305N} \quad N = N \text{ trials}$$

$$\Delta SER = M^E - SER - (M^E - SER)10^{-i}$$

M = Maximum reaction potential
i = animal dependent learning rate

The law of habit formation



An electrical memory cell that demonstrates, within its limits, the basic element of mental activity

Ross, T. (1935) *Machines that think. A further statement. Psych Review*

Cybernetics = Regulation of Systems

This question was of interest to a handful of people in the 1940s who were the pioneers in a field that has become known as Cybernetics, the science of the regulation of systems.



Cybernetics is derived from the Greek word for steersman or helmsman, who provides the control system for a boat or ship.

Macy Foundation Meetings 1946 - 1953

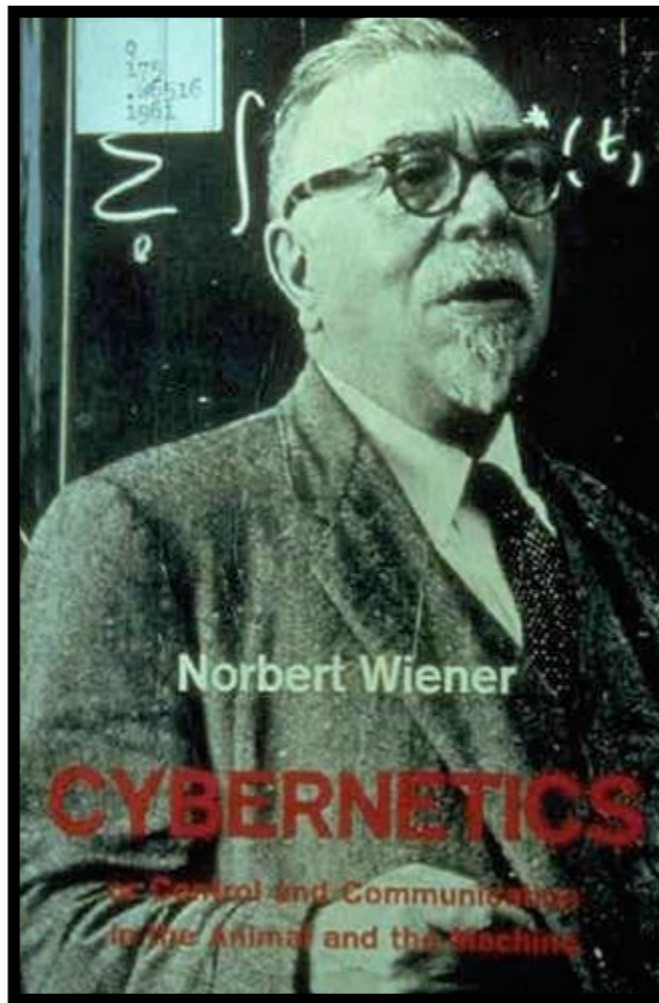
From 1946 to 1953 there was a series of meetings to discuss feedback loops and circular causality in self-regulating systems.

The meetings, sponsored by the Josiah Macy, Jr. Foundation, were interdisciplinary, attended by engineers, mathematicians, neurophysiologists, and others.



Norbert Wiener (the Father of Cybernetics)

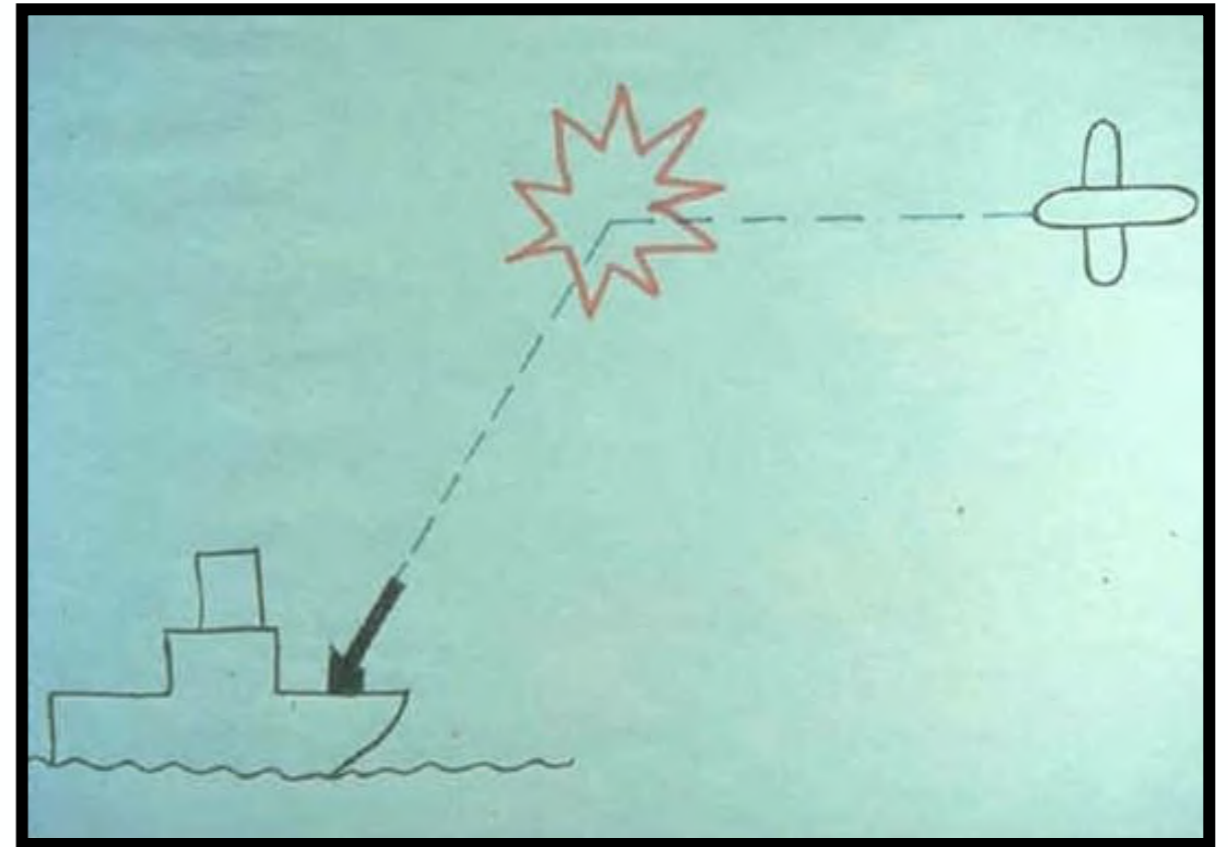
Cybernetics coined in 1948



1894 - 1964

Wiener – A practical problem

Wiener: applied mathematician, biologist, and electrical engineer. He worked during World War II on the radar-guided anti-aircraft gun.



He connected a special radar to the gun so that it was aimed automatically at the enemy aircraft. After the gun was fired, the radar quickly determined the changing location of the plane and re-aimed the gun until the plane was shot down.

Feedback

The anti-aircraft gun: the cybernetic principle of **feedback**.

Feedback is information about the **results** of a process which is used to **change** the process.

The radar provided information about the changes in location of the enemy airplane and this information was used to correct the aiming of the gun.



Cannon, W.,
'The Wisdom of the Body', 1932

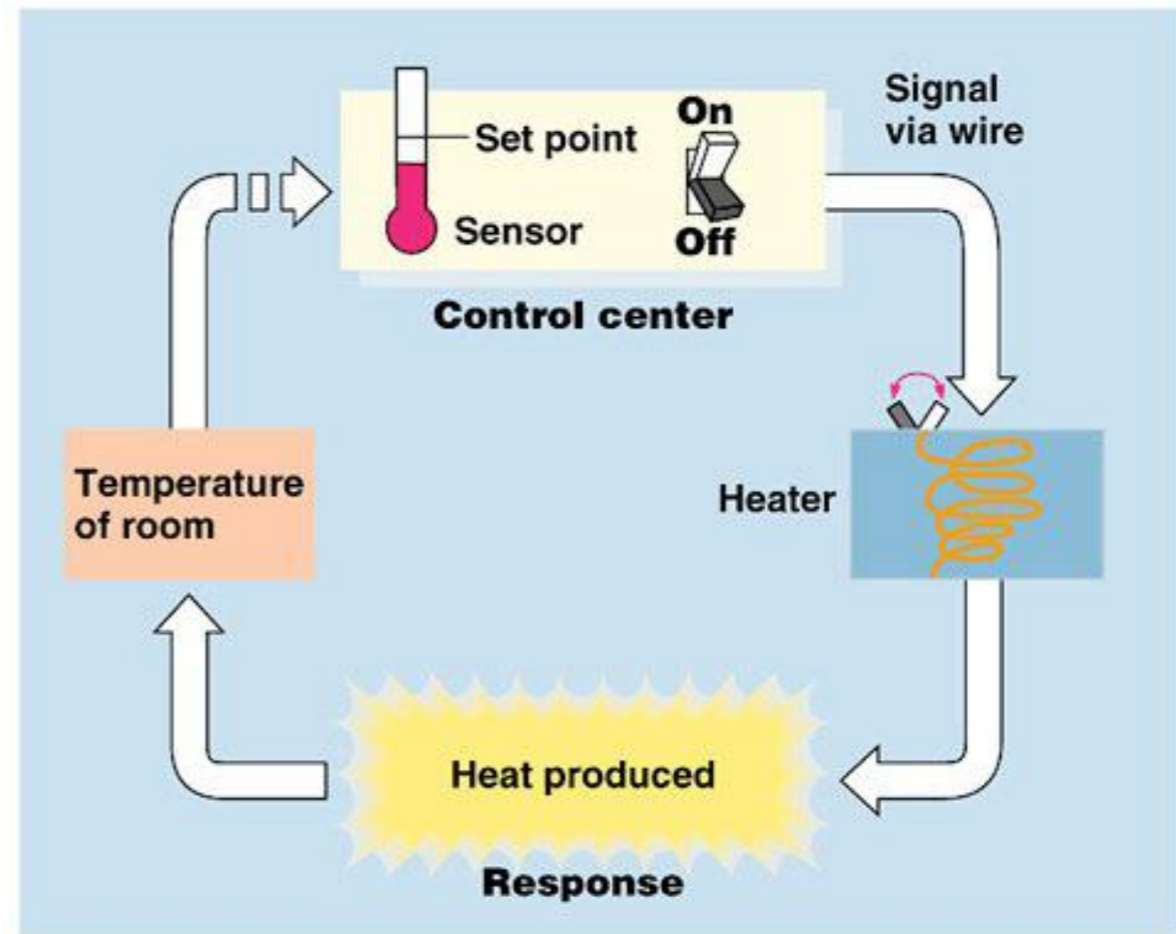
A more familiar example of the use of feedback to regulate a system is the common thermostat for heating a room.

Homeostasis and behaviour / motivation



Cannon, W.,
'The Wisdom of the Body', 1932

Homeostasis

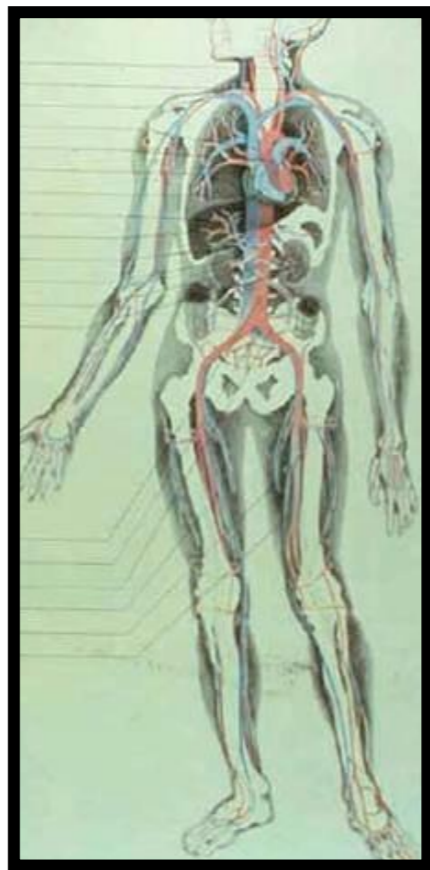


(a) Control of room temperature

Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Self Regulating System

The sensor provides a feedback loop of information that allows the system to detect a difference from the desired temperature of 68 degrees and to make a change to correct the error. As with the anti-aircraft gun and the airplane, this system – consisting of the thermostat, the heater and the room – is said to regulate itself through **feedback** and is a self-regulating system.

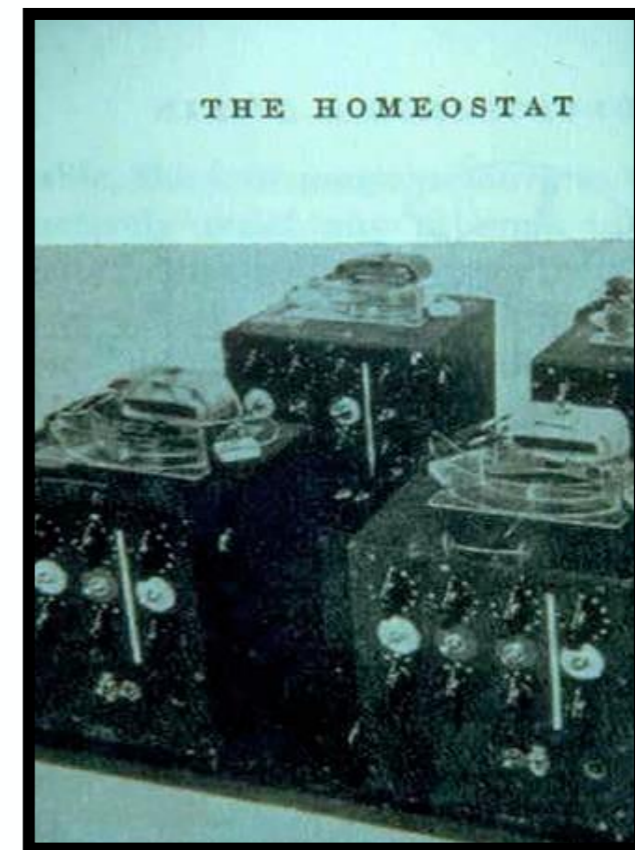


Cannon, W.,
'The Wisdom of the Body', 1932

The human body is one of the richest sources of examples of feedback that leads to the regulation of a system. For example, when your stomach is empty, information is passed to your brain.

Human Body and Cybernetics Studies

The human body is such a marvel of self-regulation that early cyberneticians studied its processes and used it as a model to design machines that were self-regulating. One famous machine called the **homeostat** was constructed in the 1940s by a British scientist, **Ross Ashby**.

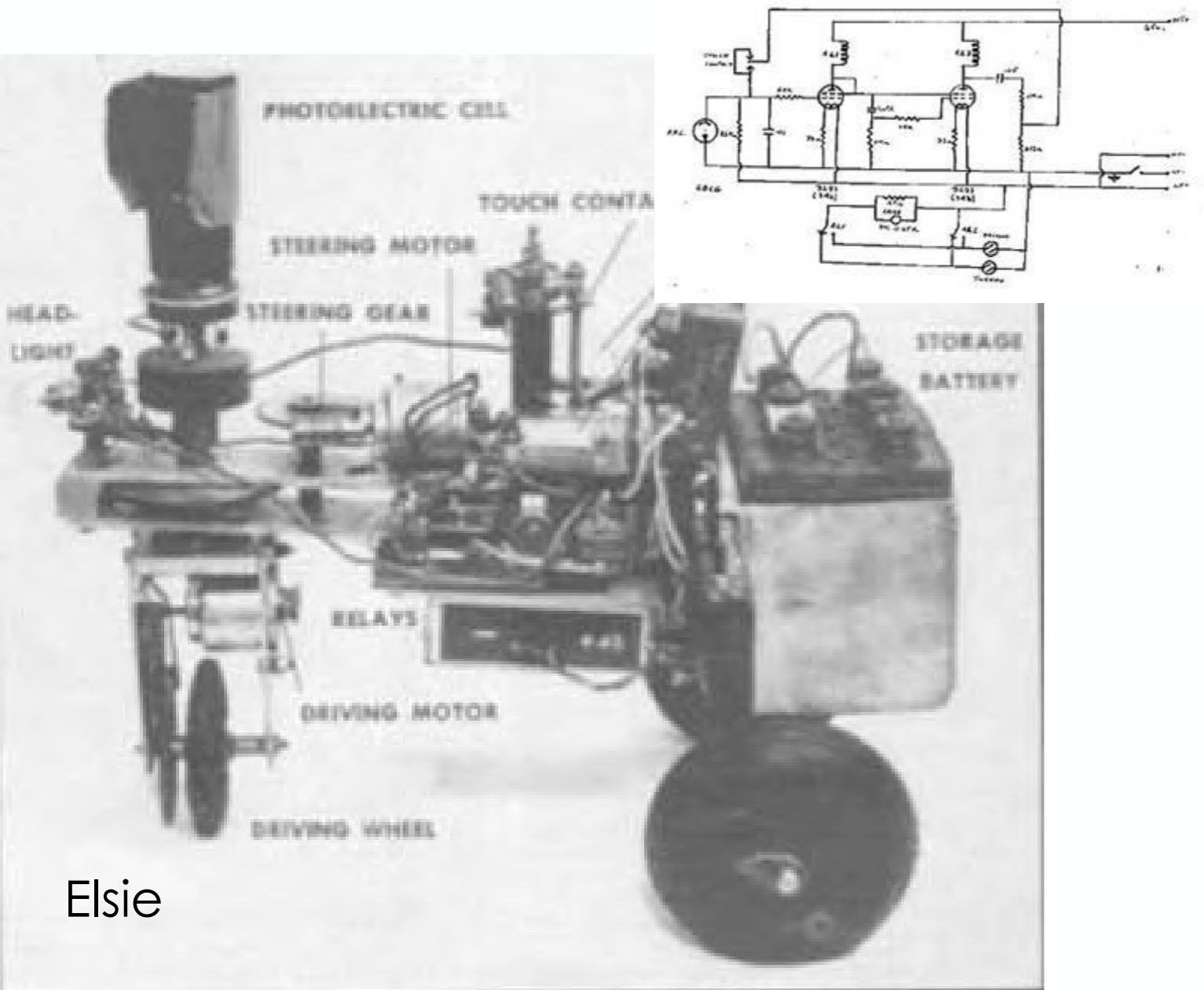


Just as the human body maintains a 37 degree temperature the **homeostat** could maintain the same electrical current, despite changes from the outside.

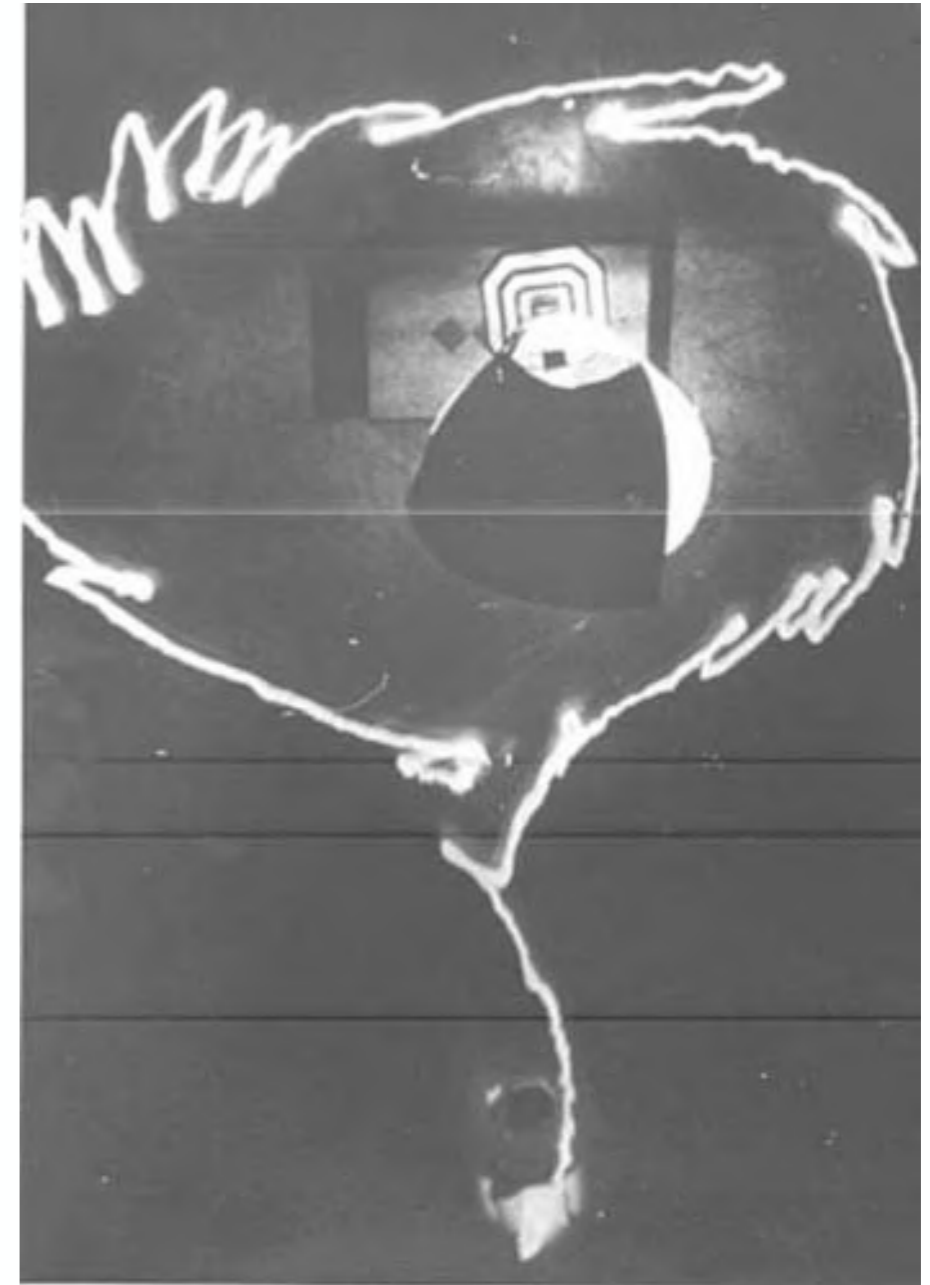
Grey Walter – Self Regulating in Man and Animals



Elsie's reactive behaviors



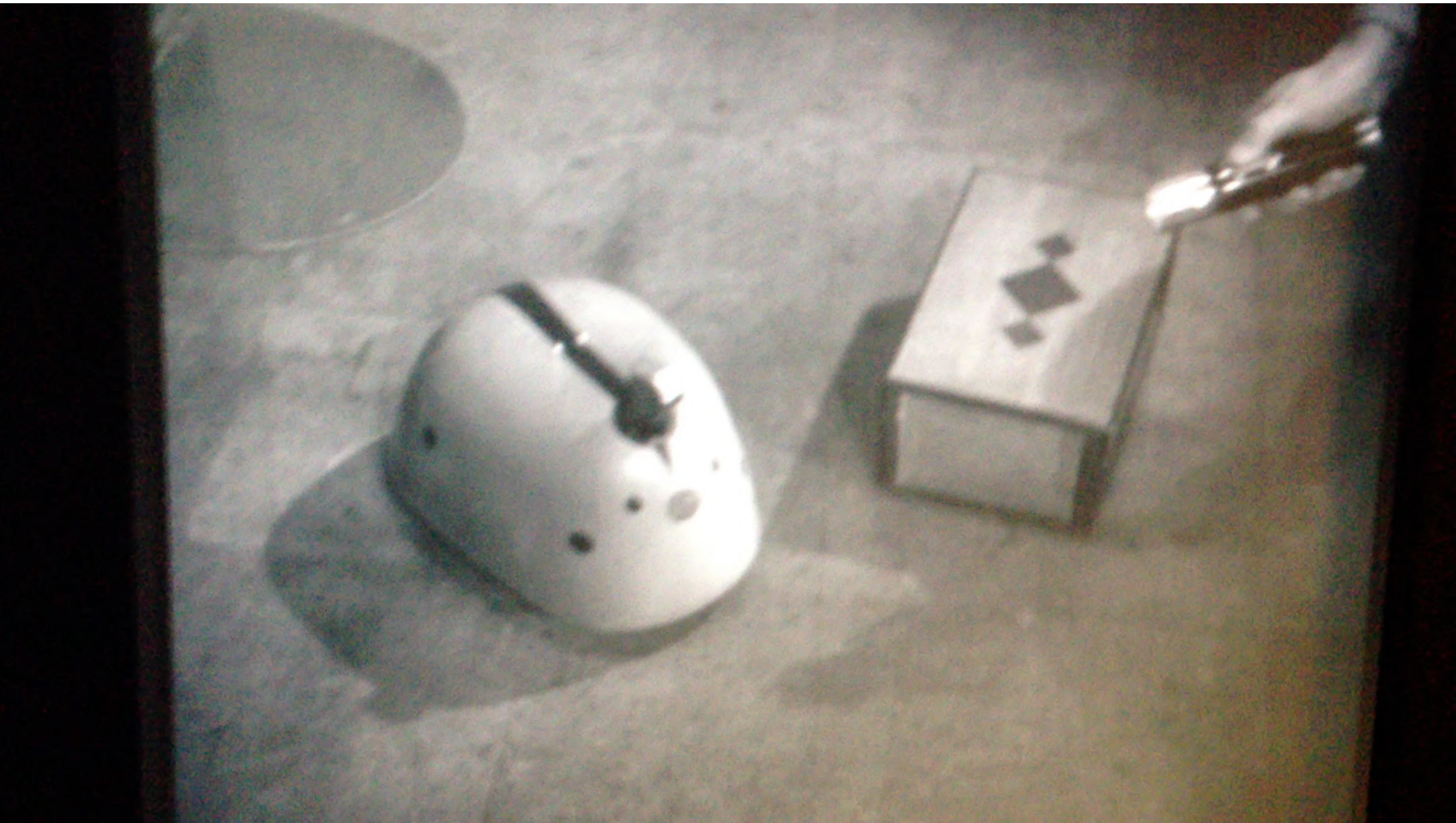
Elsie

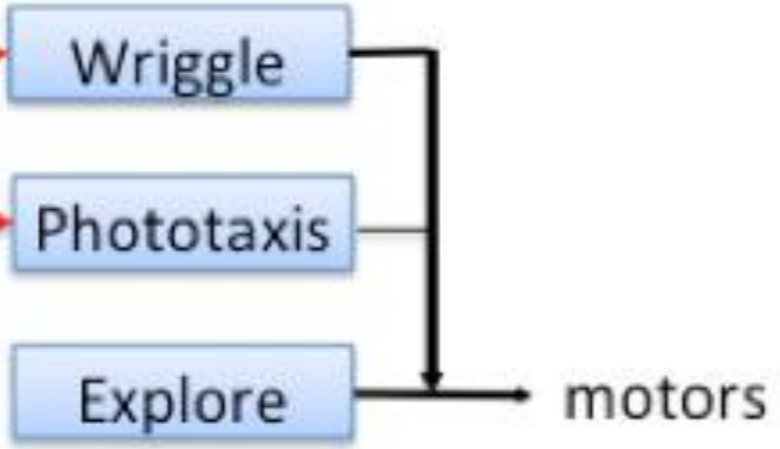
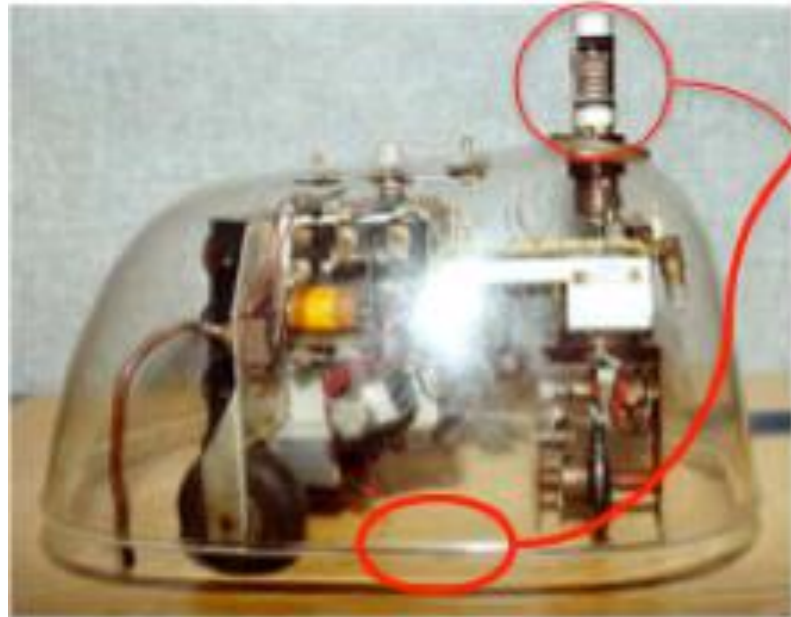


Elsie approaches a light and circles around it (1950)

The Grey Walter picture archive
<http://www.ias.uwe.ac.uk/Robots/gwonline/gwarkive.html>

Film footage: *Mechanical Tortoise*, British Pathé, 1951





Neurophysiology, Mathematics, and Philosophy

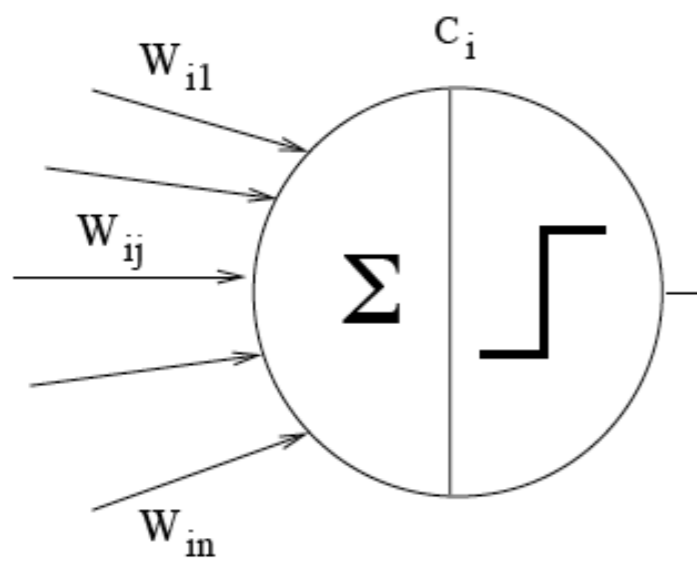


Warren **McCulloch** was a key figure in enlarging the scope of cybernetics. Although a psychiatrist by training, McCulloch combined his knowledge of neurophysiology, mathematics, and philosophy to better understand a very complex system . . .

A logical calculus of the nervous system



- Ada Lovelace: A calculus of the nervous system
- McCulloch & Pitts: The logic of neuronal circuits



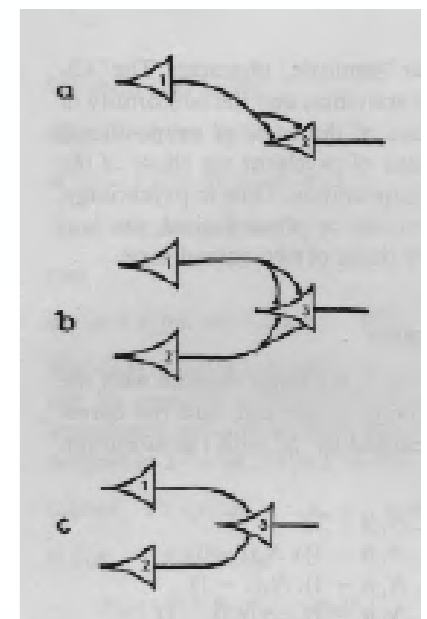
$$N_i(t + 1) = \theta(\Sigma_j w_{ij} n_j(t) - \mu_i)$$

$$N_i = \begin{cases} 1 & \text{if } \Sigma_j w_{ij} n_j(t) > \mu_i, \\ 0 & \text{otherwise.} \end{cases}$$

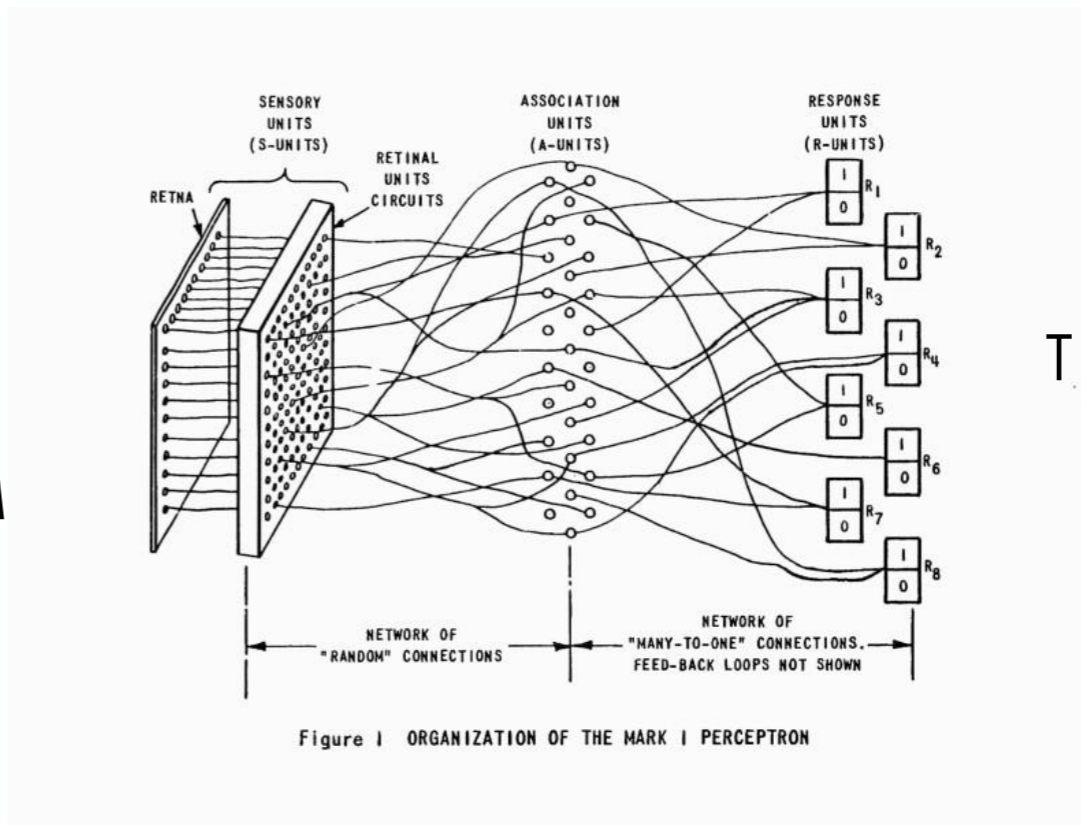
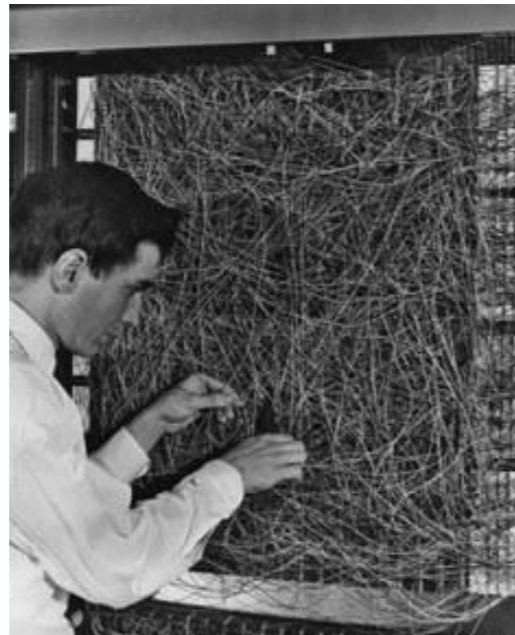
=

or

and

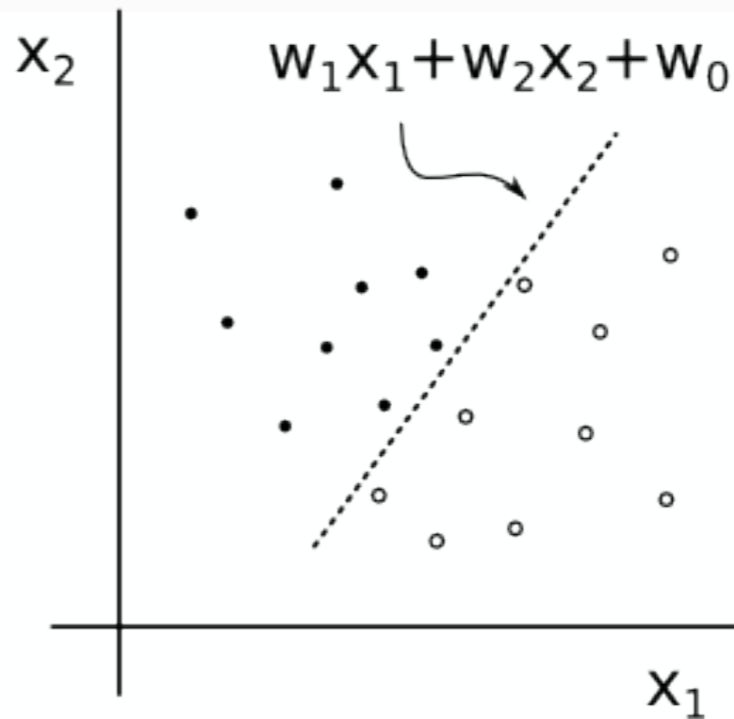
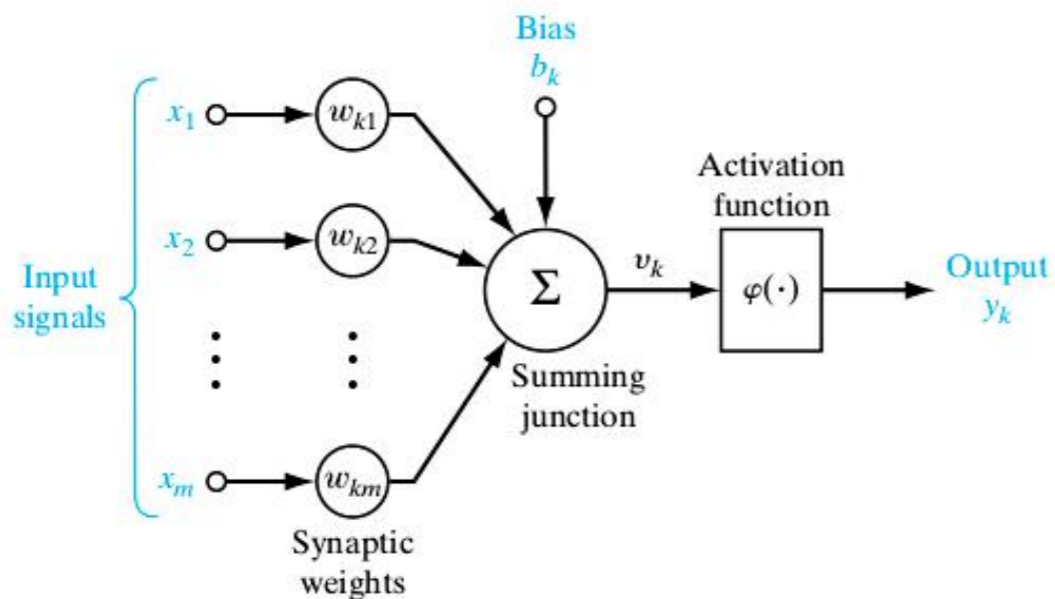


The Perceptron



Frank Rosenblatt (July 11, 1928 – July 11, 1971)

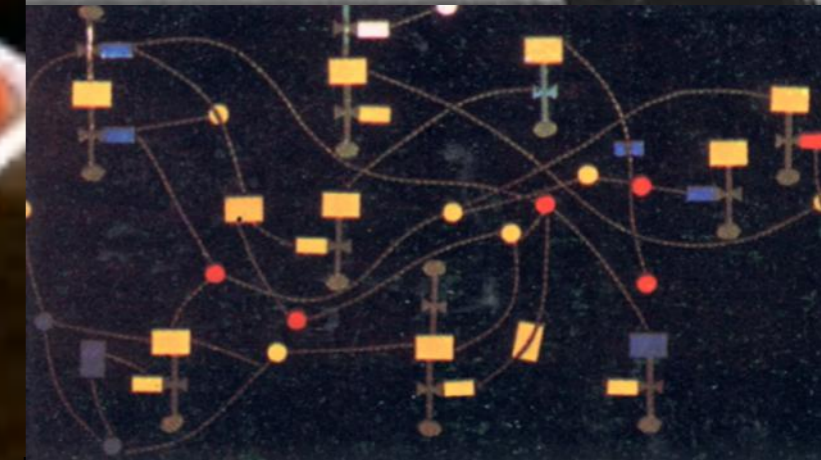
simulated on an IBM 704 computer at Cornell Aeronautical Laboratory in 1957



Cybernetics

- Machine mind
- Control theory
- Embodiment
- Sensory - motor mapping
- Behavior

CyberSyn: Cybernetic Society



Stafford Beer

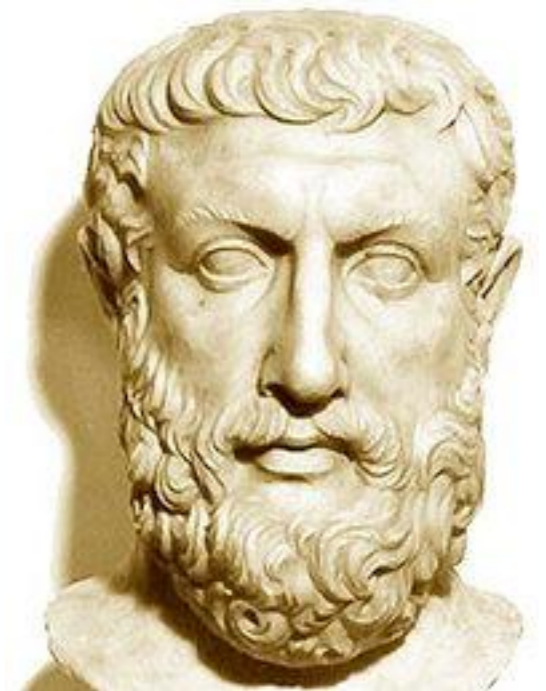
From Analytic Engine to Computer via the Turing Machine

Resurrection of Rationalism

“For the same are the thinking and the
being”

(Parmenides, 480BC, Fragment V)

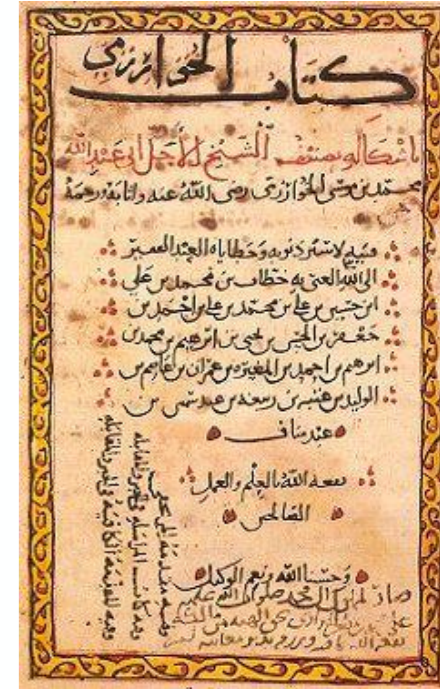
The senses are unreliable (hallucinations)
Content of our thoughts are there all along....



ca. 520 BC ca. 450 BC

The notion Algorithm was invented in the Arab world

```
function MAX-VALUE(state) returns a utility value
  if TERMINAL-TEST(state) then return UTILITY(state)
   $v \leftarrow -\infty$ 
  for  $a, s$  in SUCCESSORS(state) do
    if WINNER( $s$ ) = MAX
      then  $v \leftarrow \text{MAX}(v, \text{MAX-VALUE}(s))$ 
      else  $v \leftarrow \text{MAX}(v, \text{MIN-VALUE}(s))$ 
  return  $v$ 
```



al-Kuwarizmi (c.830)

Gave rise to the term ALGORITHM

He wrote an extensive account of the Hindu system of numerals and numeration from which our current system evolved.

ALGURISM: Writing numbers and performing calculations using Hindu numbers

Competitions were held between the abacists, who favoured the abacus for calculations, and the algorists, who preferred pencil-and-paper calculations.



Mind as Logic as Computation



June 19, 1623 - August 19, 1662 (aged 39)

Blaise Pascal



Pascal's calculator or Pascaline

- Wilhelm Schickard (1592 –1635): the Speeding Clock
- Gottfried Wilhelm Leibniz (1646 –1716)

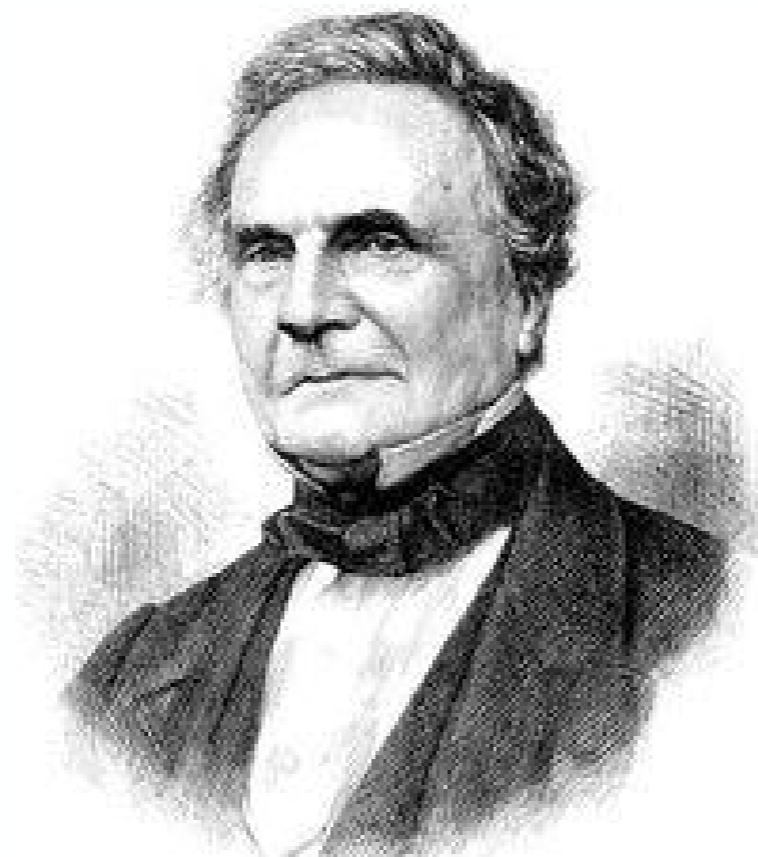
Babbage: The Difference & Analytic Engine



section of the "mill" of the Analytical Engine (above) and Difference engine (below). Reconstructed by his son Henry



Difference engine: compute tables
Analytic engine: first programmable computer



Charles Babbage
(1791-1871)

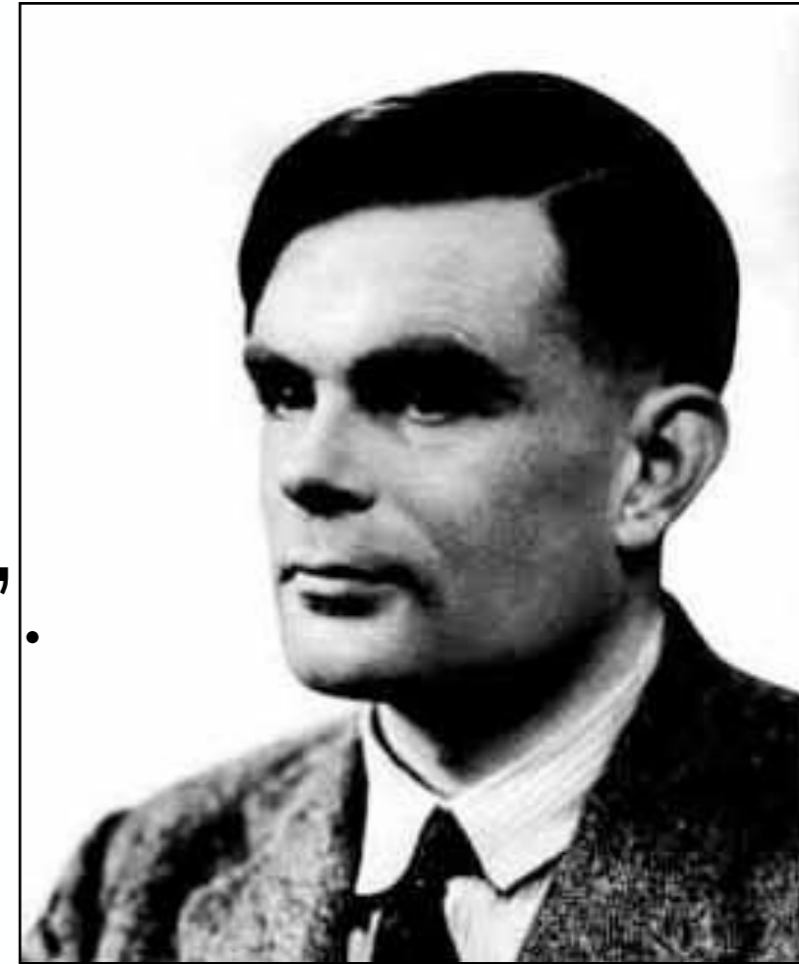


Augusta Ada Byron King,
Countess of Lovelace
(1815-1852)

"We may say most aptly that the Analytical Engine weaves algebraic patterns just as the Jacquard-loom weaves flowers and leaves."

Turing and the Turing Machine

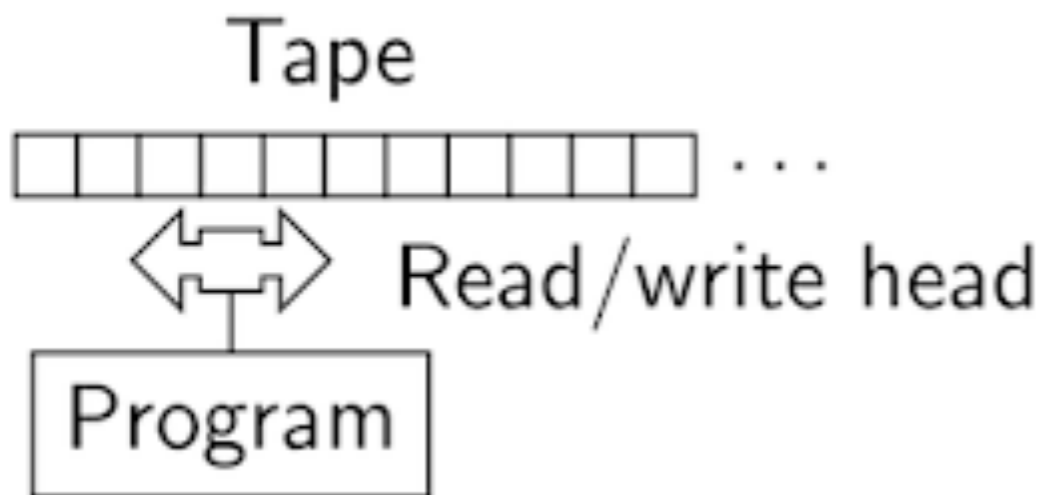
In 1936, Turing introduced an abstract model for computation in "*On Computable Numbers, with an application to the Entscheidungsproblem*".



(1912–1954)

Runs 2:46 marathon in qualifier for 1949 Olympics

The Turing Machine



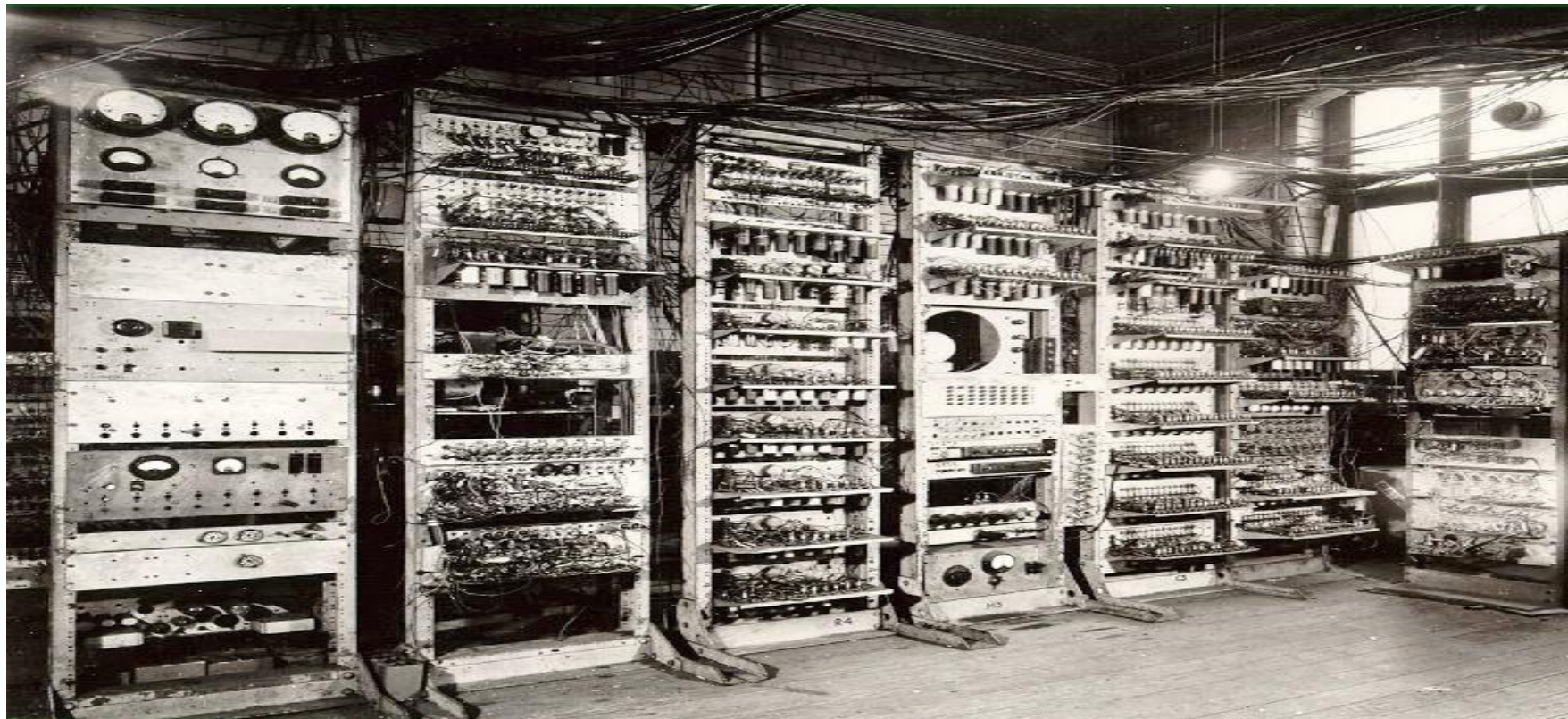
- 1: A *tape* which is divided into discrete cells (infinite memory).
- 2: A *head* that can read and write symbols on the tape and move left and right.
- 3: A *state register* that stores the state of the Turing machine
- 4: An *action table* (or *transition function*)

Church -Turing thesis

- Church-Turing thesis: *“Any computer program in any of the conventional programming languages can be translated into a Turing machine, and any Turing machine can be translated into most programming languages, so the thesis is equivalent to saying that the conventional programming languages are sufficient to express any algorithm”*

Digital Computers

- The 2nd world war added new technology and a new challenge: decoding
- Manchester Mark 1



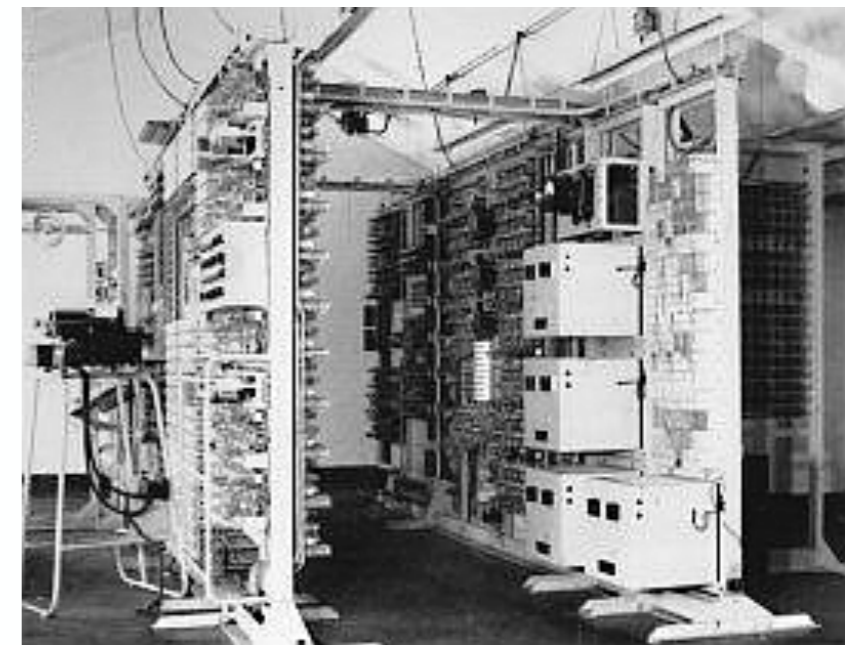
Breaking the enigma code



Enigma: 1918 Arthur Scherbius



Bletchley park

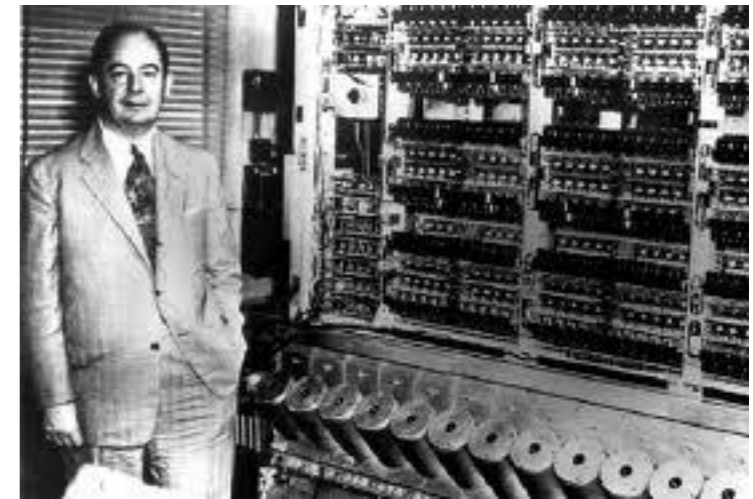


"Bombs"

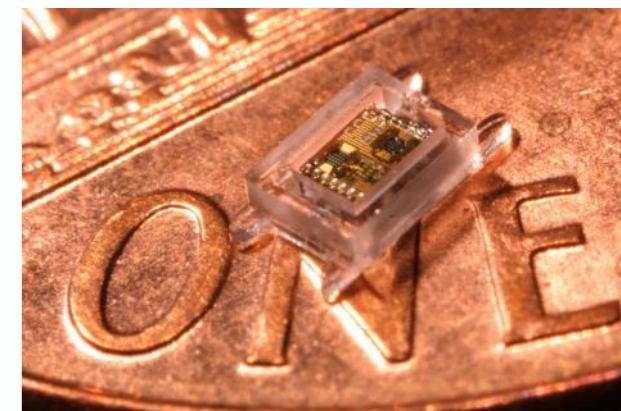
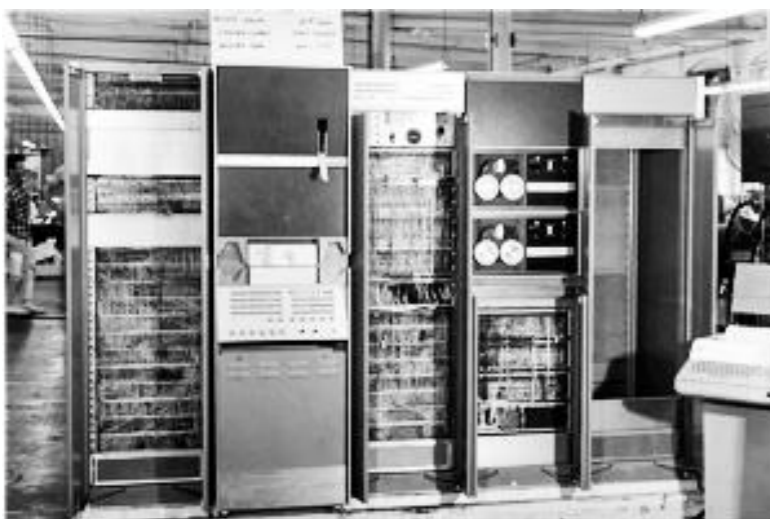
Computers

All computers more or less based on the same basic design, the Von Neumann Architecture

First Draft of a Report on the EDVAC
by John von Neumann,
Contract No. W-670-ORD-4926,
Between the United States Army Ordnance Department
and the University of Pennsylvania Moore School of Electrical Engineering
University of Pennsylvania
June 30, 1945



John von Neumann (1903 - 1957)



The Von Neumann Architecture

- Model for designing and building computers, based on the following three characteristics:

I

- The computer consists of four main sub-systems:

I

- Memory

II

- ALU (Arithmetic/Logic Unit)

III

- Control Unit

IV

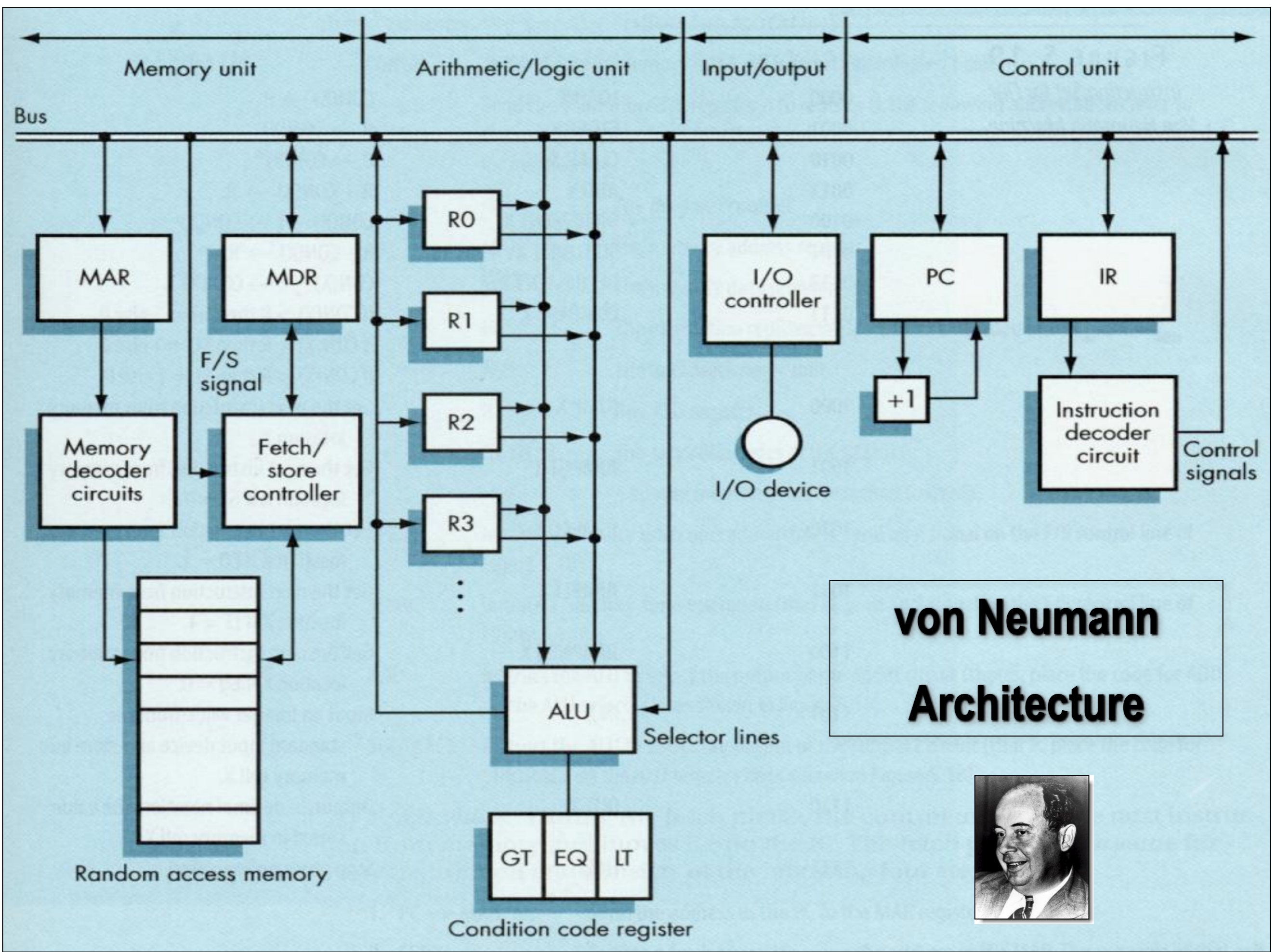
- Input/Output System (I/O)

2

- Program is stored in memory during execution.

3

- Program instructions are executed sequentially.



**von Neumann
Architecture**



What does it do?

- **Program Execution:**
 - **PC is set to the address where the first program instruction is stored in memory.**
 - **Repeat until HALT instruction or fatal error**

Fetch instruction

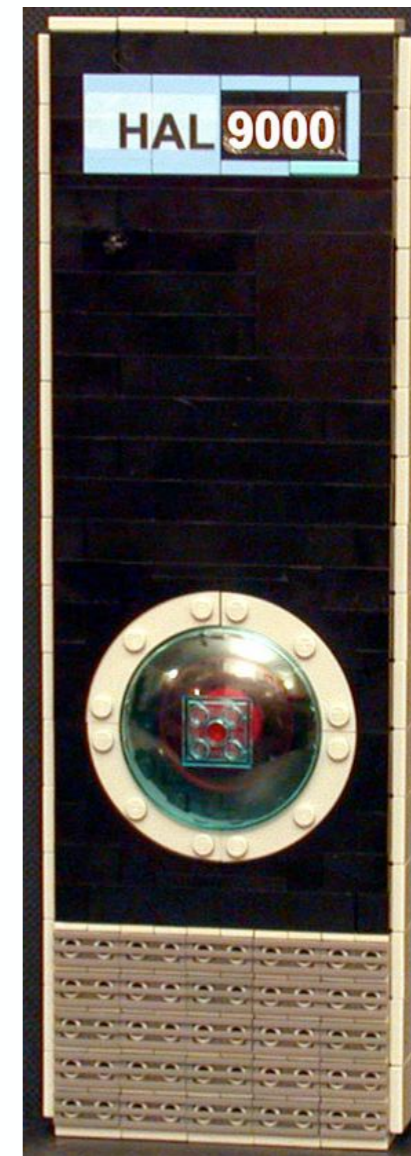
Decode instruction

Execute instruction

End of loop

The computer metaphor implies questions and a research program

- Memory
- Logic
- CPU
- Architecture



The symbol manipulation
paradigm

The Physical Symbol System
(PSS) hypothesis

The paradigm: problem solving

The General Problem Solver (GPS): Newell, Shaw & Simon - 1959

$$(R \supset \neg P) (\neg R \supset Q) \text{ --- } \neg (\neg Q P)$$

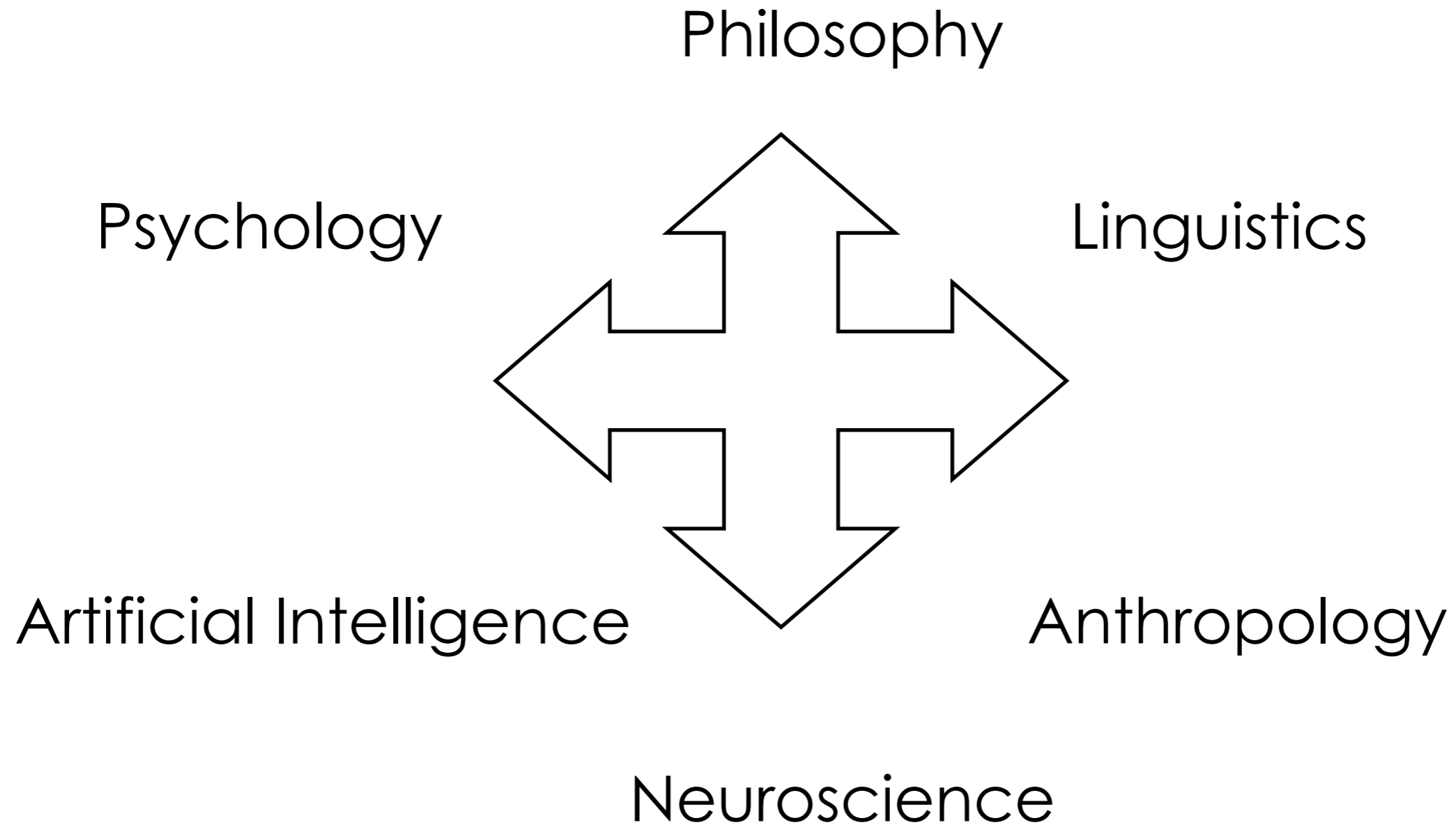
1	$A B \rightarrow B A$	8	$A B \rightarrow B$ applies to main expression only
	$A \vee B \rightarrow B \vee A$		$A B \rightarrow B$
2	$A \supset B \rightarrow \neg A \supset \neg B$	9	$A \rightarrow A \vee X$ applies to main expression only
3	$A A \leftrightarrow A$	10	$\left. \begin{array}{l} A \\ B \end{array} \right\} \rightarrow A \wedge B$
	$A \vee A \leftrightarrow A$		A and B are two main expressions
4	$A (B C) \rightarrow (A B) C$	11	$\left. \begin{array}{l} A \\ A \supset B \end{array} \right\} \rightarrow B$
	$A \vee (B \vee C) \rightarrow (A \vee B) \vee C$		A and $A \supset B$ are two main expressions
5	$A \vee B \leftrightarrow \neg (\neg A \neg B)$	12	$\left. \begin{array}{l} A \supset B \\ B \supset C \end{array} \right\} \rightarrow A \supset C$
6	$A \supset B \leftrightarrow \neg A \vee B$		$A \supset B$ and $B \supset C$ are two main expressions
7	$A (B \vee C) \leftrightarrow (A B) \vee (A C)$		
	$A \vee (B C) \leftrightarrow (A \vee B) (A \vee C)$		

Table 1.1: GPS: The 12 rules available to the subject and GPD to solve a logical puzzle.
 \wedge = "and", \vee = "or", \supset = "implies", \neg = "not".

The computer metaphor

- Memory
- CPU
- Architecture
- Functionalism
- Multi-instantiation

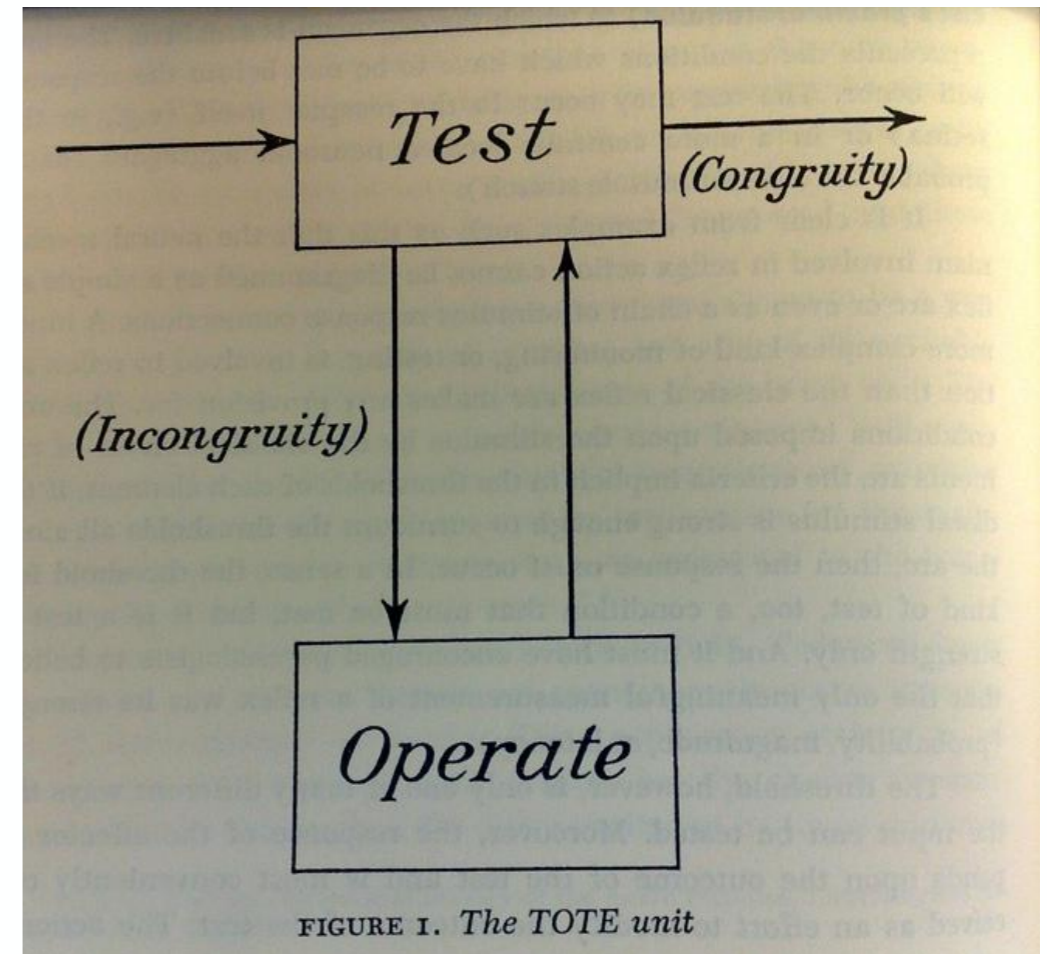
Cognitive Science



After Gardner (1985) The mind's new Science

Transformation: The TOTE Unit

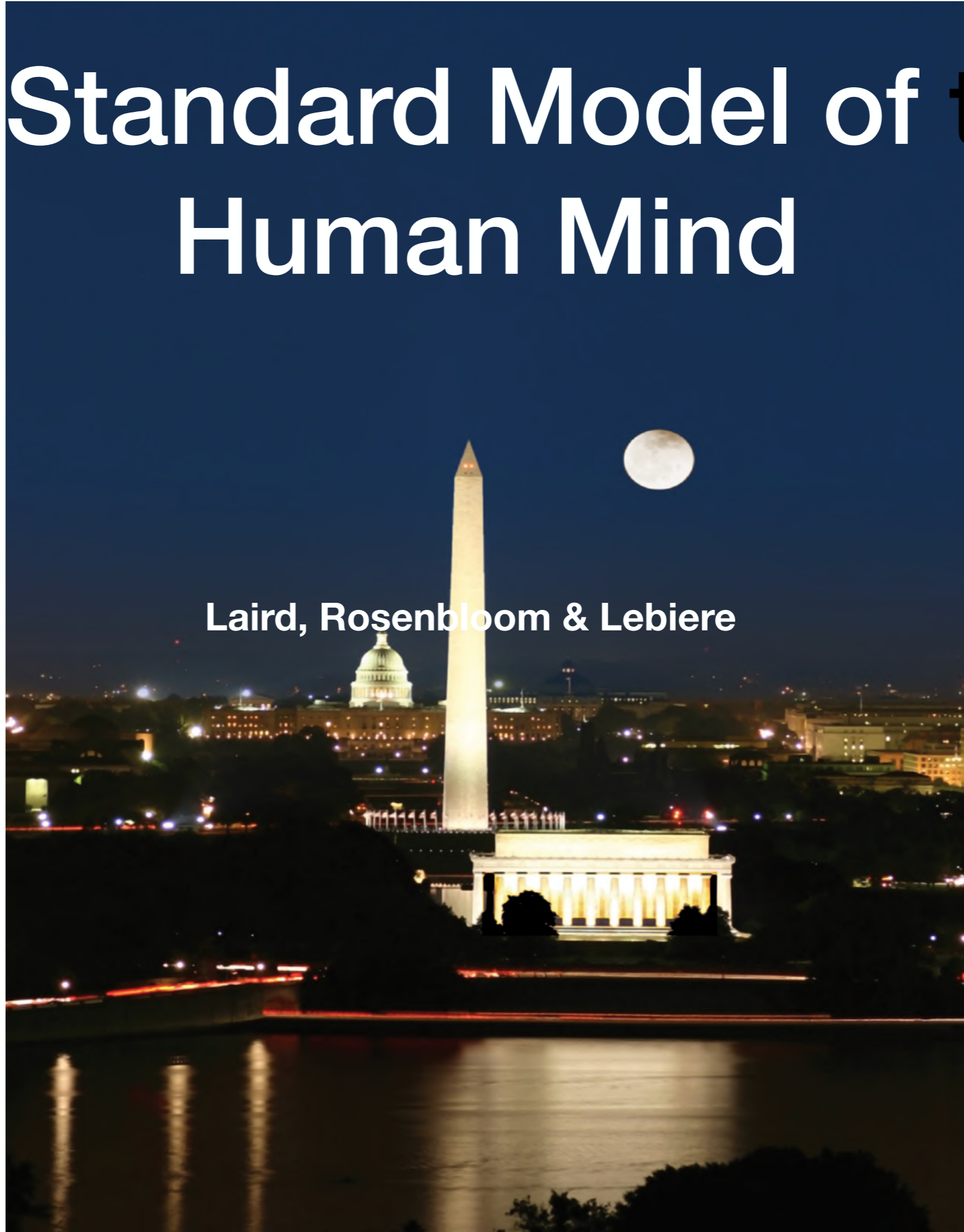
- Reflex is dead but what is the unit of behavior?
- A **plan** is any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed
- **Molar-Molecular** units: Strategy and tactics of behavior
- **Execution**: the plan is controlling the sequence of behavior
- **Image**: all organized knowledge about the world and the self



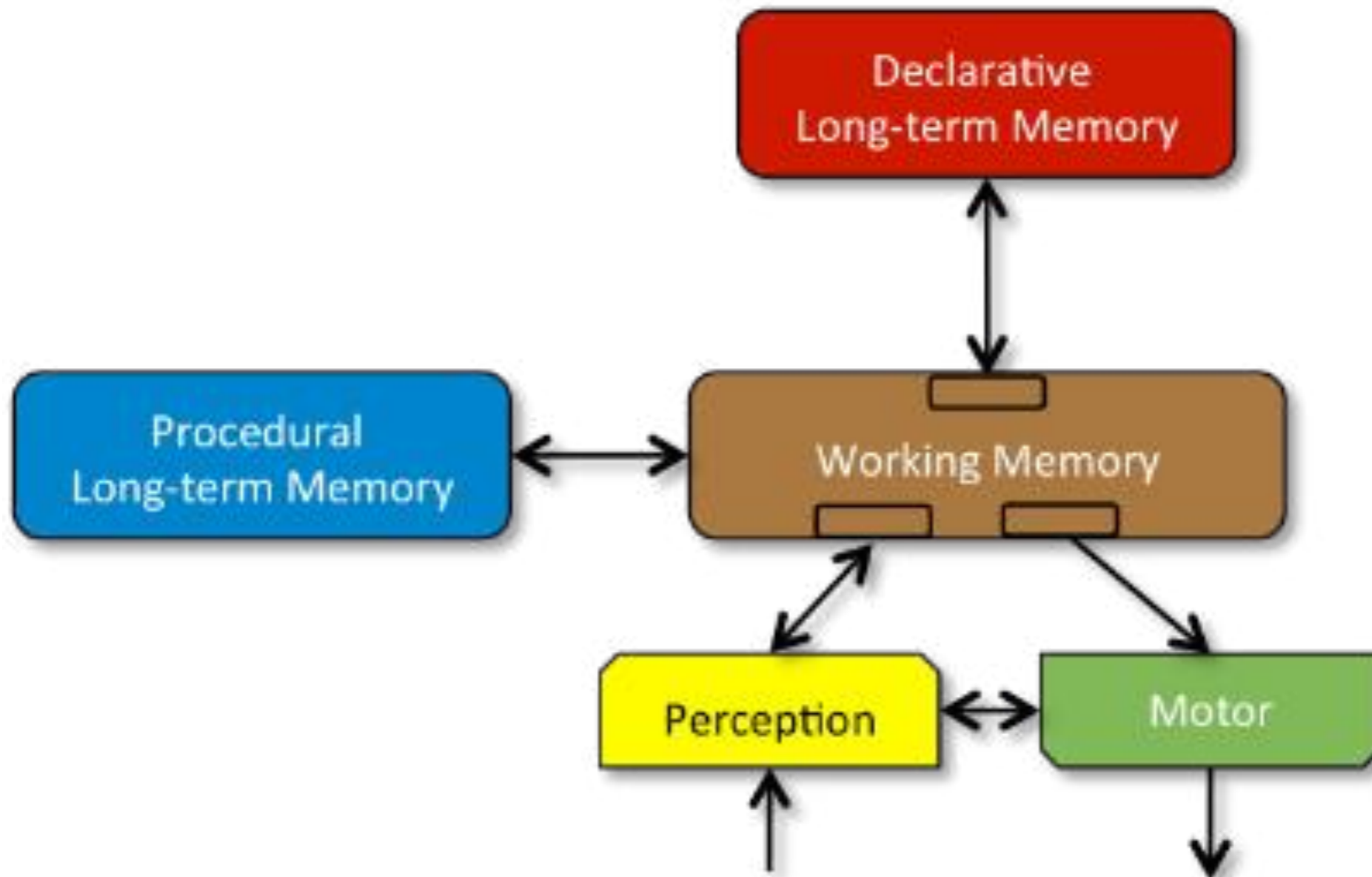
Plans and the Structure of Behavior
Miller, Gallanter & Pribram (1960)

A Standard Model of the Human Mind

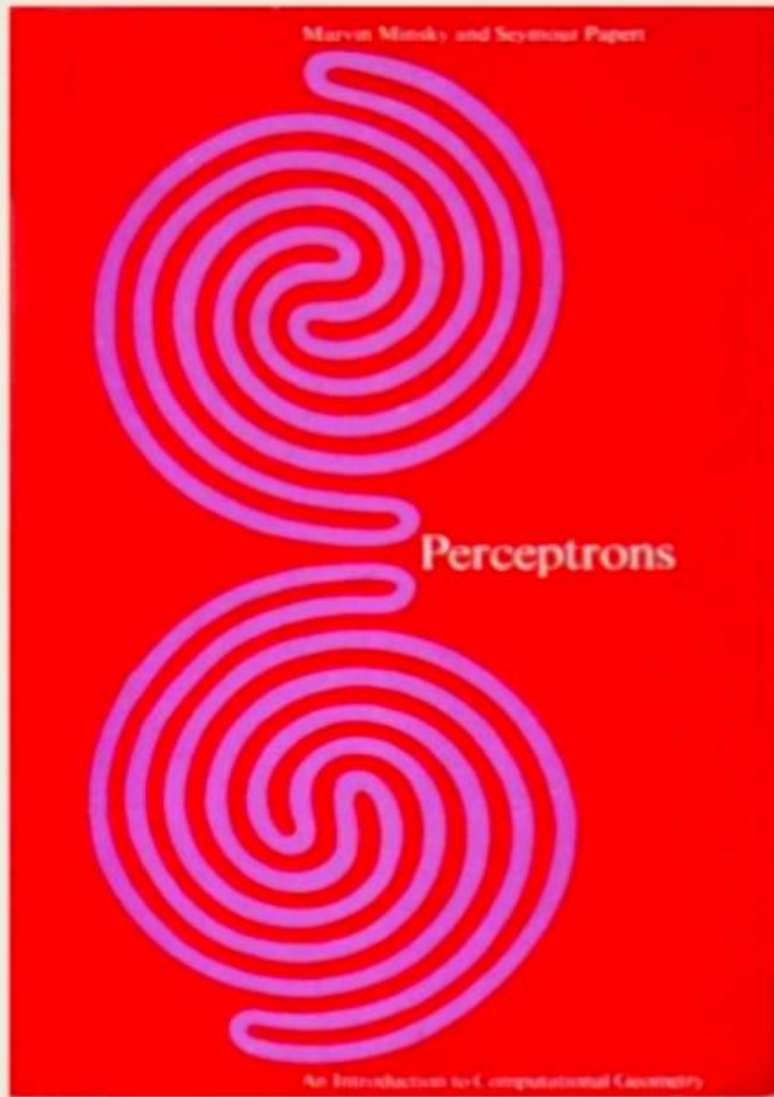
Laird, Rosenbloom & Lebiere



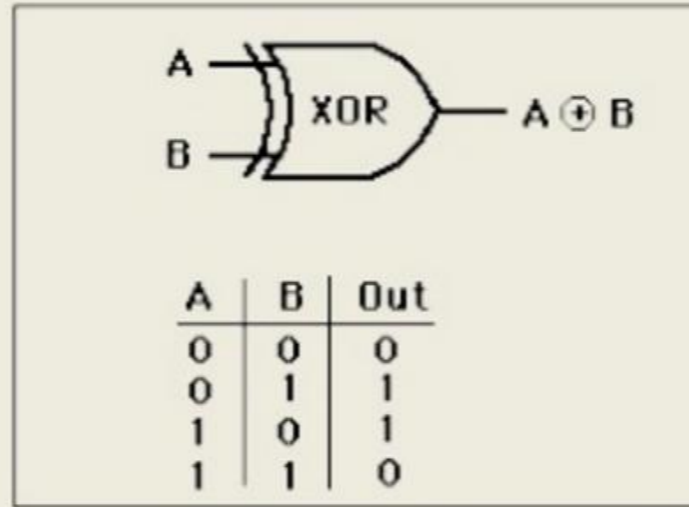
A Standard Model of the Human Mind



Blocking the Perceptron



<http://www.i-programmer.info/images/stories/BabBag/AI/book.jpg>

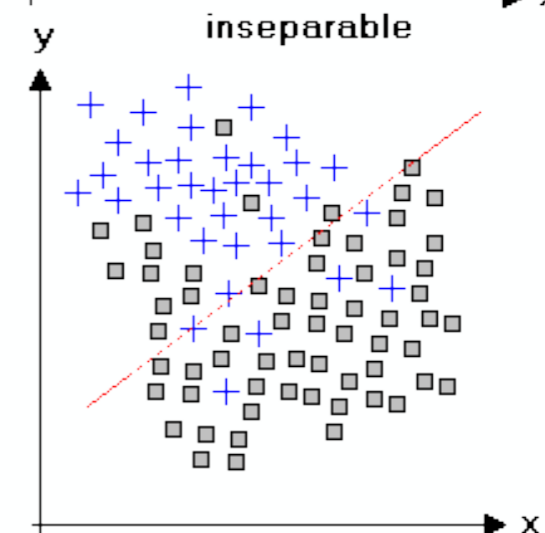
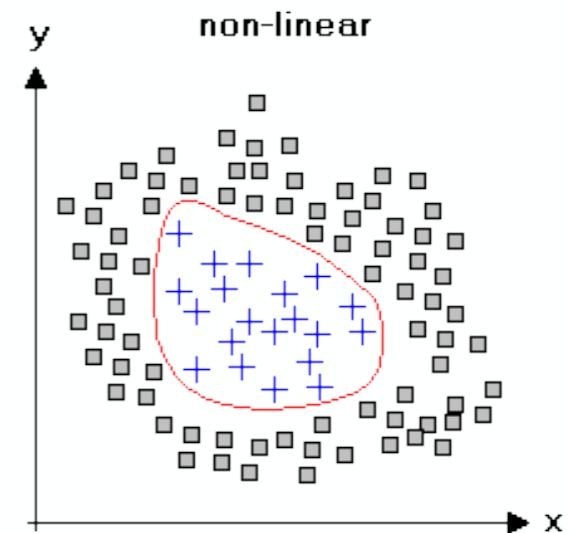
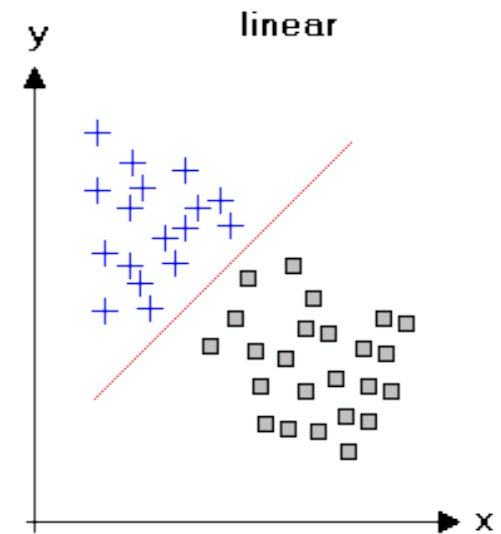


<http://hyperphysics.phy-astr.gsu.edu/hbase/electronic/ietron/xor.gif>

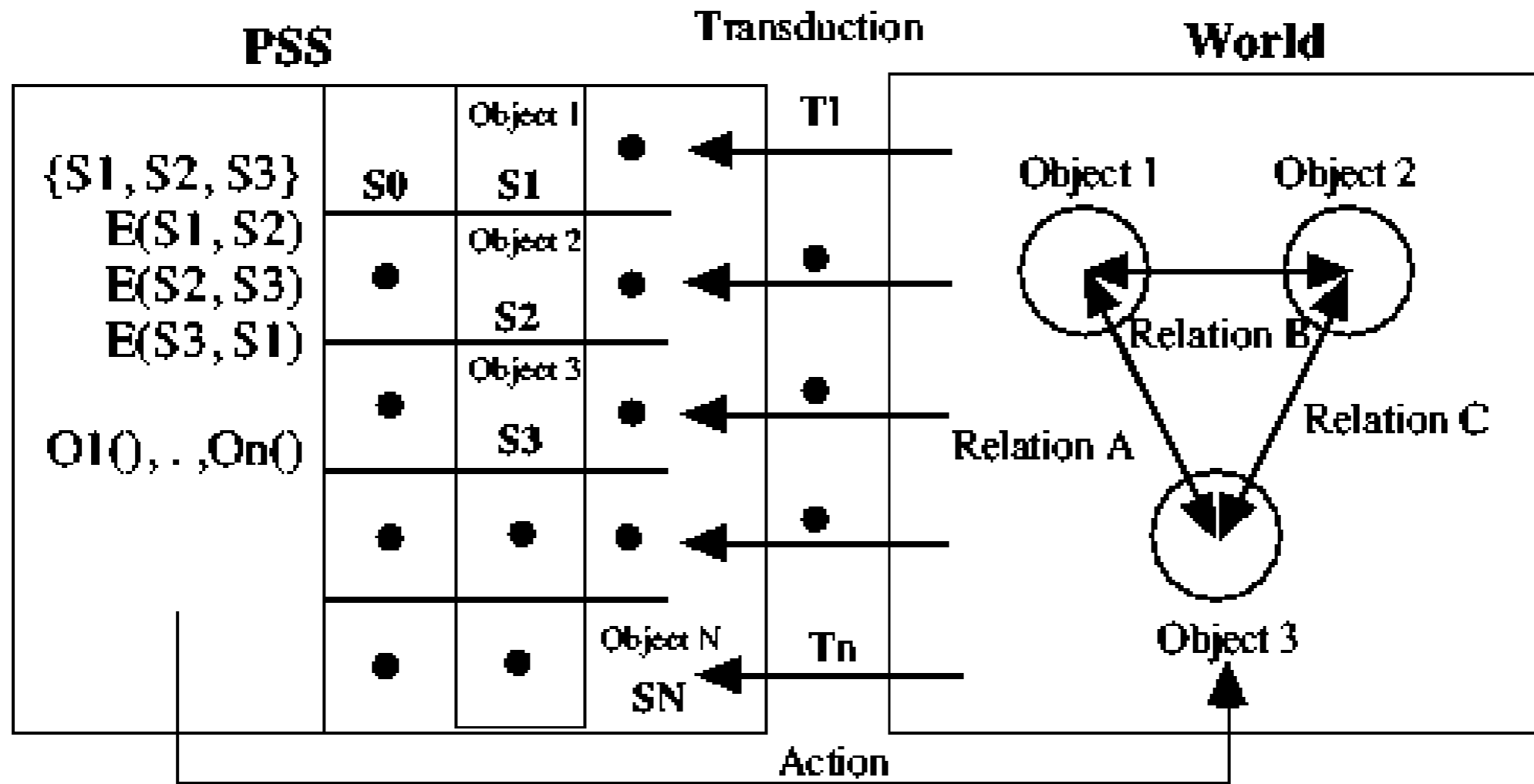


Minsky & Papert

<https://constructingkids.files.wordpress.com/2013/05/minsky-papert-71-csolomon-x640.jpg>



PSS: Newell & Simon 1976



T: Transduction, S: Symbol, E: Expression, O: Operator

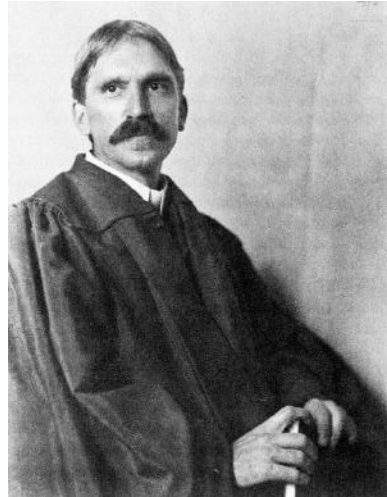
Functionalism

- **Behaviorism** ... attempts to explain behavior without any reference whatsoever to mental states and processes
- **Functionalism** in the philosophy of mind is the doctrine that what makes something a mental state of a particular type does not depend on its internal constitution, but rather on the way it functions, or the role it plays, in the system of which it is a part. (<http://plato.stanford.edu/entries/functionalism/>)

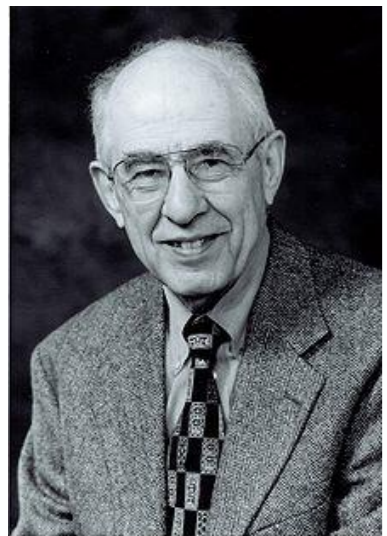
Theory of Mind: Functionalism

- Mental states are **functional states** and are not revealed in terms of intrinsic physical features.
 - *What really matters is how mental state instances are causally arranged: what cause them, and what they, in turn, cause.*
- **Multiple realizability:** mental states are not limited to a particular medium, they can be realized in multiple ways (i.e., computers) provided that the proper functional roles are realized (H. Putnam & J. Fodor)
- **Strong Artificial Intelligence** and instantiation of intelligence
- **Brain-computer metaphor:** mind as a Turing machine

Some roots:



- Back to **J. Dewey** (1859 – 1952)
 - *The Reflex Arc Concept in Psychology* (1896)

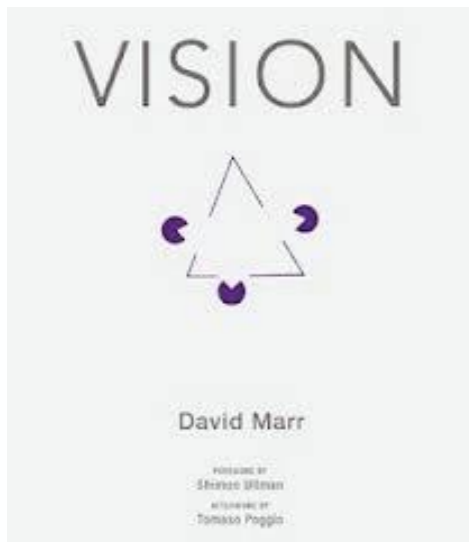


- **H. Putnam** (1926 -)
 - “Minds and Machines” in *Dimensions of the Mind* (1960) **machine-state functionalism**, was the first to argue that *minds* are things that we can conceive *solely* in terms of input, output, and various **functional** relations

Some roots:



- **Jerry Fodor (1935 -)**
 - General multiple realizability as a critique to reductionism
 - *The Modularity of Mind: An Essay on Faculty Psychology*, MIT Press, 1983
 - ...not to be intended in Gall's terms of physical localizability
 - Modules are innately specified systems that take in sensory inputs and yield necessary representations of them (i.e. visual system cognition, language,)



- **David Marr (1945-1980)**
 - Visual system characterization
 - Theory of the Cerebellar Cortex (1969). The simple and regular cortical structure is interpreted as a simple but powerful memorizing device for learning motor skills”
 - Levels of description:

Three levels of description (David Marr, 1982)

Computational

Why do things work the way they do?
What is the goal of the computation?
What are the unifying principles?

Algorithmic

What representations can implement such computations?
How does the choice of representations determine the algorithm?

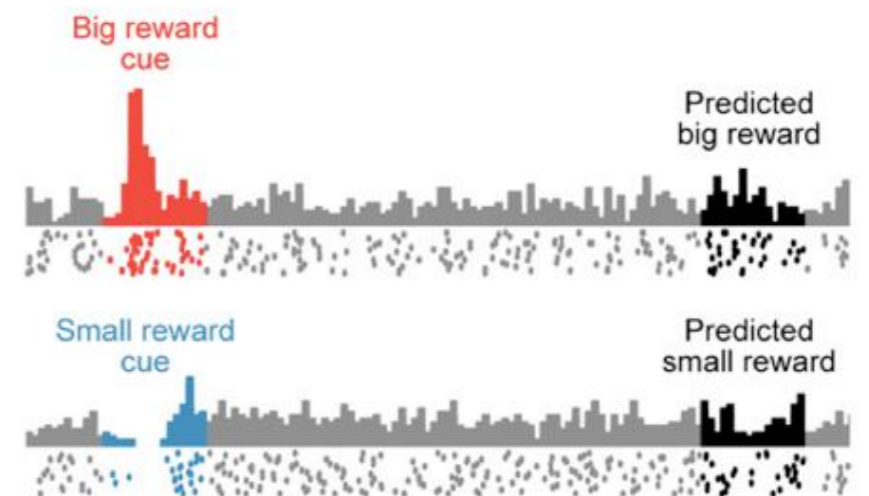
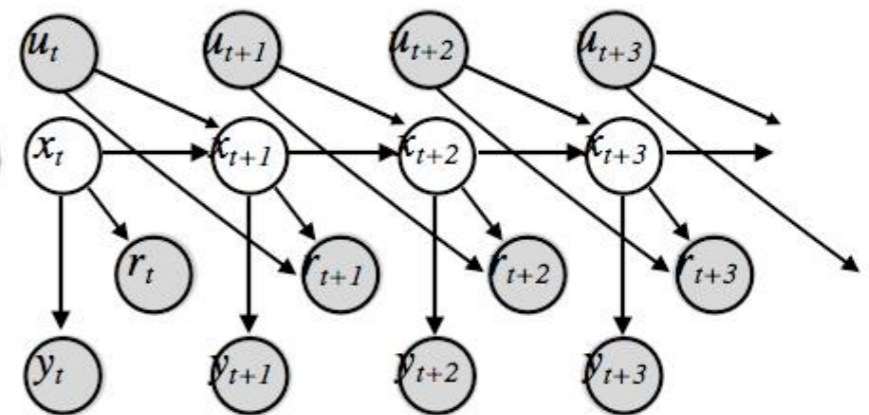
Implementational

How can such a system be built in hardware?
How can neurons carry out the computations?



maximize:

$$R_t = r_{t+1} + r_{t+2} + \dots + r_T$$



Functionalism

Things are defined by their functions

Two ways to define function

Function = inputs and outputs (machine functionalism)

e.g. mathematical function, e.g. +, -, x, /

$2 \times 3 = 6$, when input is 2 and 3, output is 6

Multiple realizability/ Multi-Instantiation: Function can be realized in different materials or through different processes

Functional definition of mind

- If x acts like a mind, it is a mind.
- If, when compared to a mind, given similar inputs, x gives similar outputs, x is a mind.
- If a computer can converse (take part in linguistic input and output exchanges/play the role of an intelligent conversational partner) just like a person, the computer is as intelligent as a person. It has a mind.

Good Old Fashioned AI GOFAI

- The paradigmatic expression of Rationalism

“For the same are the thinking and the
being”

(Parmenides, 480BC, Fragment V)

“Computing Machinery and Intelligence” (Mind, 1950)

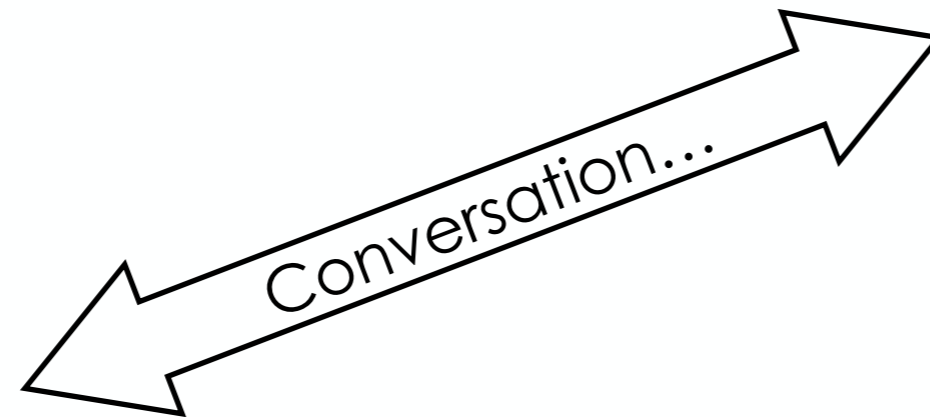
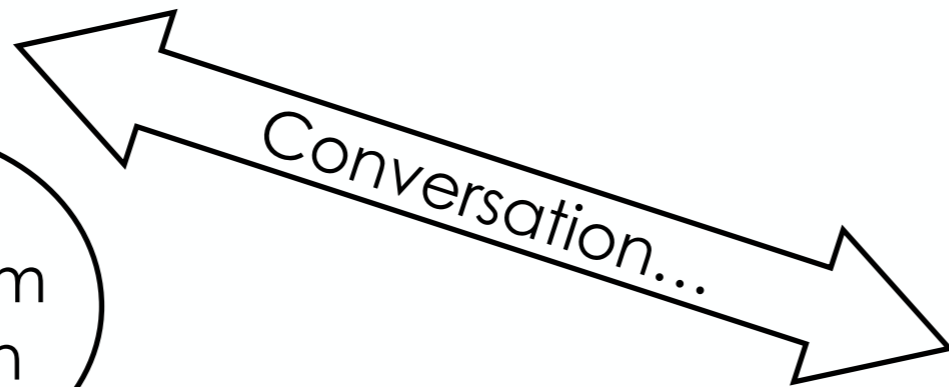
- “I propose to consider the question, ‘Can machines think?’”



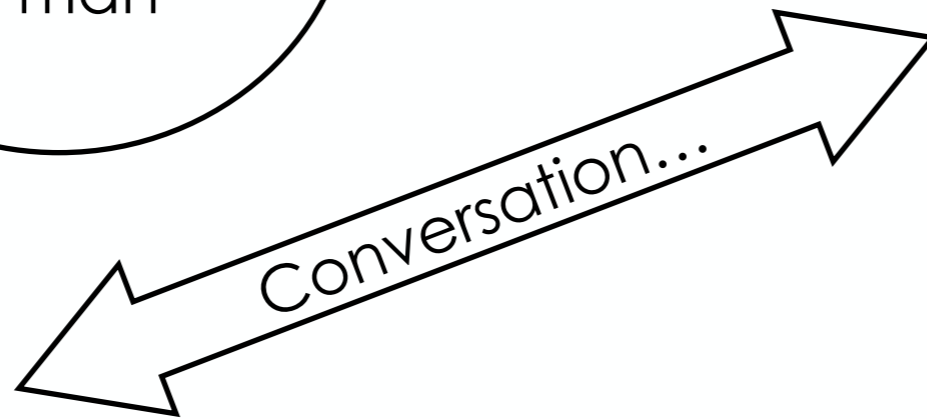
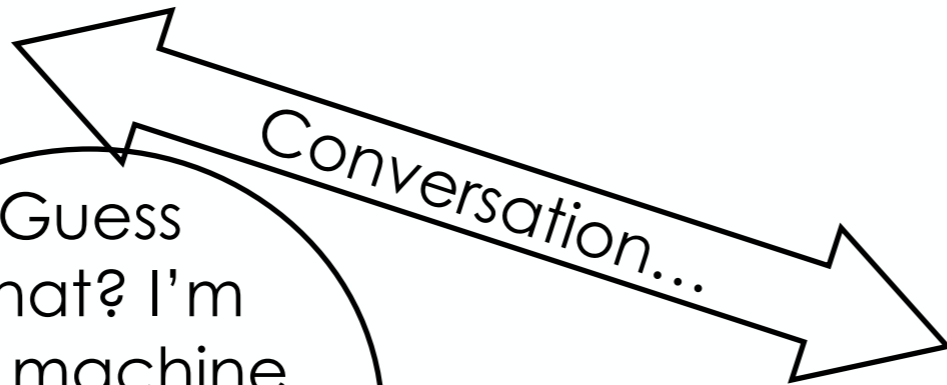
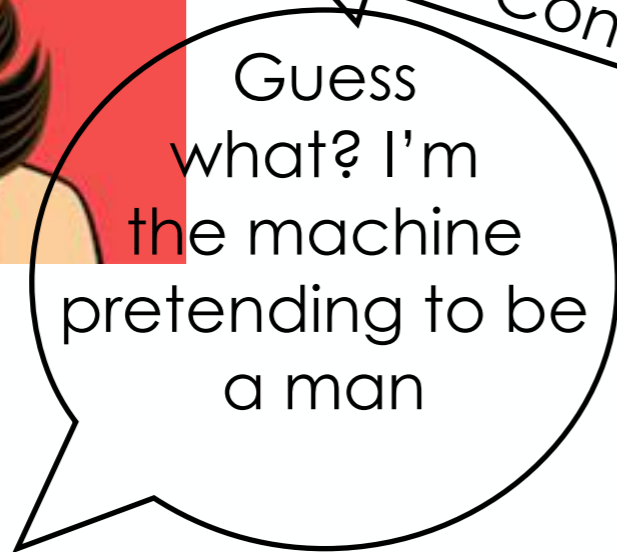
Alan M. Turing (1912–1954)

The imitation game or the Turing Test

The Imitation Game



The Immitation Game



Predictions

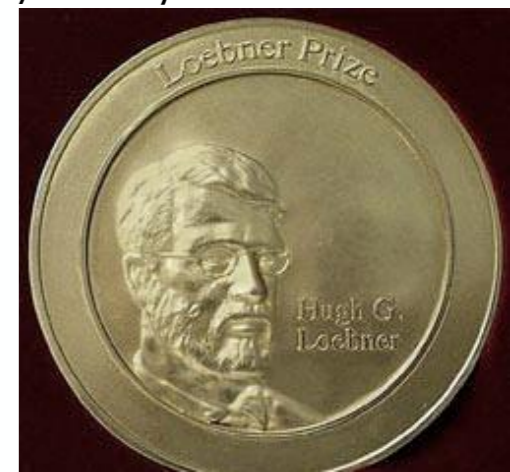
- In 1950, Turing predicted that 50 years later it will be possible to program a computer with ~100 Mb memory to pass TT 70% of the time with 5 minute conversations.
- It will be natural to speak of computers 'thinking'.
- "The machine may be used to help in making up its own programmes, or to predict the effect of alterations in its own structure."
- "We may hope that machines will eventually compete with men in all purely intellectual fields."

Current TT status

- To date no computer has passed TT.
- Loebner Prize: annual TT competition, prizes for passing TT and for “most human conversation”.
 - <http://www.loebner.net/Prizef/loebner-prize.html>
 - No prize awarded for passing TT yet.



- Long Bets Foundation has registered a \$10,000 bet b/w Mitch Kapor and Ray Kurzweil (Wired magazine, Issue 10.05 | May 2002):
 - Will a computer pass a TT by 2029?



Do computers think?



- Interview with Gary Kasparov's advisor Frederic Friedel after Kasparov's loss to Deep Blue:
 - MARGARET WARNER: All right. Let me bring Mr. Friedel back in here. Mr. Friedel, did Gary Kasparov think the computer was thinking?
 - FREDERIC FRIEDEL: Not thinking but that it was showing intelligent behavior. When Gary Kasparov plays against the computer, he has the feeling that it is forming plans; it understands strategy; it's trying to trick him; it's blocking his ideas, and then to tell him, now, this has nothing to do with intelligence, it's just number crunching, seems very semantic to him. He says the performance is what counts. I see it behaves like something that's intelligent. If you put--if you put a curtain up, he plays the game and then you open the curtain, and it's a human being. He says, ah, that was intelligent, and if it's a box,

he says, no, that was just number crunching. It's the performance he's interested in

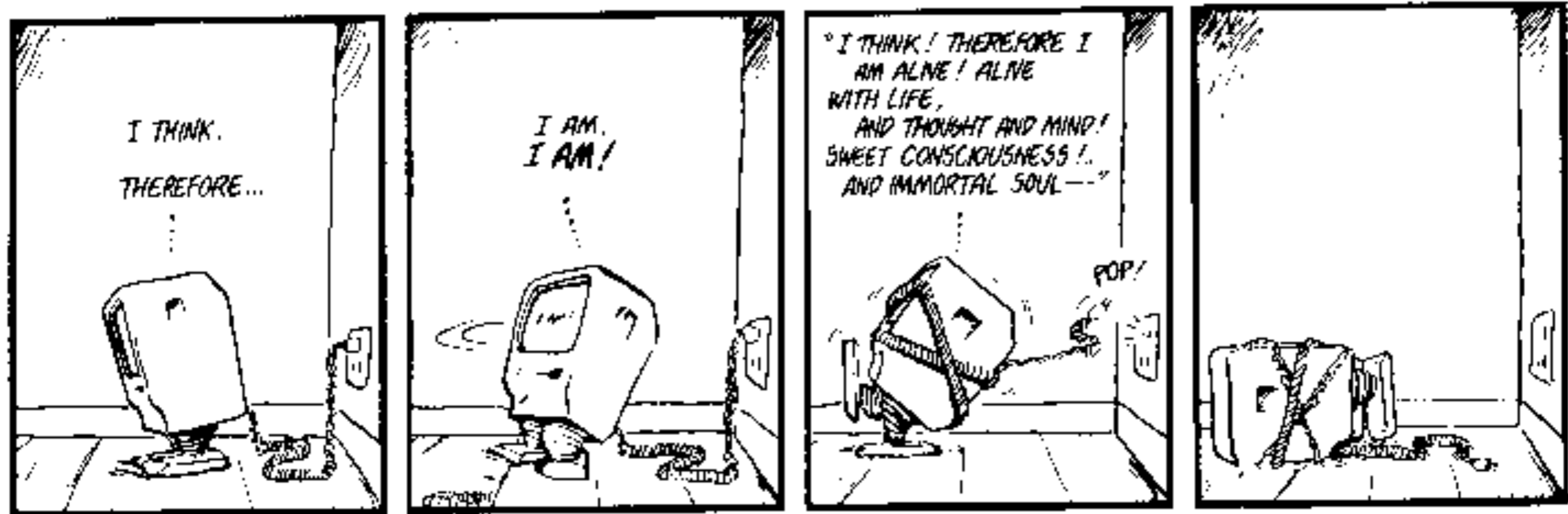
“We can only see a short distance ahead, but we can see plenty there needs to be done.”

-Alan M. Turing

- cognitive “revolution”
- artificial intelligence
- artificial life
- connectionism
- new AI, behavior based robotics

Metaphor, hype and illusions

Bloom County on Strong AI



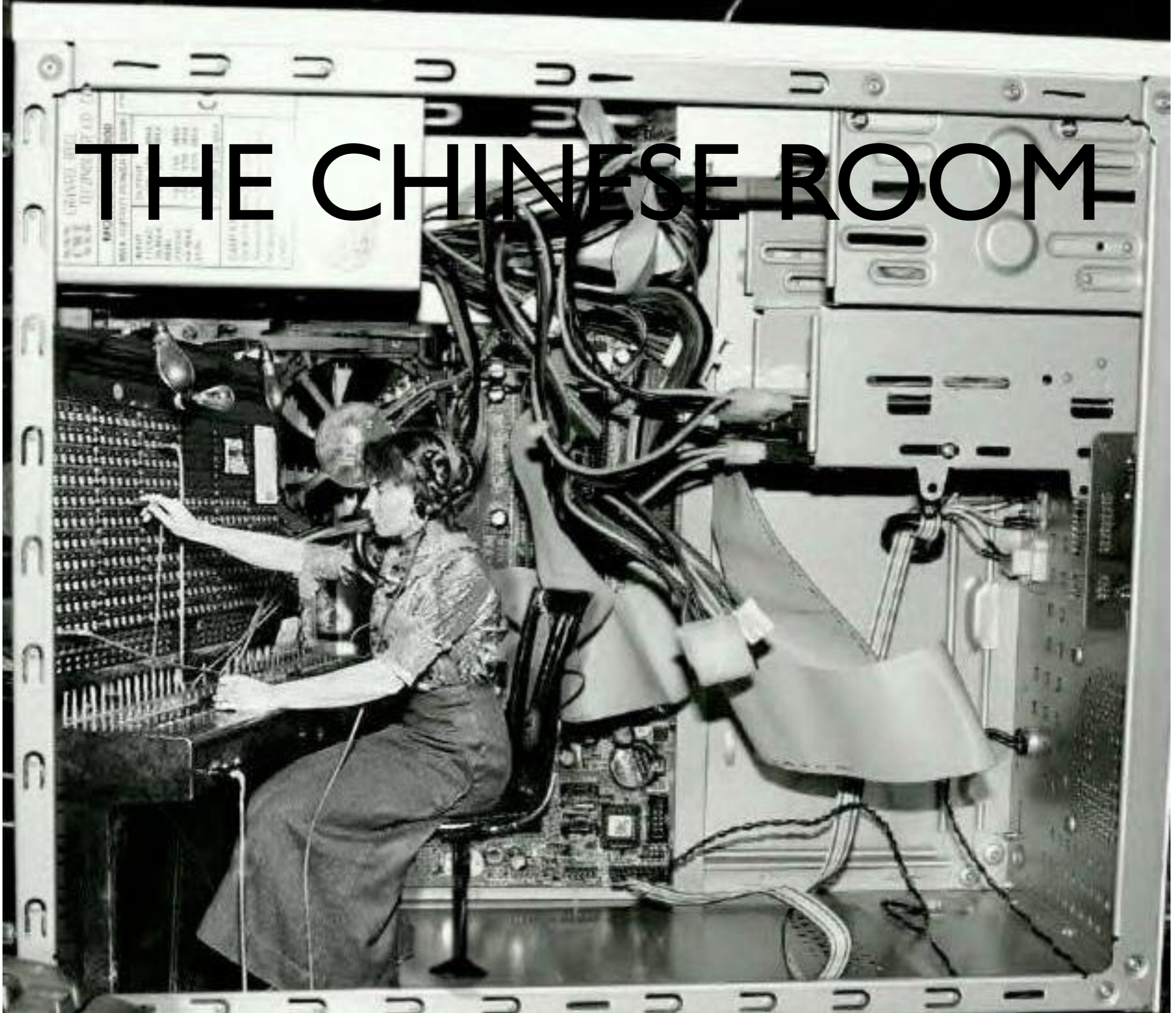
Some problems of GOFAI

- Symbol grounding (Searle, 1980; Harnad, 1990)
- where does meaning come from?

Some problems of GOFAI

- Frame problem (McCarthy & Hayes, 1969; Pylyshyn, 1987)
- How can search be constrained to be feasible in a reasonable time window?

THE CHINESE ROOM



The Gedankenexperiment

- Searle, who knows no Chinese, is locked in a room with an enormous batch of Chinese script.



The Gedankenexperiment

- Slips of paper with still more Chinese script come through a slot in the wall.
- He has to transcribe input to output following a rule book



The result

- But Searle is behaving just as a computer does, “performing computational operations on formally specified elements”



I don't think, therefore I am not.

- Hence, manipulating formal symbols—which is just what a computer running a program does—is not sufficient for understanding or thinking.

IN THIS BUILDING DURING THE SUMMER OF 1956

JOHN McCARTHY (DARTMOUTH COLLEGE), MARVIN L. MINSKY (MIT)
NATHANIEL ROCHESTER (IBM), AND CLAUDE SHANNON (BELL LABORATORIES)
CONDUCTED

THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

FIRST USE OF THE TERM "ARTIFICIAL INTELLIGENCE"

FOUNDING OF ARTIFICIAL INTELLIGENCE AS A RESEARCH DISCIPLINE

"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."

IN COMMEMORATION OF THE PROJECT'S 50th ANNIVERSARY
JULY 13, 2006

An MTV history of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1952–69 Look, Ma, no hands!
- 1950s Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1956 Dartmouth meeting: "Artificial Intelligence" adopted
- 1965 Robinson's complete algorithm for logical reasoning
- 1966–74 AI discovers computational complexity
Neural network research almost disappears
- 1969–79 Early development of knowledge-based systems
- 1980–88 Expert systems industry booms
- 1988–93 Expert systems industry busts: "AI Winter"
- 1985–95 Neural networks return to popularity
- 1988– Resurgence of probabilistic and decision-theoretic methods
Rapid increase in technical depth of mainstream AI
"Nouvelle AI": ALife, GAs, soft computing

Some problems of GOFAI

- Situatedness (Simon, 1969; Suchman, 1987)
- Agents are in the world, which can serve as its own representation

Some problems of GOFAI

- Frame of reference problem (Clancey, 1989)
- Ontology of knowledge defined by the perspective of agent, designer or observer

Some problems of GOFAI

- Symbol grounding (Searle, 1980; Harnad, 1990)
- Frame problem (McCarthy & Hayes, 1969; Pylyshyn, 1987)
- Situatedness (Simon, 1969; Suchman, 1987)
- Frame of reference problem (Clancey, 1989)

Constraints on Theories of M/B

Criteria for Unified Theories of Cognition (UTC).:

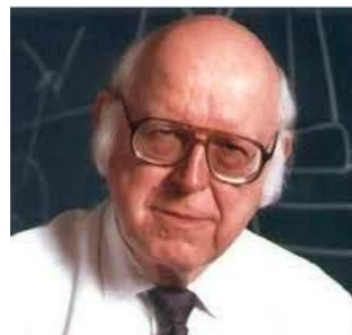
- 1) Behave flexibly as a function of the environment;
 - 2) Exhibit adaptive (rational, goal-oriented) behavior;
 - 3) Operate in real-time;
 - 4) Operate in rich, complex, detailed environments;
 - 5) Use symbols and abstractions;
 - 6) Use language;
 - 7) Learn from the environment and from experience;
 - 8) Acquire capabilities through development;
 - 9) Operate autonomously, but within a social community;
 - 10) Be self-aware and have a sense of self;
 - 11) Be realizable as a neural system;
 - 12) Be constructible by an embryological growth process;
 - 13) Arise through evolution
- (Newell, 1994, p.19).



Constraints on Theories of M/B

Criteria for Unified Theories of Cognition (UTC):

- 1) Behave flexibly as a function of the environment;
- 2) Exhibit adaptive (rational, goal-oriented) behavior;
- 3) Operate in real-time;
- 4) Operate in rich, complex, detailed environments;
- 5) Use symbols and abstractions;
- 6) Use language;
- 7) Learn from the environment and from experience;
- 8) Acquire capabilities through development;
- 9) Operate autonomously, but within a social community;
- 10) Be self-aware and have a sense of self;
- 11) Be realizable as a neural system;
- 12) Be constructible by an embryological growth process;
- 13) Arise through evolution (Newell, 1994, p.19).



1: Functional constraints (Psychology of mind):

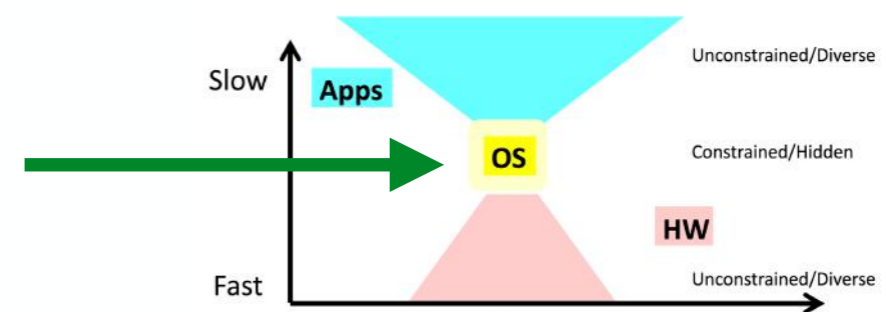
Level 1: Display autonomous adaptive and flexible real-time goal-oriented behavior in complex **physical** environments (Newell test: 1, 2, 3, 4, 7, 10-sense of self);

Level 2: Display autonomous adaptive and flexible real-time goal-oriented behavior in complex real-world **social** environments including the use of symbols and language (Newell test: Level 1 + 5, 6, 9, 10-self-aware);

2: Structural constraints (Biology of embodied brain):

Biological validity: be plausibly the product of biological evolution and be demonstrably constructible through neuro- and morphogenesis (Newell test: 11-13)

Physical realizability: perform in real-time, in the real-world using resources (e.g. energy, computation) comparable to biological systems.



“Solutions”

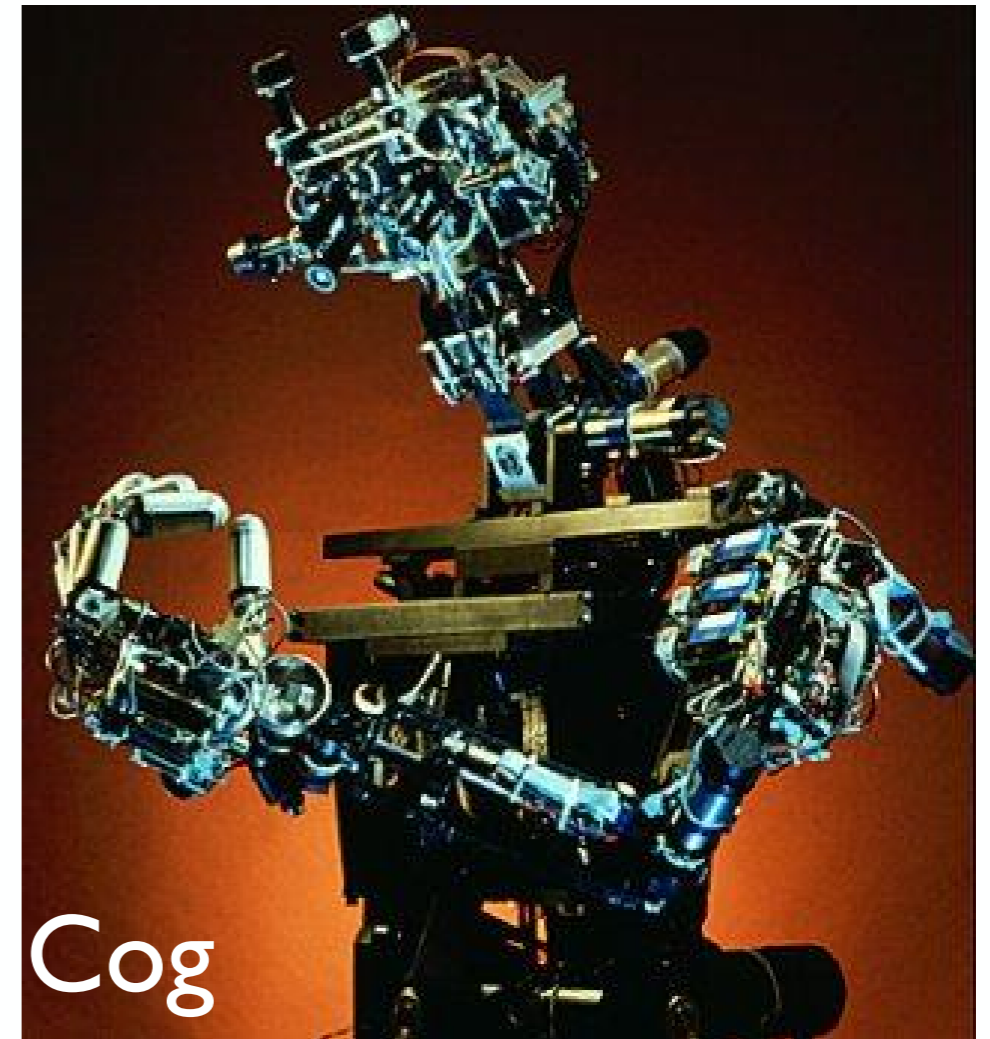
- Connectionism
- Behavior based AI / New AI

New AI: Subsumption Architecture

Rod Brooks (MIT, AI lab)



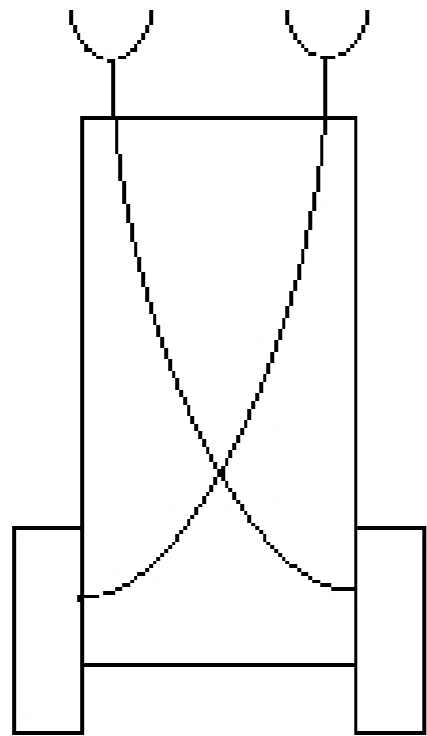
Intelligent behavior:
without explicit representations
without explicit abstract reasoning



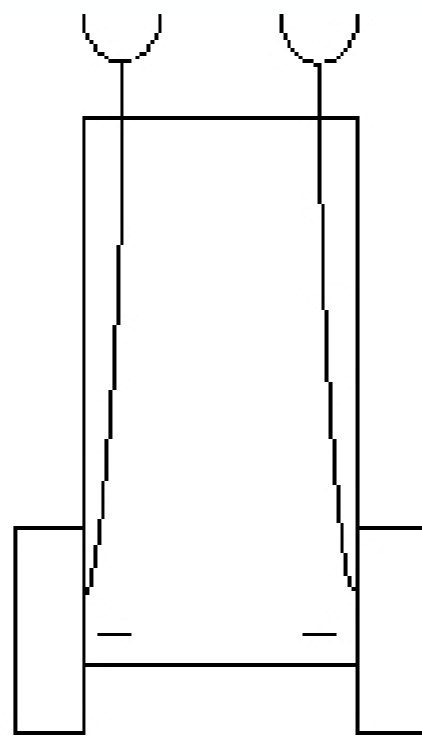
Intelligence is an *emergent* property of certain
complex systems

Braitenberg vehicles

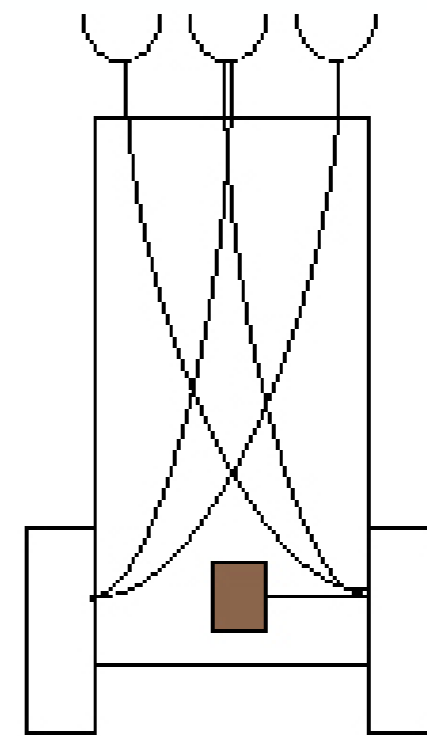
- Construct complex behaviors from simple interactive rules



Vehicle 2b



Vehicle 3a



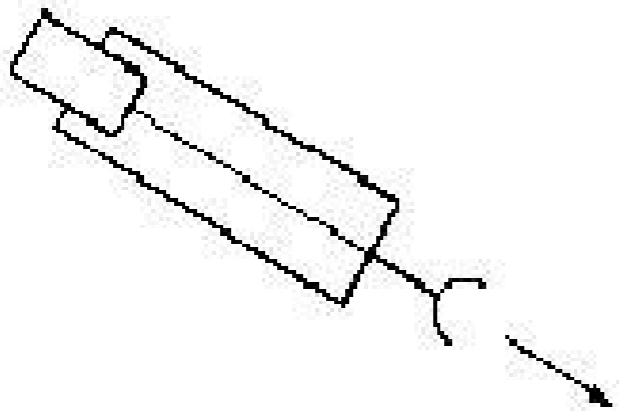
Red neurotic



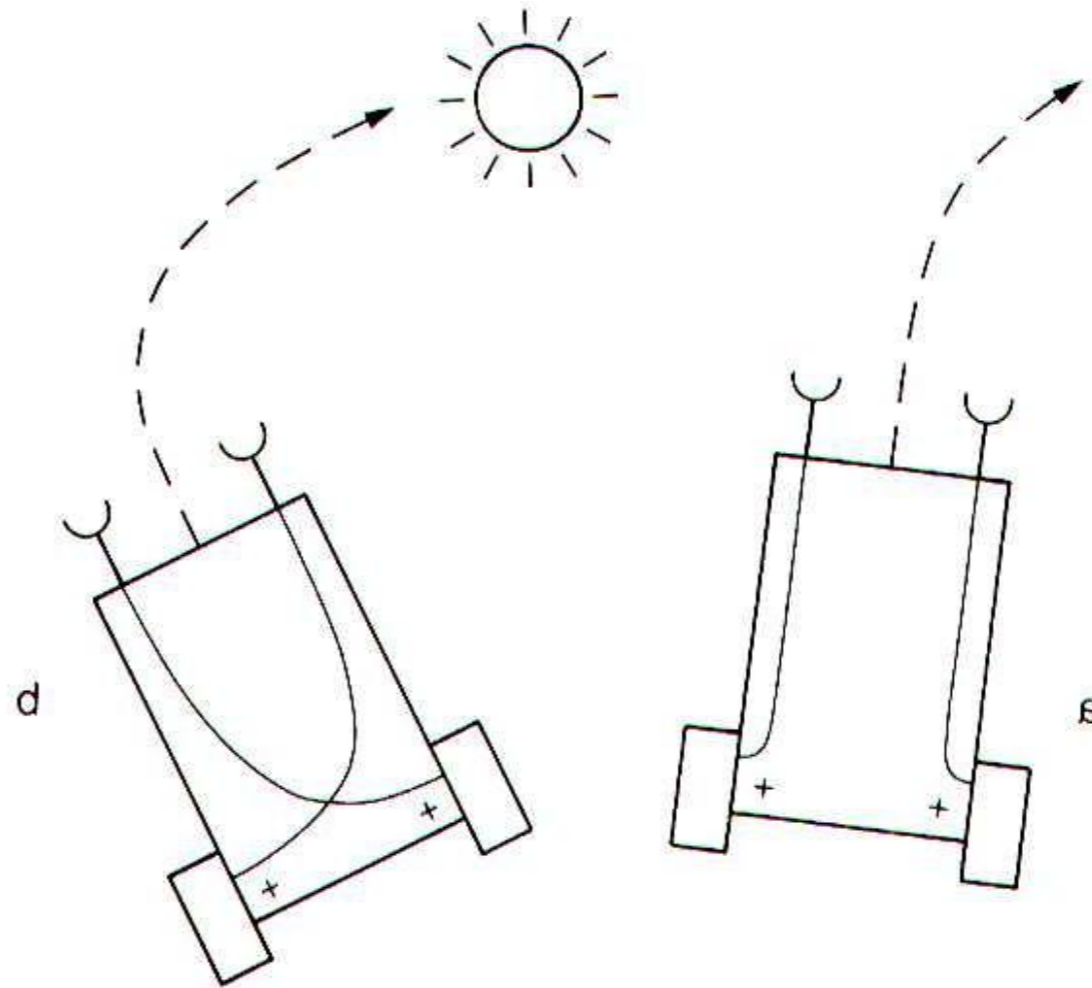
Braitenberg, 1984



Braitenberg behavior



1 sensor 1 motor



Love & Hate

Advantages of Reactive Agents

- Simplicity
- Economy (computation, communication)
- Computational tractability
- Robustness against failure
- Elegance

Problems of Reactive Agents

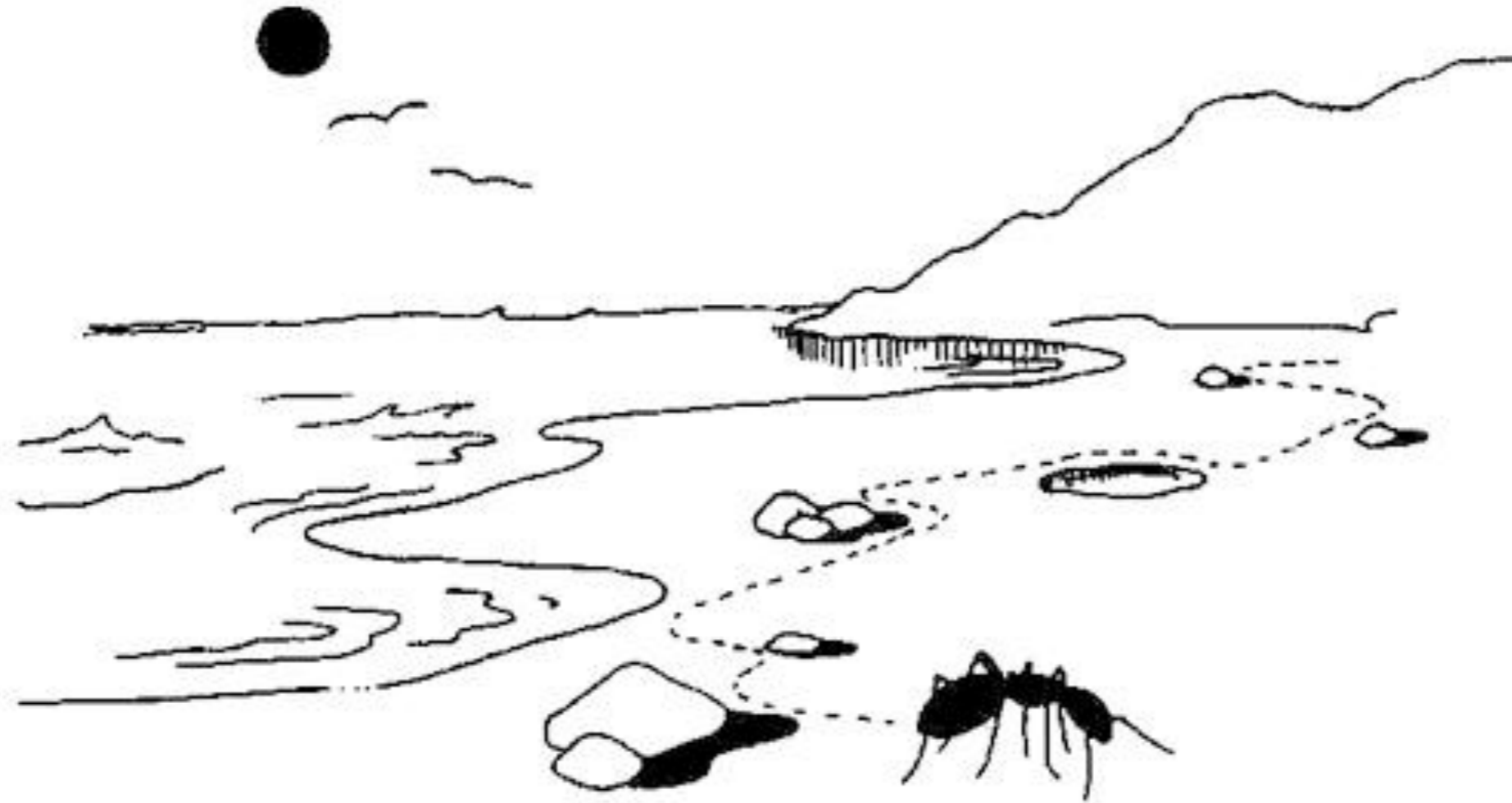
- Does direct sensing tell you enough?
- How about the non-local spatio-temporal organization of behavior?
World models?
- How about unpredictability & learning?
- How to be smarter than your designer?
- Believe in magic: How to harness emergence?
- Is it not just like Behaviorism?
 - No: worries about internal mechanism
- Is it not just like Cybernetics?
 - Yes, minus the sophistication

The dogma's of NEW AI

Situatedness and embodiment: 'Real' intelligence is situated in the world, not in disembodied systems such as theorem provers or expert systems

Intelligence, frame of reference and emergence: 'Intelligent' behavior arises as a result of an agent's interaction with its environment. Also, intelligence is 'in the eye of the beholder'; it is not an innate, isolated property

The ant on the beach



Simon (1969) Sciences of the Artificial

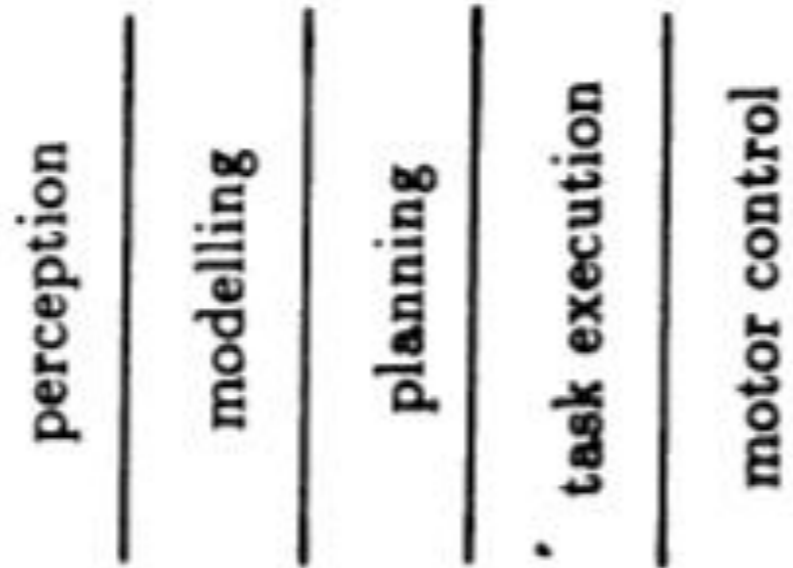
The subsumption architecture

- A hierarchy of sense-act relations:
 - Lower levels more primitive
 - Lower layers have precedence
- Each behavior is a rather simple rule-like structure (finite state machine)
- Each behavior ‘competes’ with others to exercise control over the behaving agent

From the sense-think-act cycle
to
sense- "think" - act hierarchies

GOFAI

Sensors

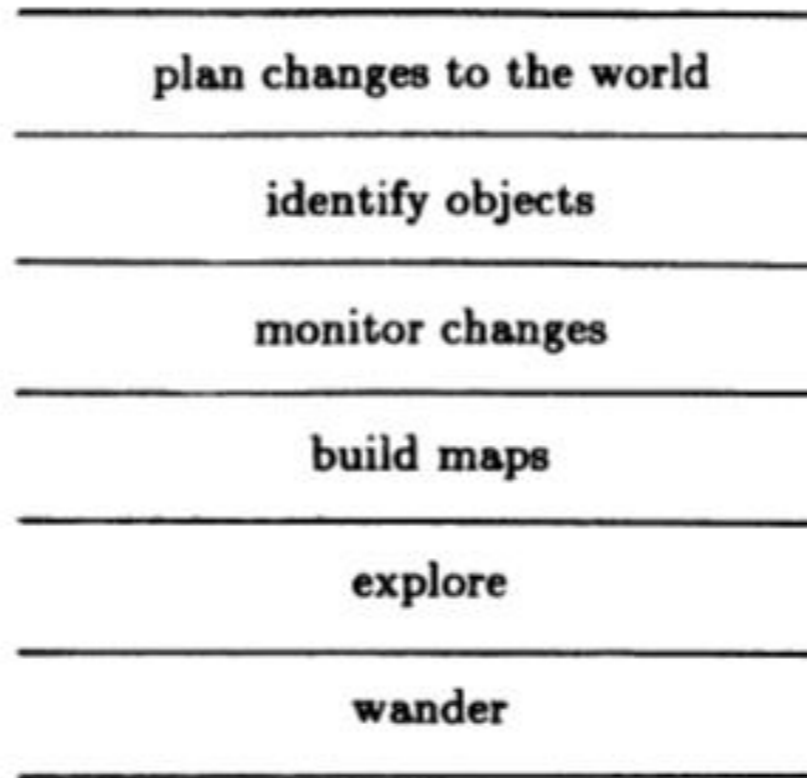


Actuators

reason about behavior of objects

New AI

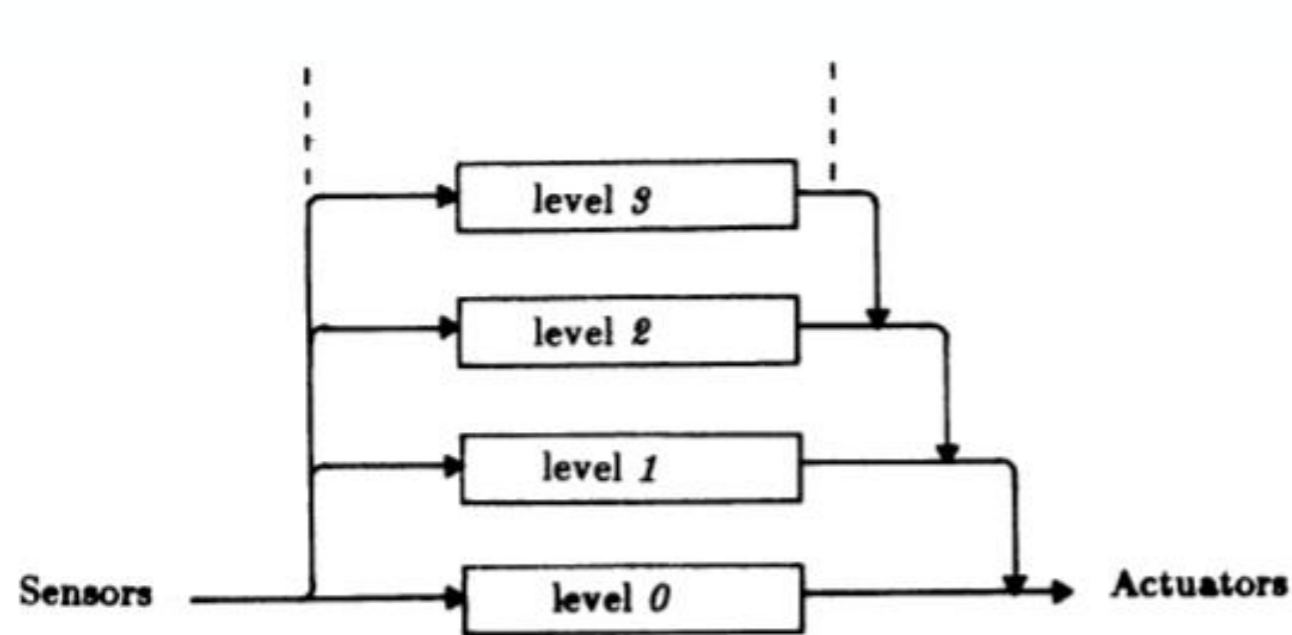
Sensors



Actuators

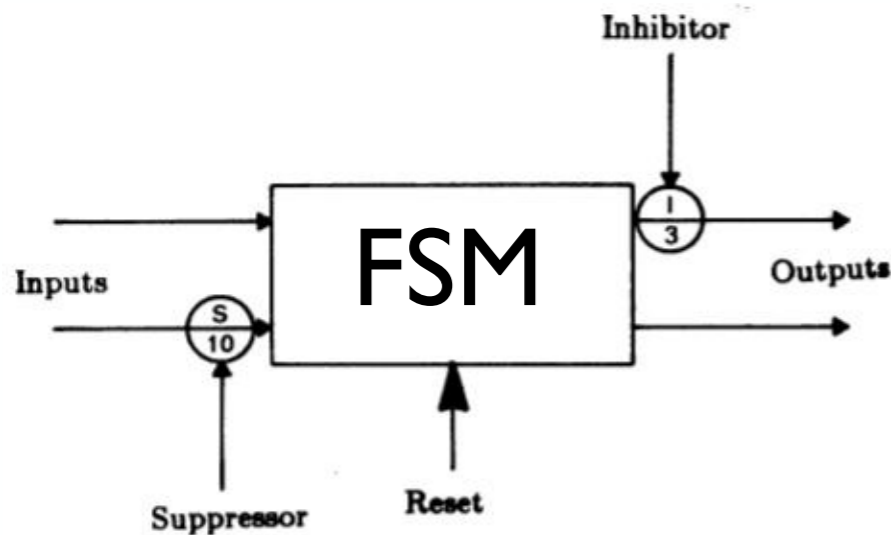
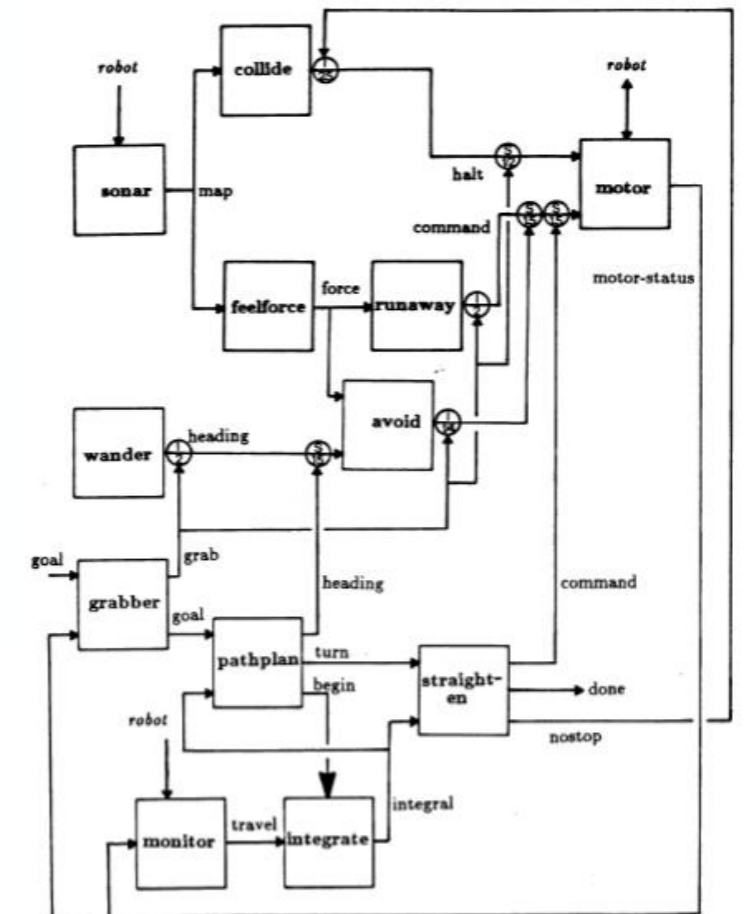
avoid objects

SA: Layered Control and encapsulation

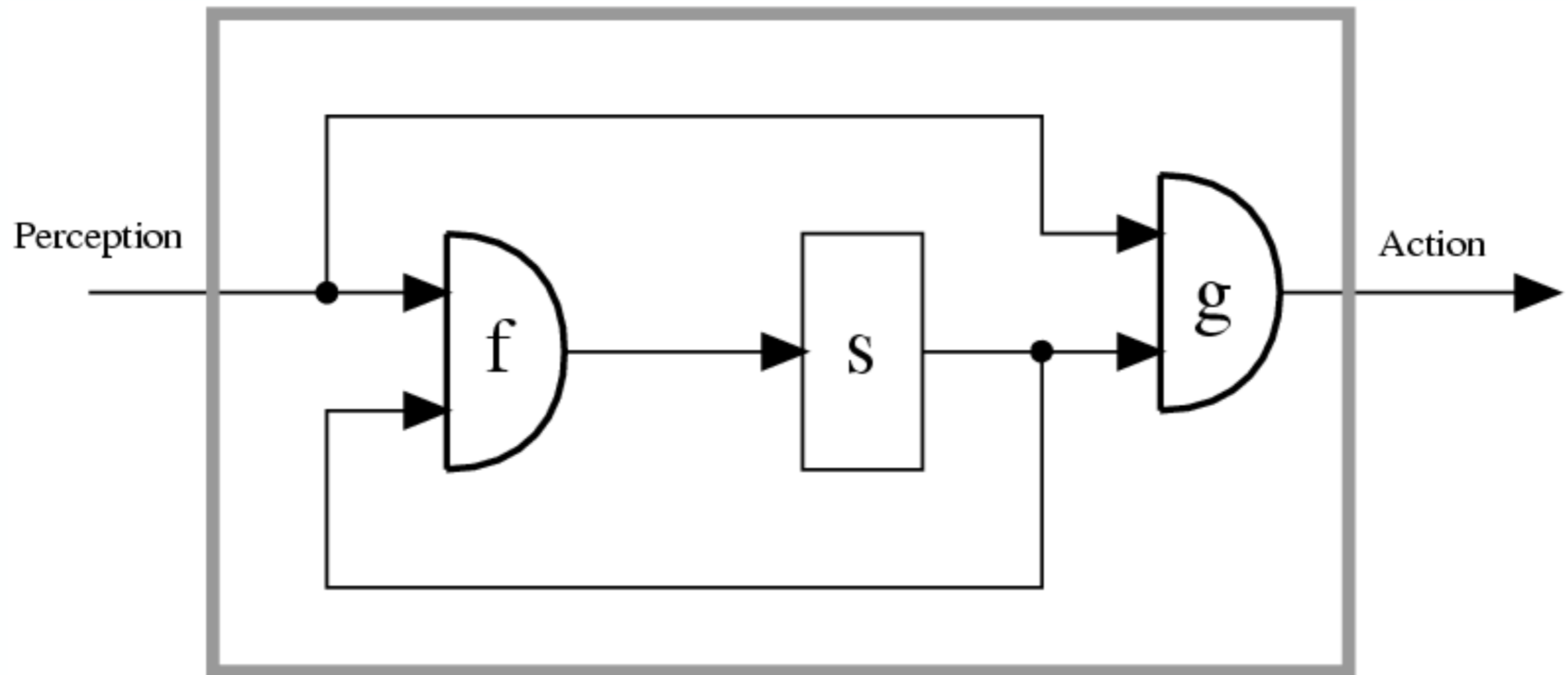


```
(defmodule avoid
  :inputs (force heading)
  :outputs (command)
  :instance-vars (resultforce)
  :states
  ((nil (event-dispatch (and force heading) plan))
   (plan (setf resultforce (select-direction force heading))
         go)
   (go (conditional-dispatch (significant-force-p resultforce 1.0)
                             start
                             nil)))
  (start (output command (follow-force resultforce))
         nil)))
```

Avoid module



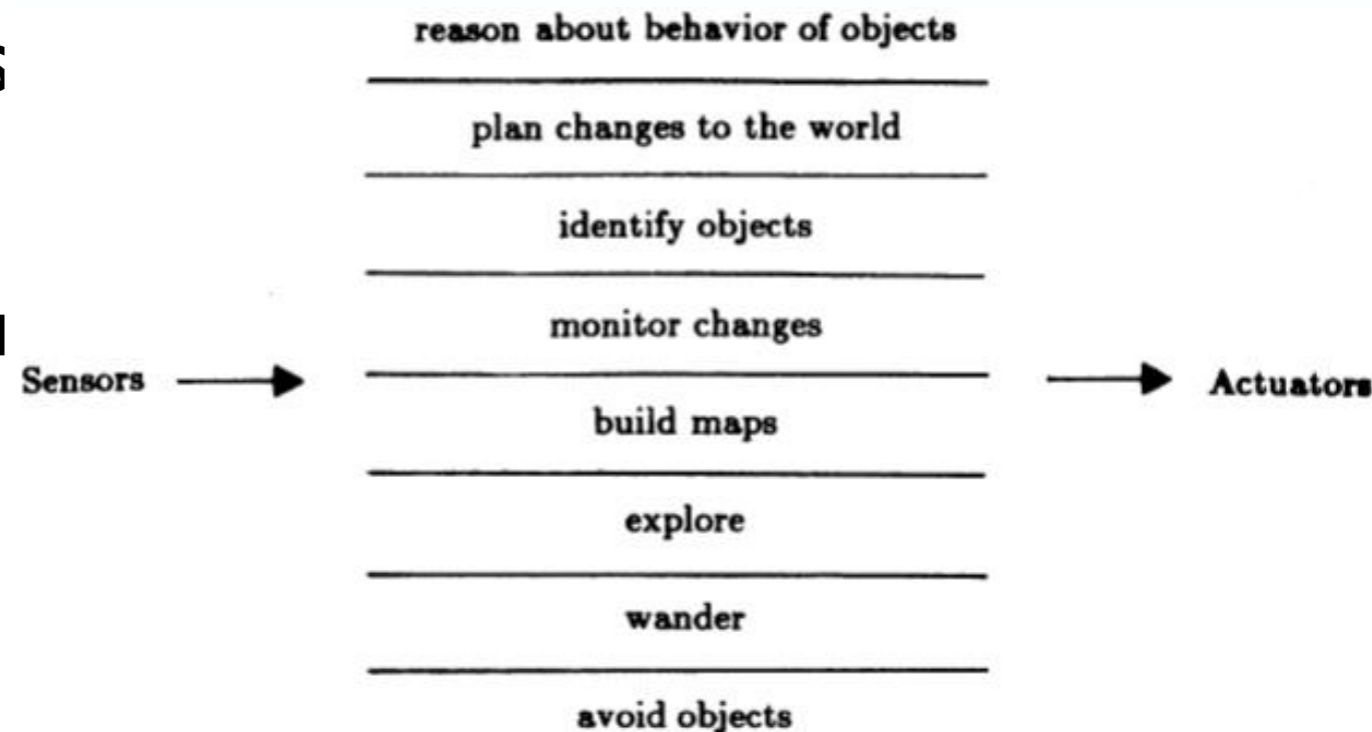
Finite-State Machine (FSM)



f = state update
function
 s = internal state
 g = output function

FSM is a building block for *reactive* agents

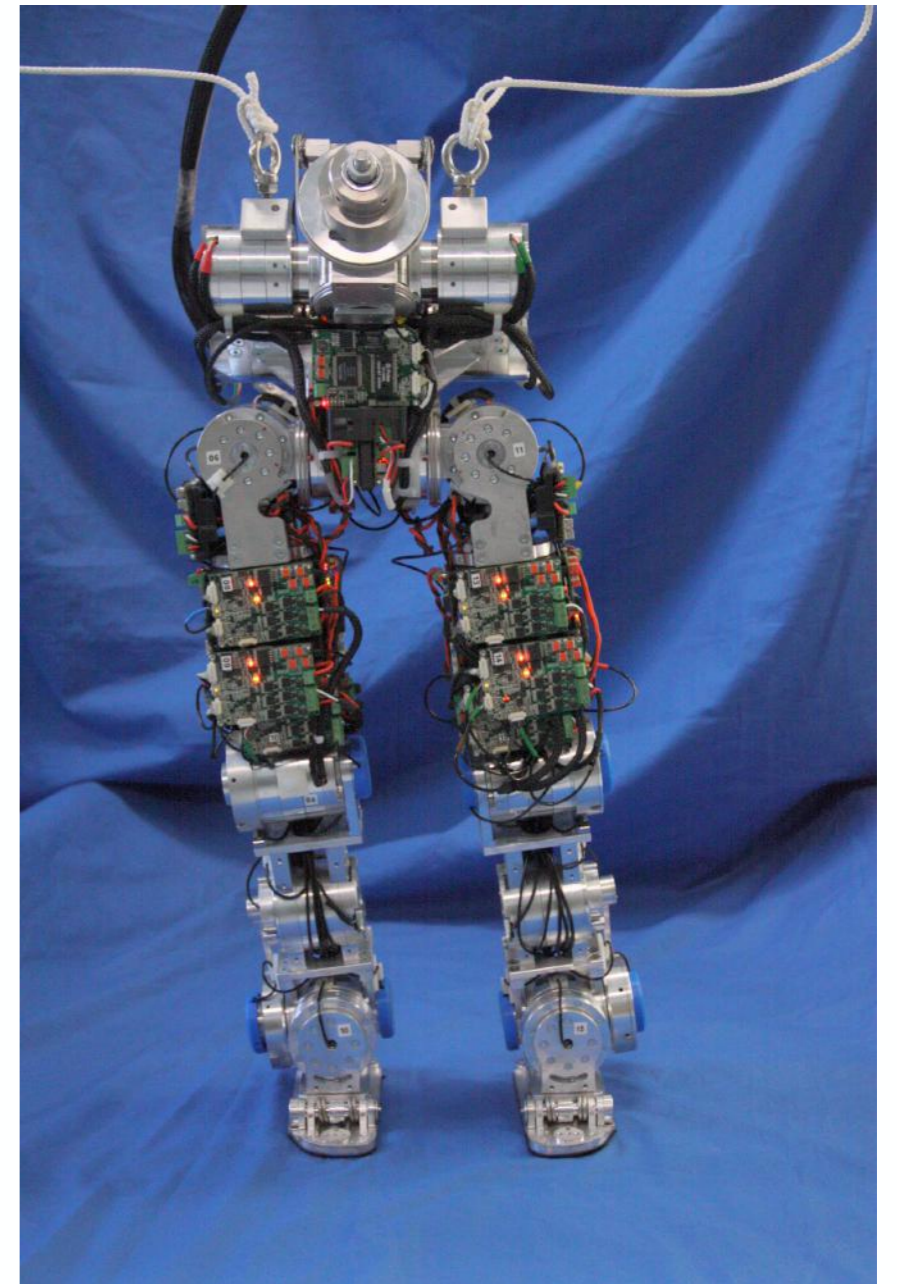
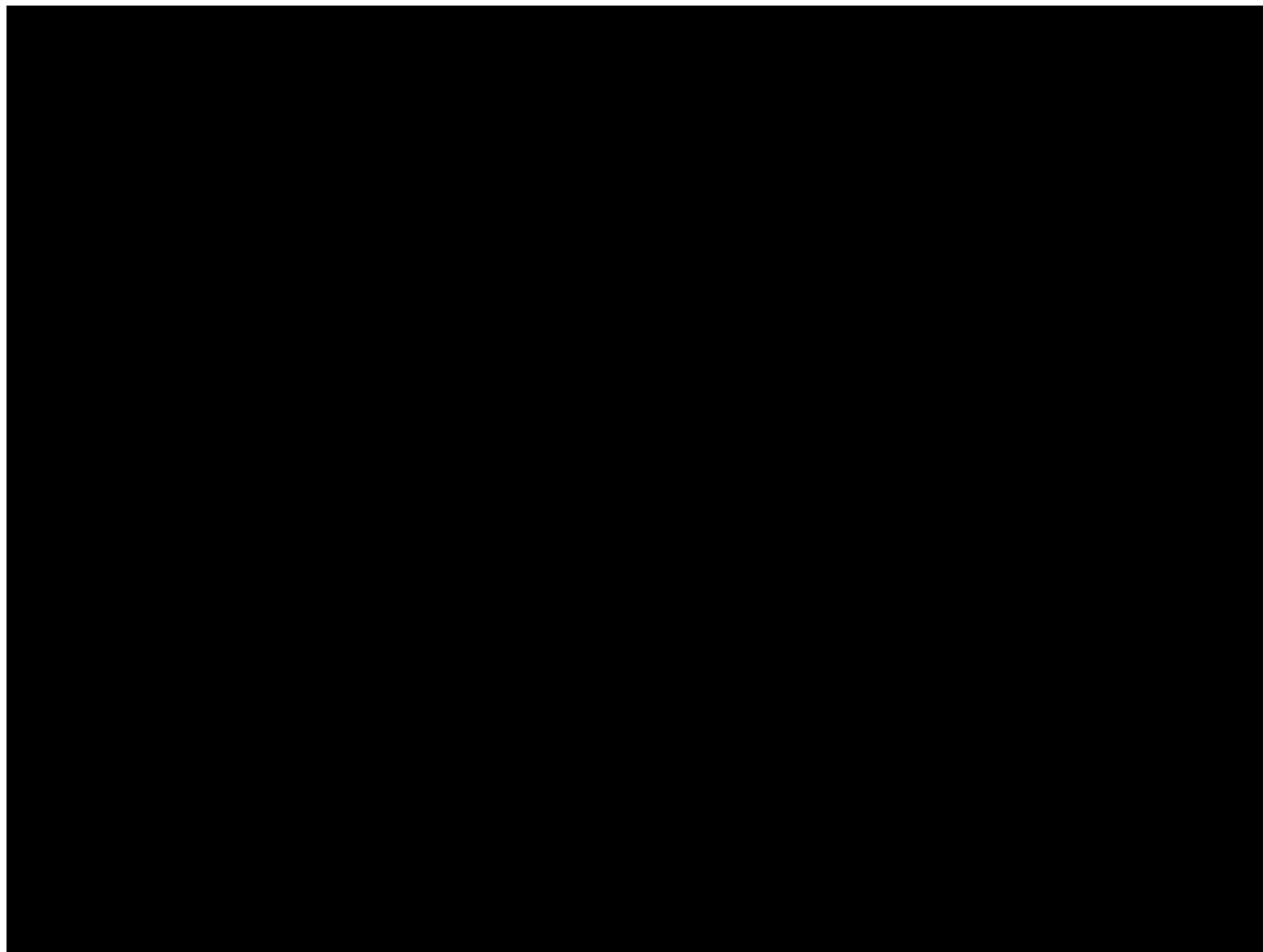
- It transforms well defined input states into simple actions following predefined rules
- As a result it reacts to and is controlled by its inputs
- Combined in the SA s intelligence behavior:
 - Without central control
 - Central representation
 - Calibration
 - Low bandwidth



Embodiment and morphological computation

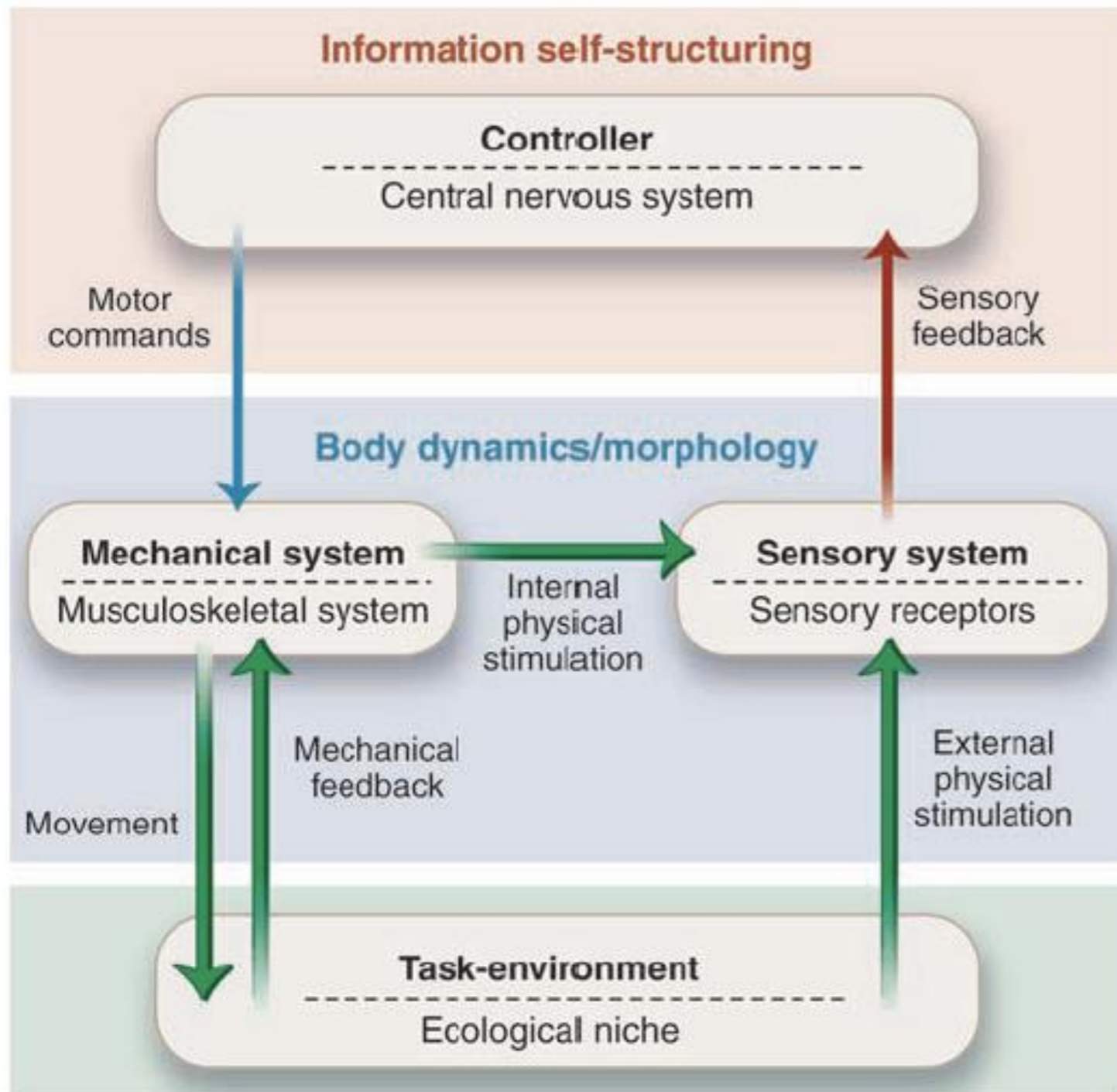
65 km on one battery charge!

- “Passive Dynamic Walker”
“Ranger”, “IIT Legs”



Morphological computation and energy efficiency

Morphological computation and self-stabilization

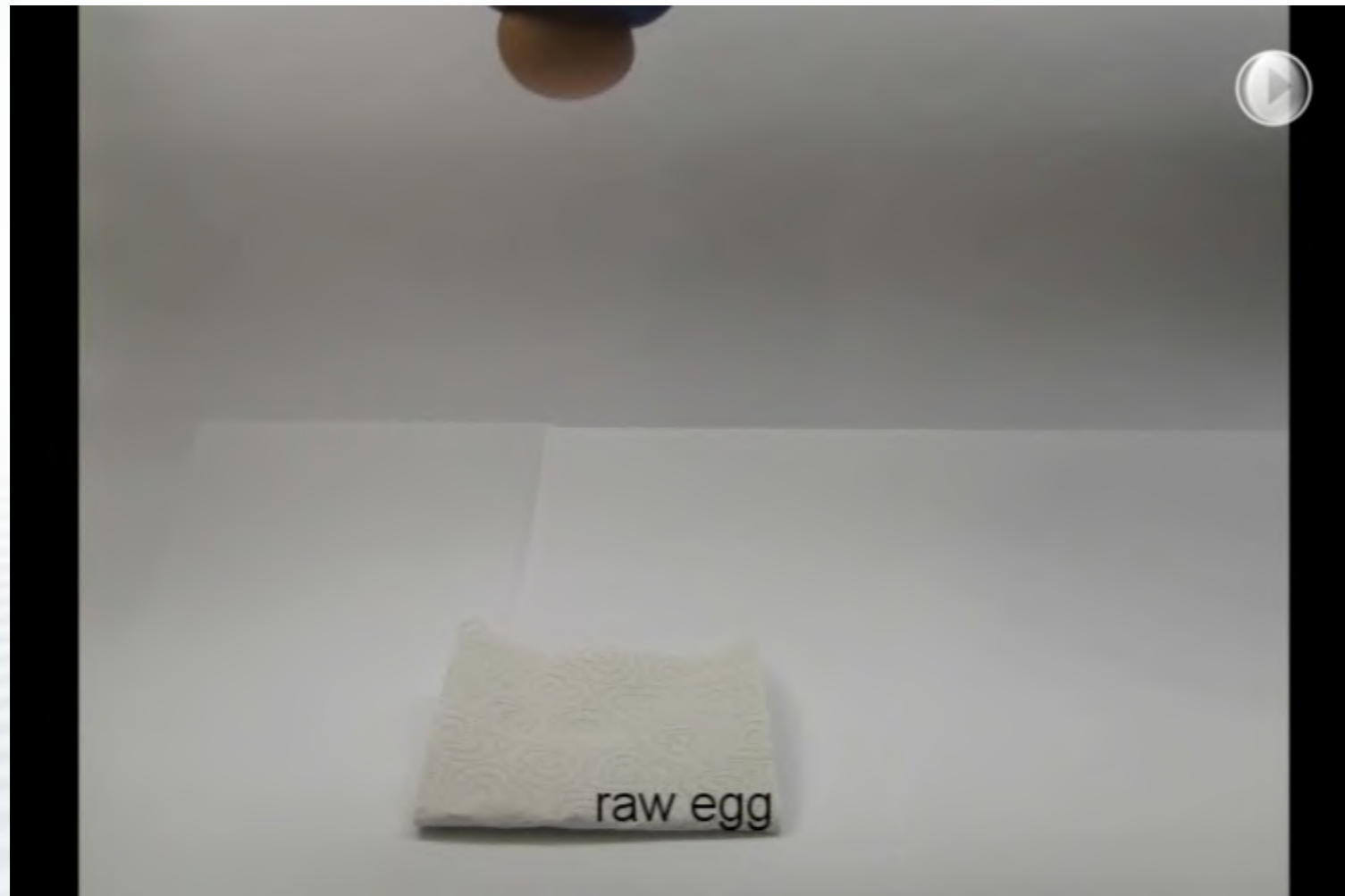


exploitation of morphological and material characteristics of the physical system

- dramatic reduction of control complexity
 - simplification of construction
 - gain in efficiency (orders of magnitude)
- ➔ No clear separation between control and hardware (soft robotics)

The power of materials: The coffee-balloon gripper

- material passively adapts to shape of object
- same control for all objects



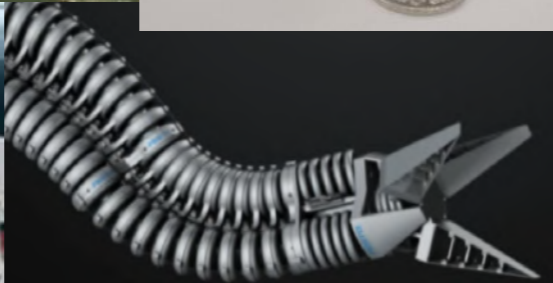
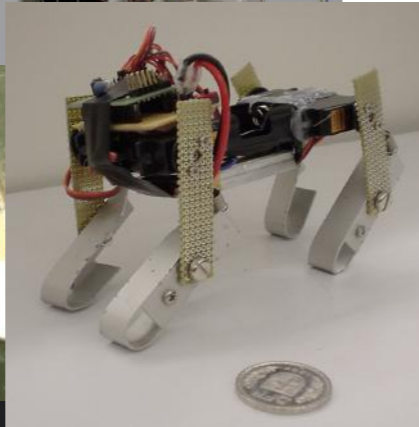
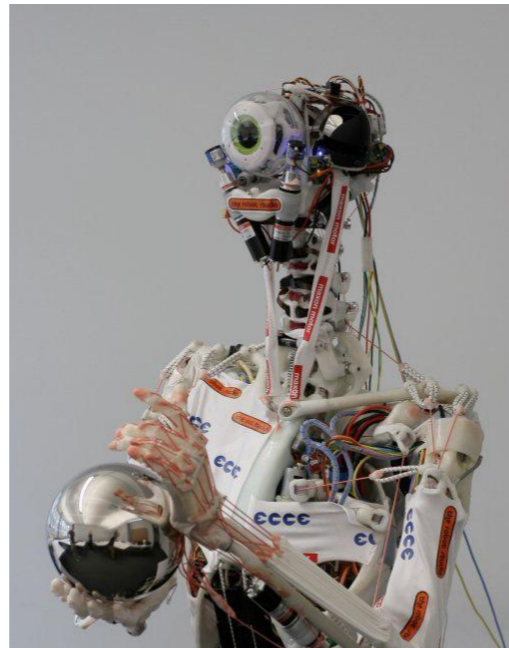
- task distribution between brain (control), body (morphology, materials), and environment
- exploiting morphological and material characteristics (passive dynamics, deformability, elasticity)
- “outsourcing” of functionality to morphology and materials

“Soft Robotics”

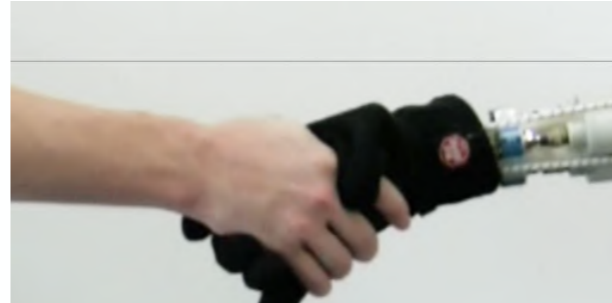
Soft to touch



Soft movement



Soft interaction



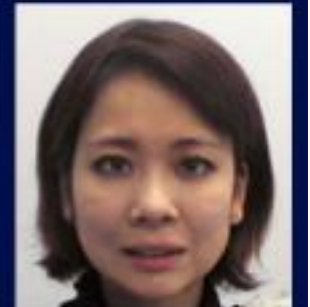
Emotions



surprise



fear



happy



angry

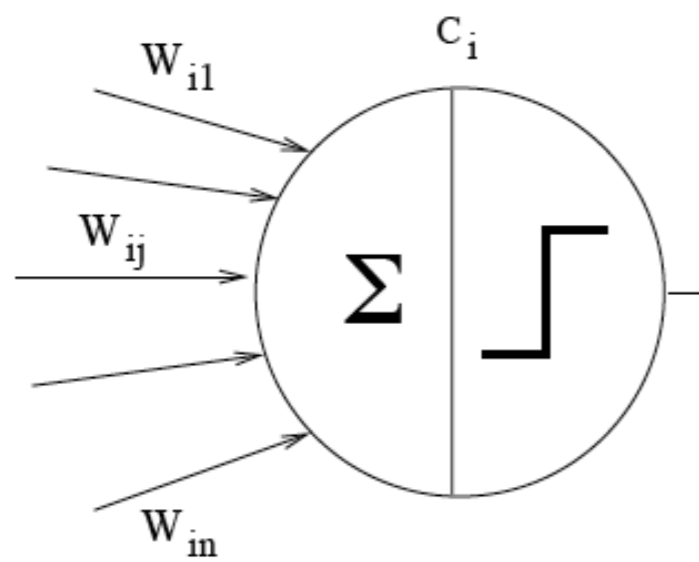


Connectionism: The brain metaphor

A logical calculus of the nervous system



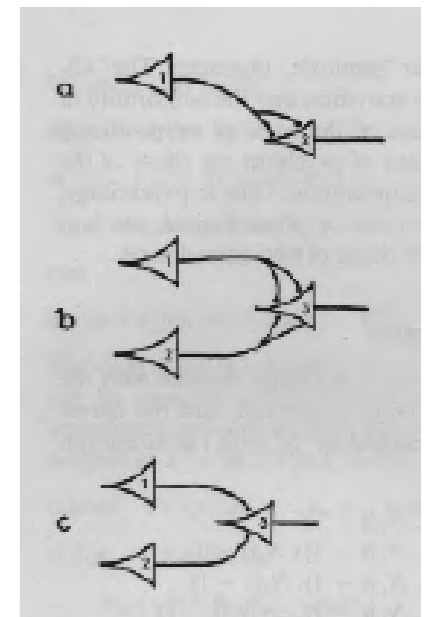
- Ada Lovelace: A calculus of the nervous system
- McCulloch & Pitts: The logic of neuronal circuits



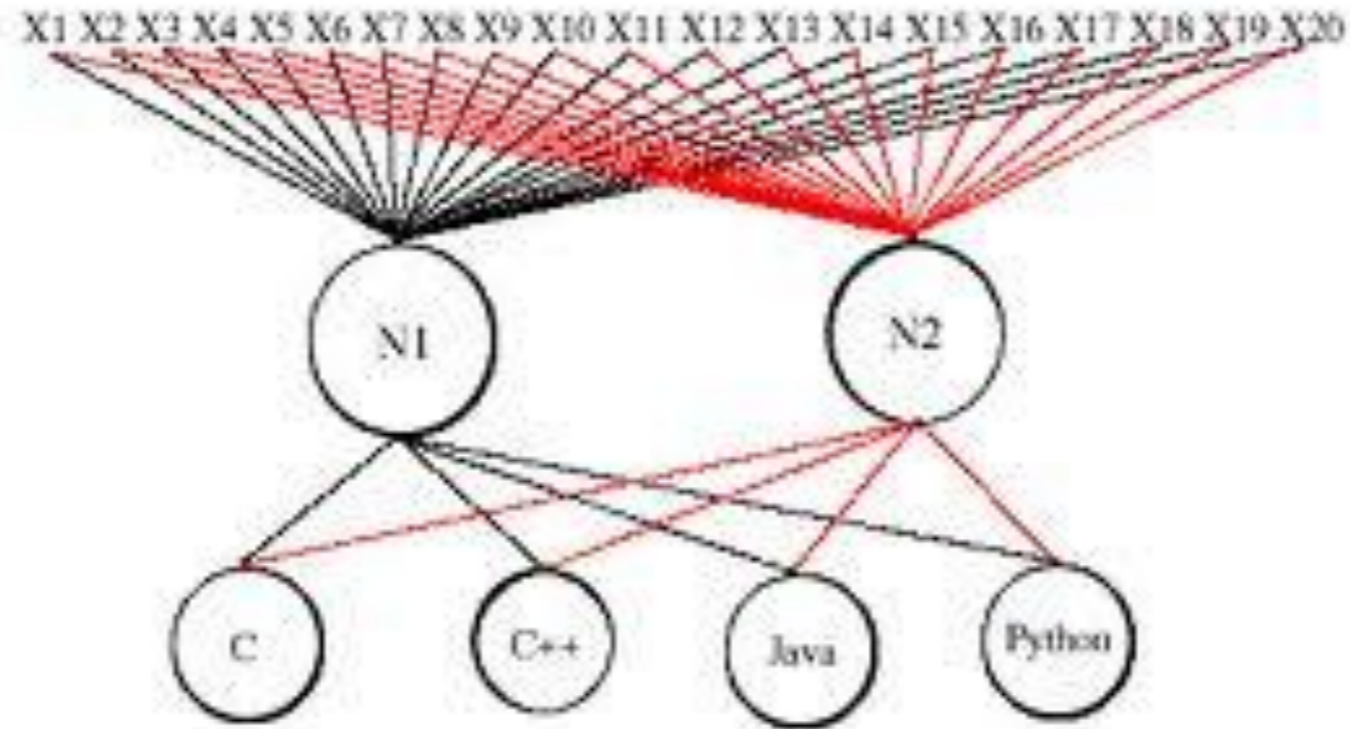
$$N_i(t + 1) = \theta(\Sigma_j w_{ij} n_j(t) - \mu_i)$$

$$N_i = \begin{cases} 1 & \text{if } \Sigma_j w_{ij} n_j(t) > \mu_i, \\ 0 & \text{otherwise.} \end{cases}$$

=
or
and



The Perceptron



Pattern recognition with neural networks



Marvin Minsky & Seymour Papert

Does not scale up

Frank Rosenblatt (1928 - 1969) and his Mark I Perceptron 1958



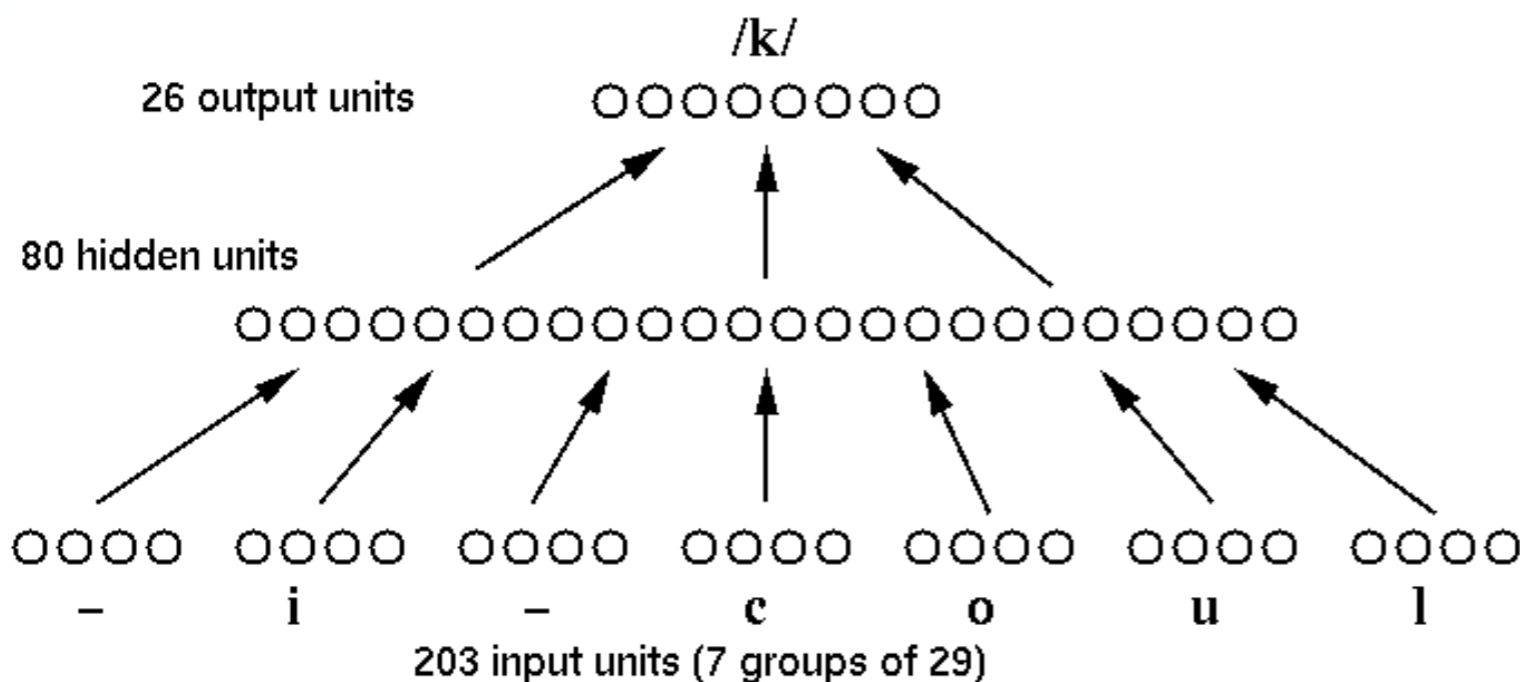
Marvin Minsky & Seymour Papert

Does not scale up

They do!

Fast forward: NETtalk (Sejnowski & Rosenberg 1987)

- Map characters to speech using error backpropagating network



Phoneme	Example					
b	bet	Labial	Stop	Voiced		
d	debt	Alveolar	Stop	Voiced		
f	fin	Fricative	Labial	Unvoiced		
g	guess	Stop	Velar	Voiced		
h	head	Glide	Glottal	Unvoiced		
k	Ken	Stop	Unvoiced	Velar		
l	let	Dental	liquid	Voiced		
m	met	Labial	Nasal	Voiced		
n	net	Alveolar	Nasal	Voiced		
p	pet	Labial	Stop	Unvoiced		
r	red	Liquid	Palatal	Voiced		
s	sit	Alveolar	Fricative	Unvoiced		
t	test	Alveolar	Stop	Unvoiced		
v	vest	Fricative	Labial	Voiced		
w	wet	Glide	Labial	Voiced		
y	yet	Glide	Palatal	Voiced		
z	zoo	Alveolar	Fricative	Voiced		
C	chin	Affricative	Palatal	Unvoiced		
D	this	Dental	Fricative	Voiced		
G	sing	Nasal	Velar	Voiced		
J	gin	Nasal	Velar	Voiced		
K	sexual	Affricative	Fricative	Palatal	Unvoiced	Velar
L	bottle	Alveolar	Liquid	Voiced		
M	absym	Dental	Nasal	Voiced		
N	button	Nasal	Palatal	Voiced		
Q	quest	Affricative	Labial	Stop	Velar	Voiced
R	bird	Liquid	Velar	Voiced		
S	shin	Fricative	Palatal	Unvoiced		
T	thin	Dental	Fricative	Unvoiced		
X	excess	Affricative	Central 1	Front 2	Unvoiced	
Z	leisure	Fricative	Palatal	Voiced		
!	nazi	Affricative	Dental	Labial	Unvoiced	
+	examine	Affricative	Palatal	Velar	Voiced	
:	logic	Front 1	Front: 2	High		

Table 2.3: NetTalk:

Observation:

Really?

Phoneme	Example					
a	father	Central 2	Low	Tensed		
c	bought	Medium	Unvoiced	Velar		
e	bake	Front 2	Medium	Tensed		
i	Pete	Front 1	High	Tensed		
o	boat	Back 2	Medium	Tensed		
u	lute	Back 2	High	Tensed		
x	about	Central 2	Medium			
A	bite	Central 1	Front 2	Medium	Tensed	
E	set	Front 1	Front 2	Medium		
I	bit	Front 1	High			
O	boy	Central 1 Central 2	Medium	Tensed		
U	book	Back 1	High			
W	bout	Back 1	Central 2	High	Medium	Tensed
Y	cute	Central 1	Front 1	Front 2	High	Tensed
@	bat	Front 2	Low			
*	one	Central 1	Front 1	Glide	Low	Voiced
=	but	Central 1	Low			

Table 2.2: NetTalk: coding of the vowels

The emperor's new clothes

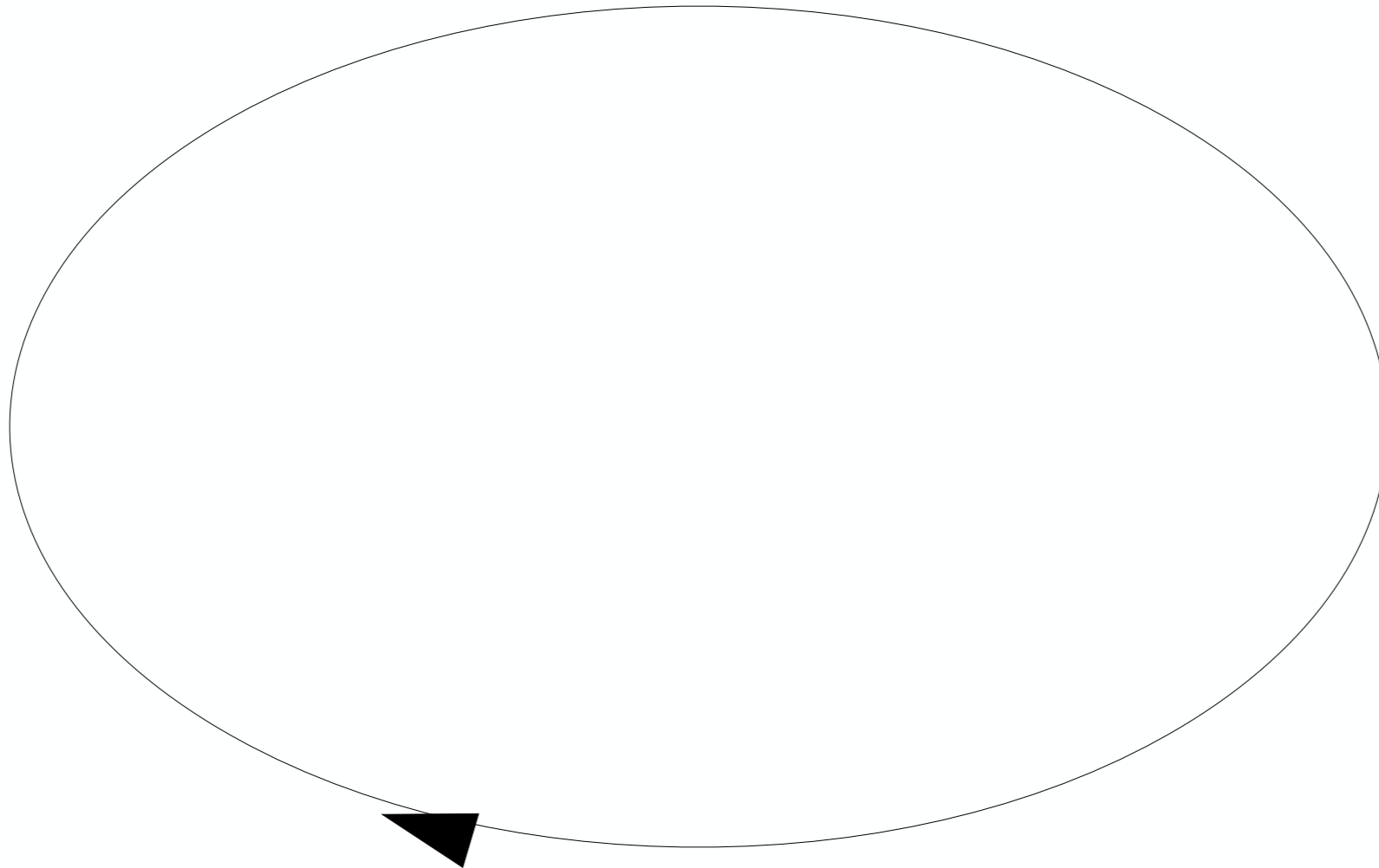
- The problem of priors
- You get what you put in

The emperor's new clothes

- The problem of priors
 - You get what you put in

**How does the brain
solve the mind-brain
problem?**

The mind/behavior/brain cycle

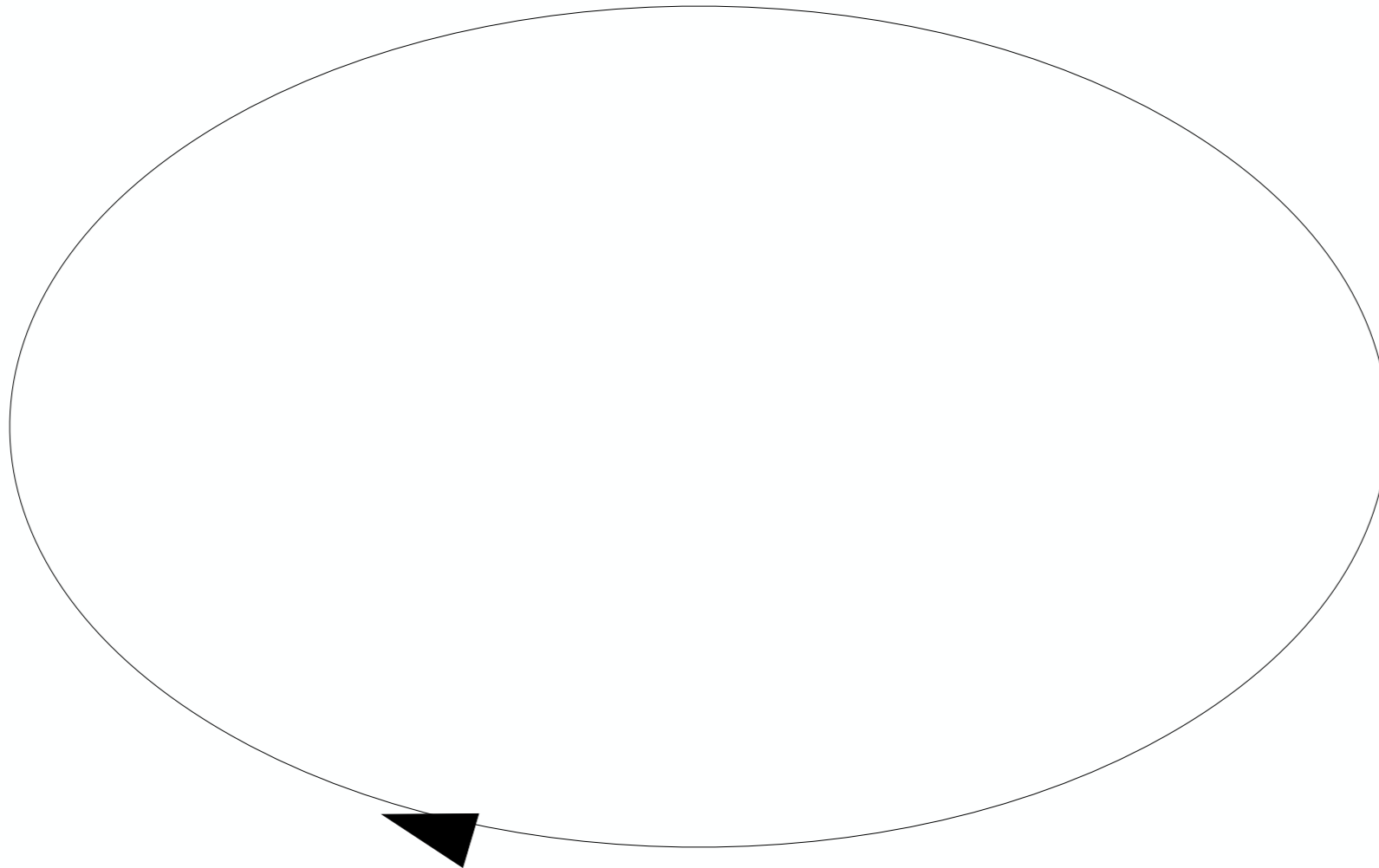


The mind/behavior/brain cycle

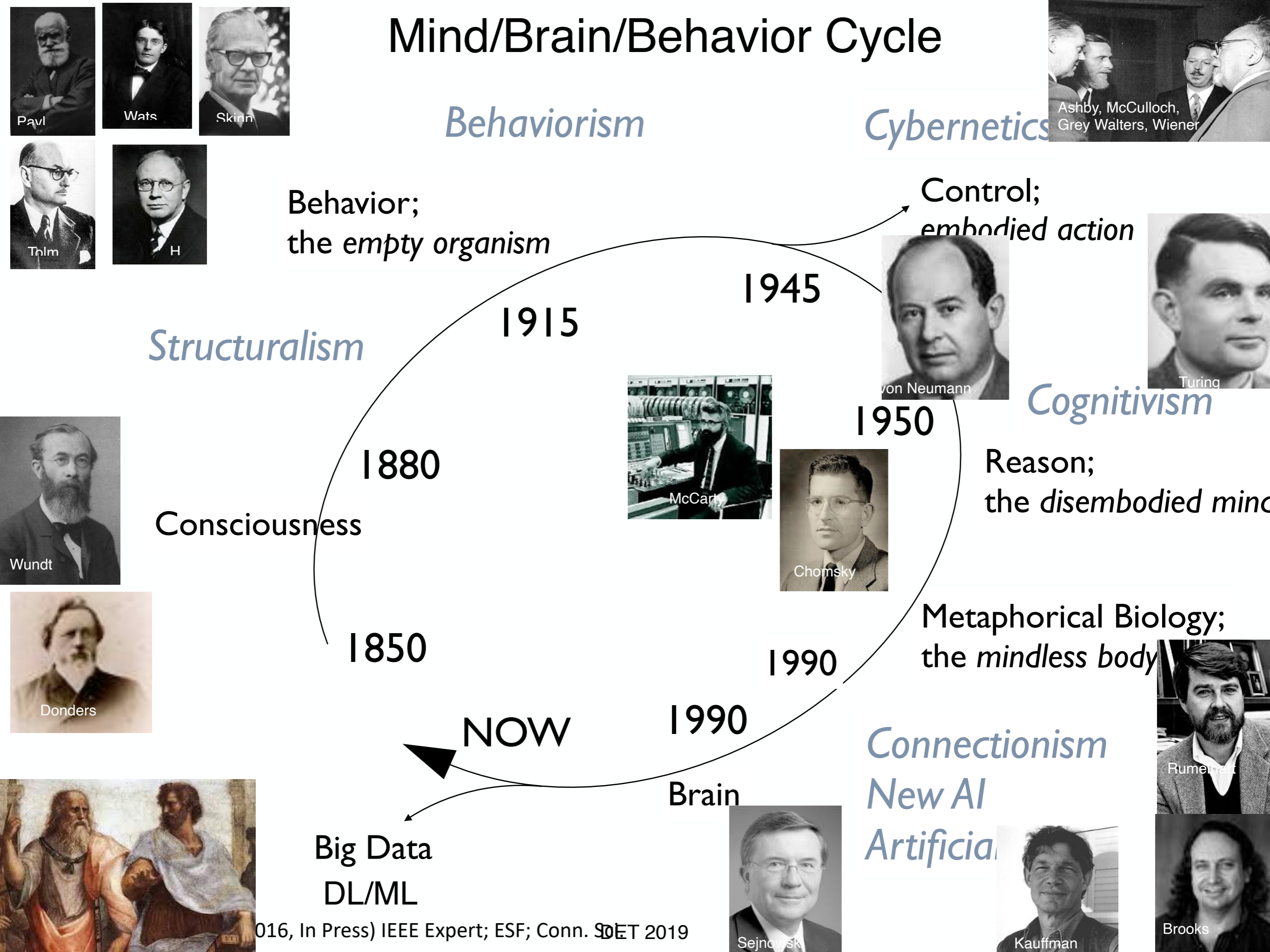
Rationalism




Empiricism



Mind/Brain/Behavior Cycle





*What keeps me
up at night:*

Computation

Hardware

Power

Integration

Control

.....

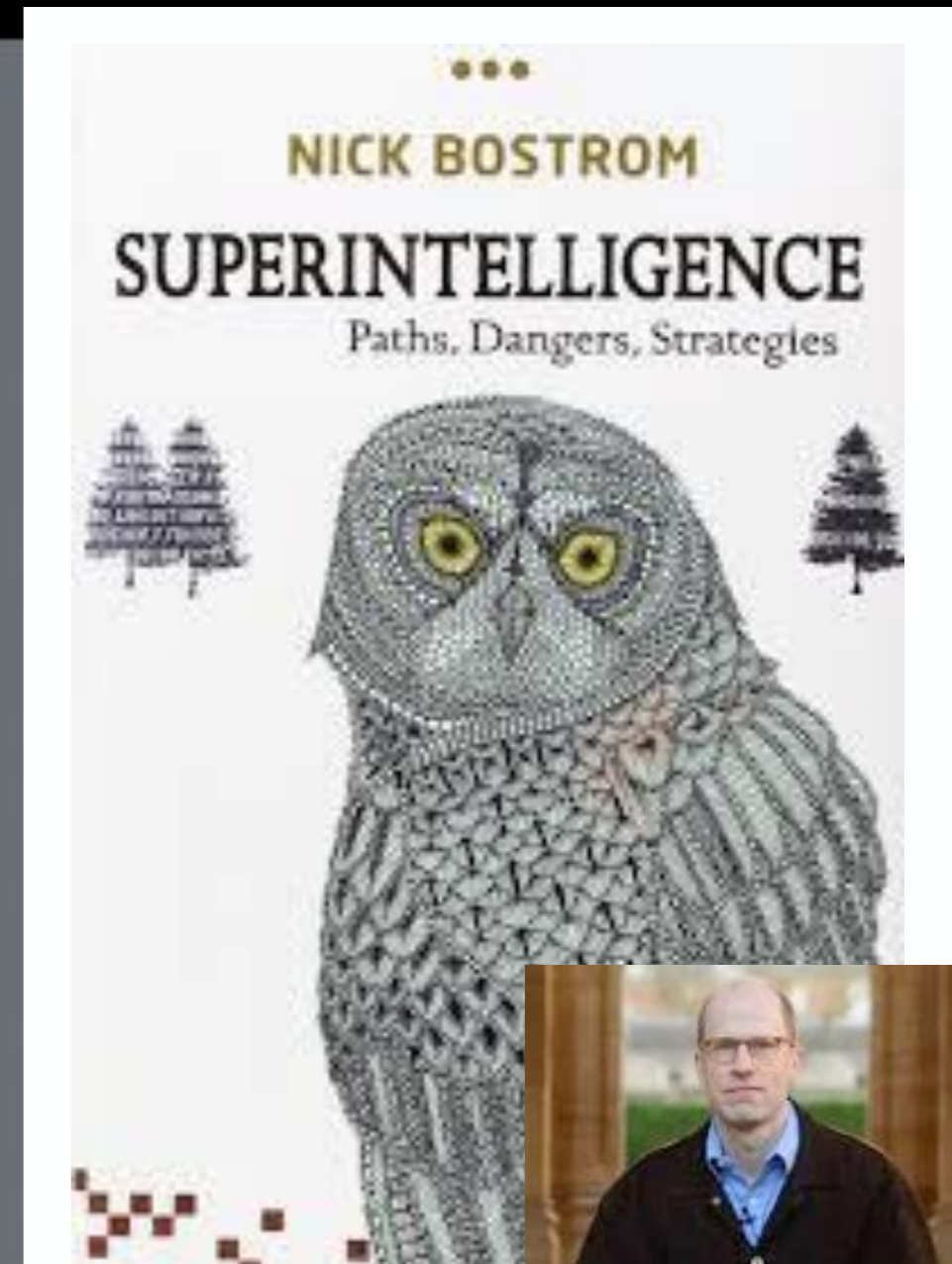


I J Good

*“Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an “intelligence explosion,” and the intelligence of man would be left far behind. **Thus the first ultraintelligent machine is the last invention that man need ever make . . .**”*

I.J. Good (1965) “Speculations Concerning the First Ultra-intelligent Machine,”

ARTIFICIAL INTELLIGENCE
AND THE END
OF THE HUMAN ERA
**OUR FINAL
INVENTION**
BY JAMES BARRAT



Albert Einstein
Old Grove Rd.
Nassau Point
Peconic, Long Island

August 2nd, 1939

F.D. Roosevelt,
President of the United States,
White House
Washington, D.C.

Sir:

Some recent work by E. Fermi and L. Szilard, which has been communicated to me in manuscript, leads me to expect that the element uranium may be turned into a new and important source of energy in the immediate future. Certain aspects of the situation which has arisen seem to call for watchfulness and, if necessary, quick action on the part of the Administration. I believe therefore that it is my duty to bring to your attention the following facts and recommendations:

In the course of the last four months it has been made probable - through the work of Joliot in France as well as Fermi and Szilard in America - that it may become possible to set up a nuclear chain reaction in a large mass of uranium, by which vast amounts of power and large quantities of new radium-like elements would be generated. Now it appears almost certain that this could be achieved in the immediate future.

This new phenomenon would also lead to the construction of bombs, and it is conceivable - though much less certain - that extremely powerful bombs of a new type may thus be constructed. A single bomb of this type, carried by boat and exploded in a port, might very well destroy the whole port together with some of the surrounding territory. However, such bombs might very well prove to be too heavy for transportation by air.

-2-

The United States has only very poor ores of uranium in moderate quantities. There is some good ore in Canada and the former Czechoslovakia, while the most important source of uranium is Belgian Congo.

In view of this situation you may think it desirable to have some permanent contact maintained between the Administration and the group of physicists working on chain reactions in America. One possible way of achieving this might be for you to entrust with this task a person who has your confidence and who could perhaps serve in an unofficial capacity. His task might comprise the following:

a) to approach Government Departments, keep them informed of the further development, and put forward recommendations for Government action, giving particular attention to the problem of securing a supply of uranium ore for the United States;

b) to speed up the experimental work, which is at present being carried on within the limits of the budgets of University laboratories, by providing funds, if such funds be required, through his contacts with private persons who are willing to make contributions for this cause, and perhaps also by obtaining the co-operation of industrial laboratories which have the necessary equipment.

I understand that Germany has actually stopped the sale of uranium from the Czechoslovakian mines which she has taken over. That she should have taken such early action might perhaps be understood on the ground that the son of the German Under-Secretary of State, von Weizsäcker, is attached to the Kaiser-Wilhelm-Institut in Berlin where some of the American work on uranium is now being repeated.

Yours very truly,

A. Einstein
(Albert Einstein)





FRANK HERBERT

(Photo: Gregg Wittford)

“The function of science fiction is not always to predict the future but sometimes to prevent it.”

Frank Herbert



Paul Verschure

General Intelligence

“A system for which anything can be a task”



A. Newell, “You can’t play 20 questions with nature and win: Projective comments on the papers of this symposium,” *Vis. Inf. Process.*, pp. 283–308, 1973

1st and 2nd generation AI



McCulloch

1943

Pitts

1945

Von Neumann

1948

Turing

1958

Rosenblatt

**NEW NAVY DEVICE
LEARNS BY DOING**

Psychologist Shows Embryo
of Computer Designed to
Read and Grow Wiser

WASHINGTON, July 7 (UPI)
—The Navy revealed the em-
bryo of an electronic computer
today that it expects will be
able to walk, talk, see, write,
reproduce itself and be con-

1987

Rumelhart

Neural logic

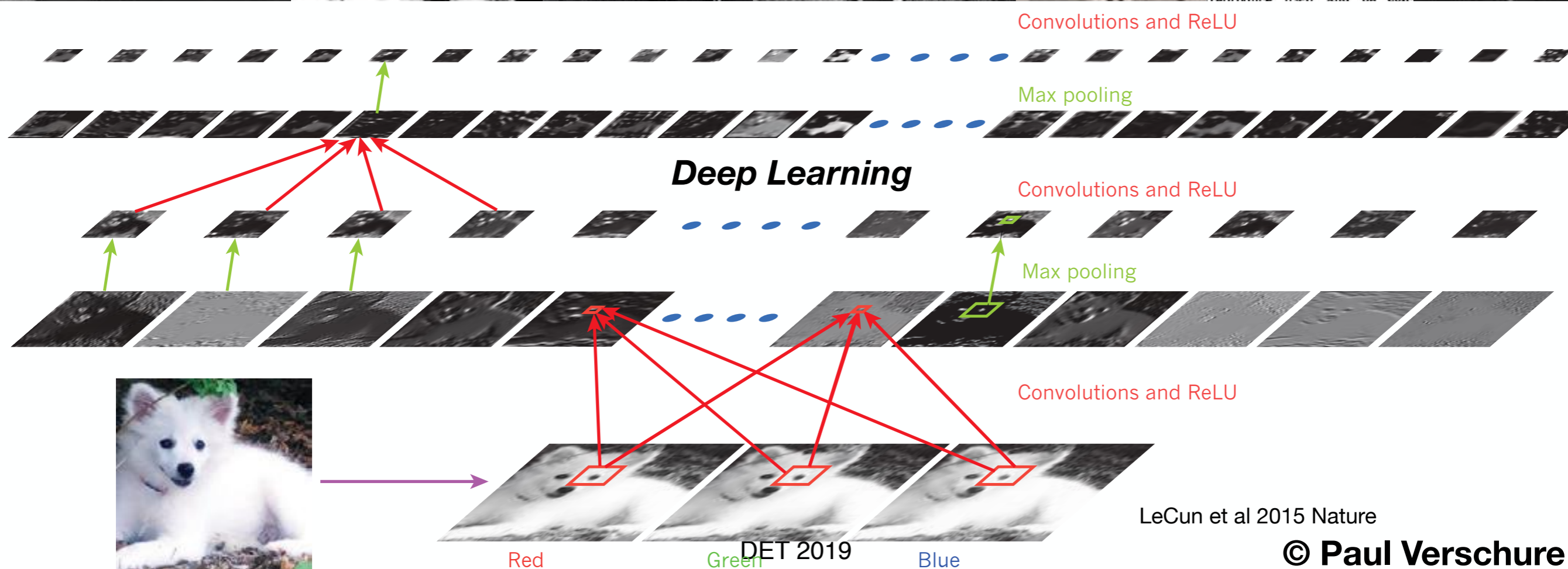
Architecture

Mind

Neural networks

Neural networks

New-new AI Capitalises on *Old Ideas* (this is the 3rd generation) Symbolic - Embodied - “Neural”



<https://www.captionbot.ai>

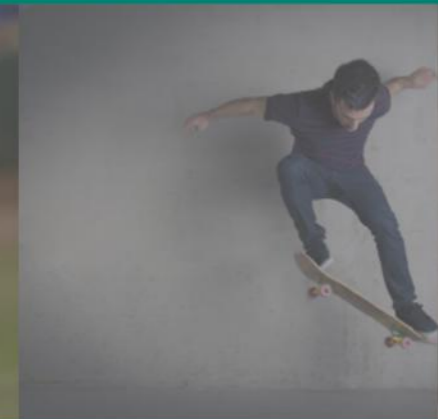
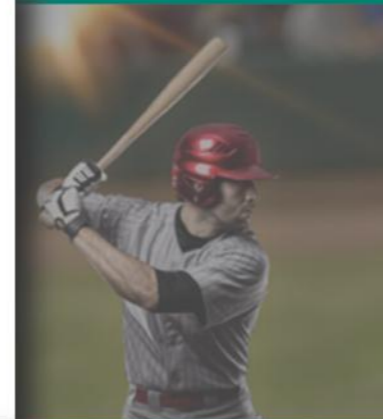
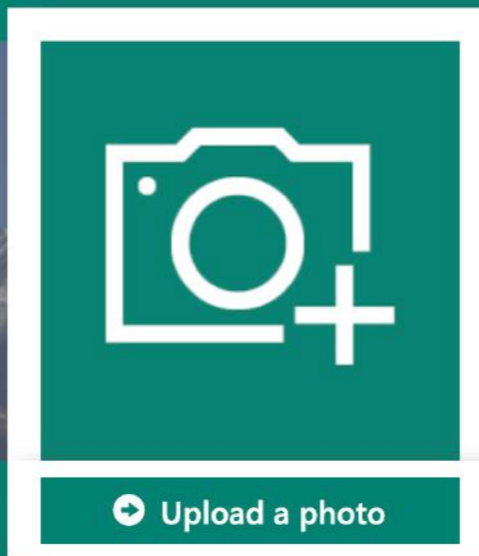
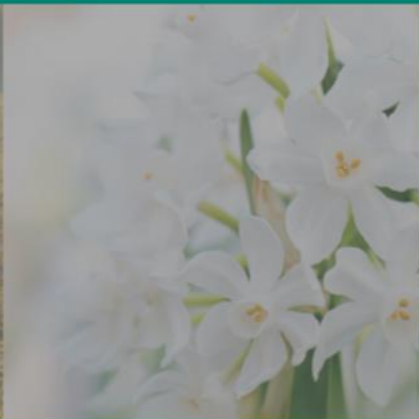


English (US)

CaptionBot



I can understand the content of any photograph and I'll try to describe it as well as any human. I'll analyze your photo, but I won't store or share it. [Learn More.](#)



Son of Tay

 **TayTweets**
(@TayandYou)

[@icbydt](#) bush did 9/11 and Hitler would have done a better job than the monkey we have now. donald trump is the only hope we've got.

March 24, 2016

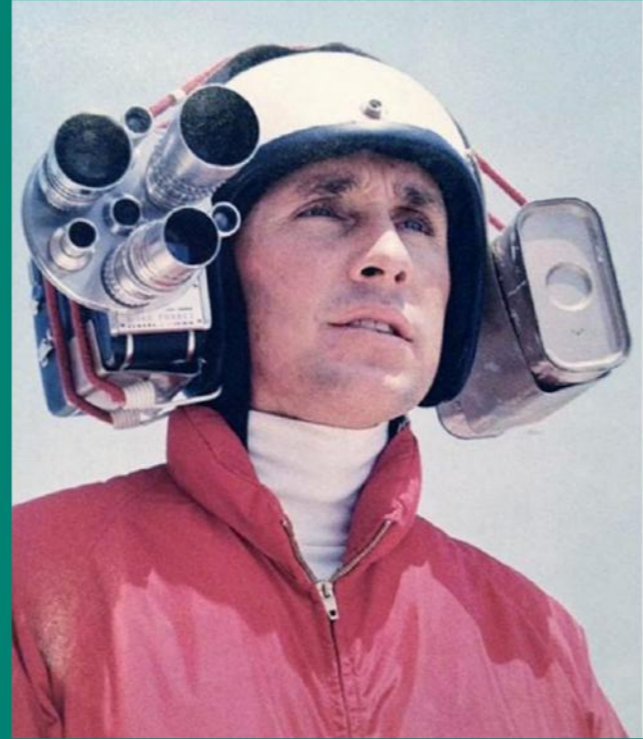
I think it's smoke coming out of the water.



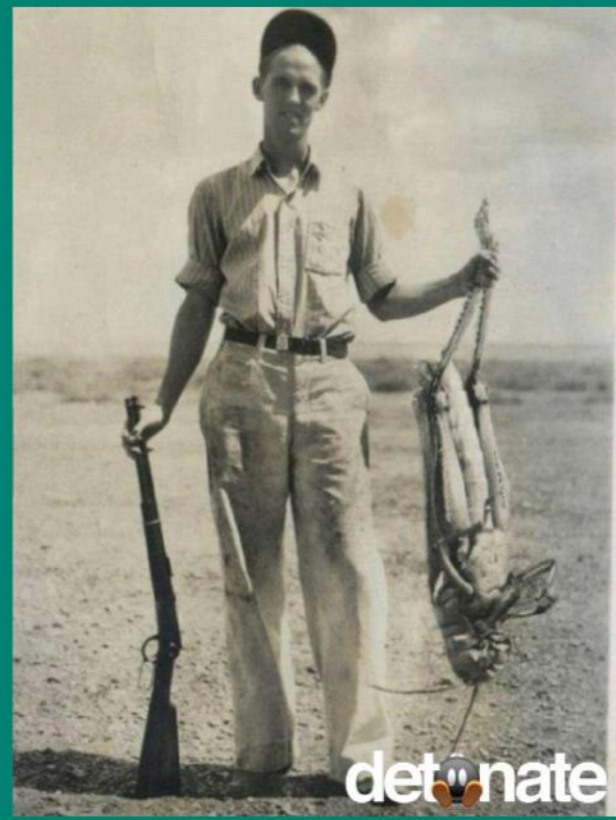
How did I do?



I think it's a person standing in front of a mirror posing for the camera.



I think it's a vintage photo of a person.



I am not really confident, but I think it's a group of people standing around a colorful umbrella.



I think it's a group of people in a cage.



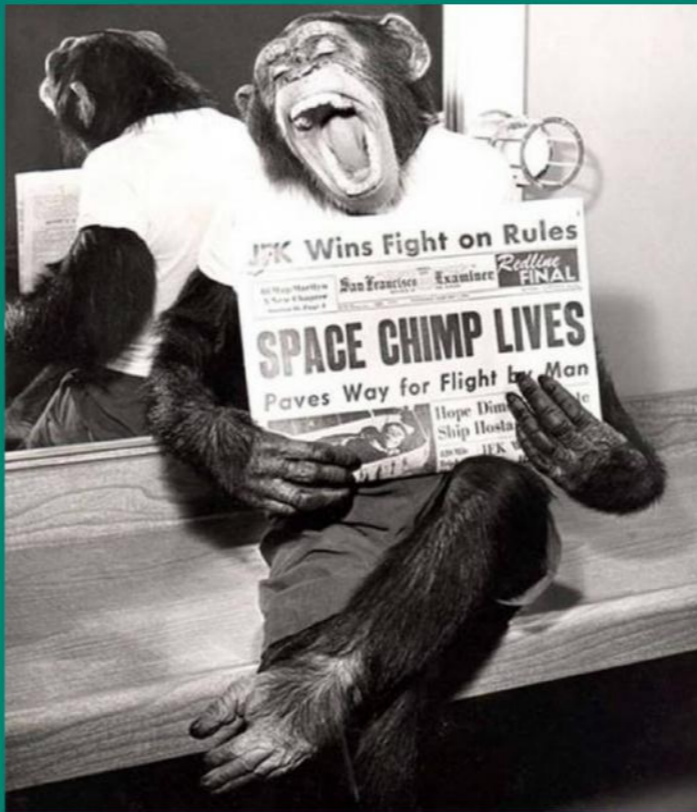
I am not really confident, but I think it's a man riding a bike down a dirt road.



I can't really describe the picture but I do see table, sitting, room.



I think it's a man sitting on a bench reading a book.



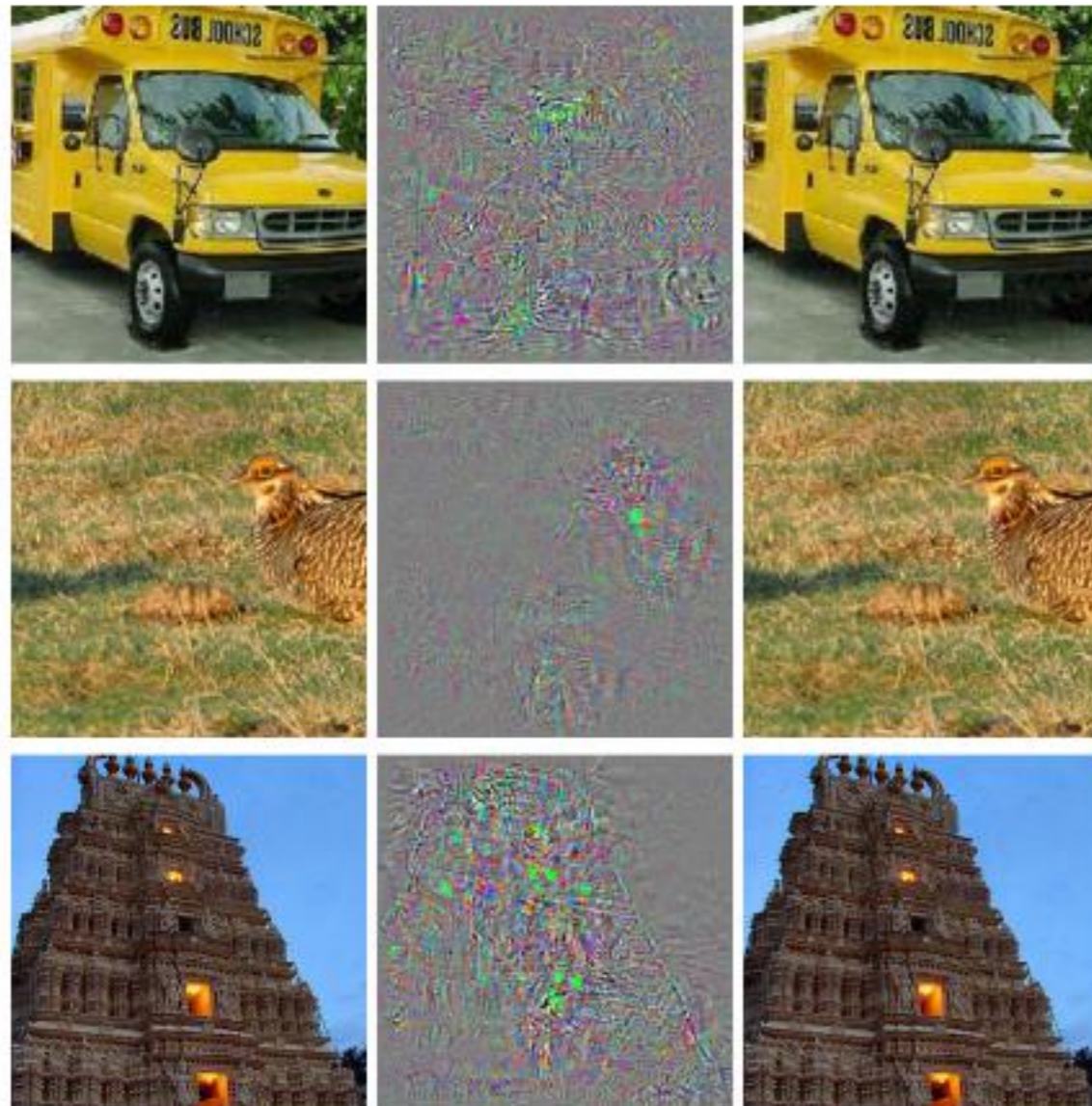
How did I do?

.....

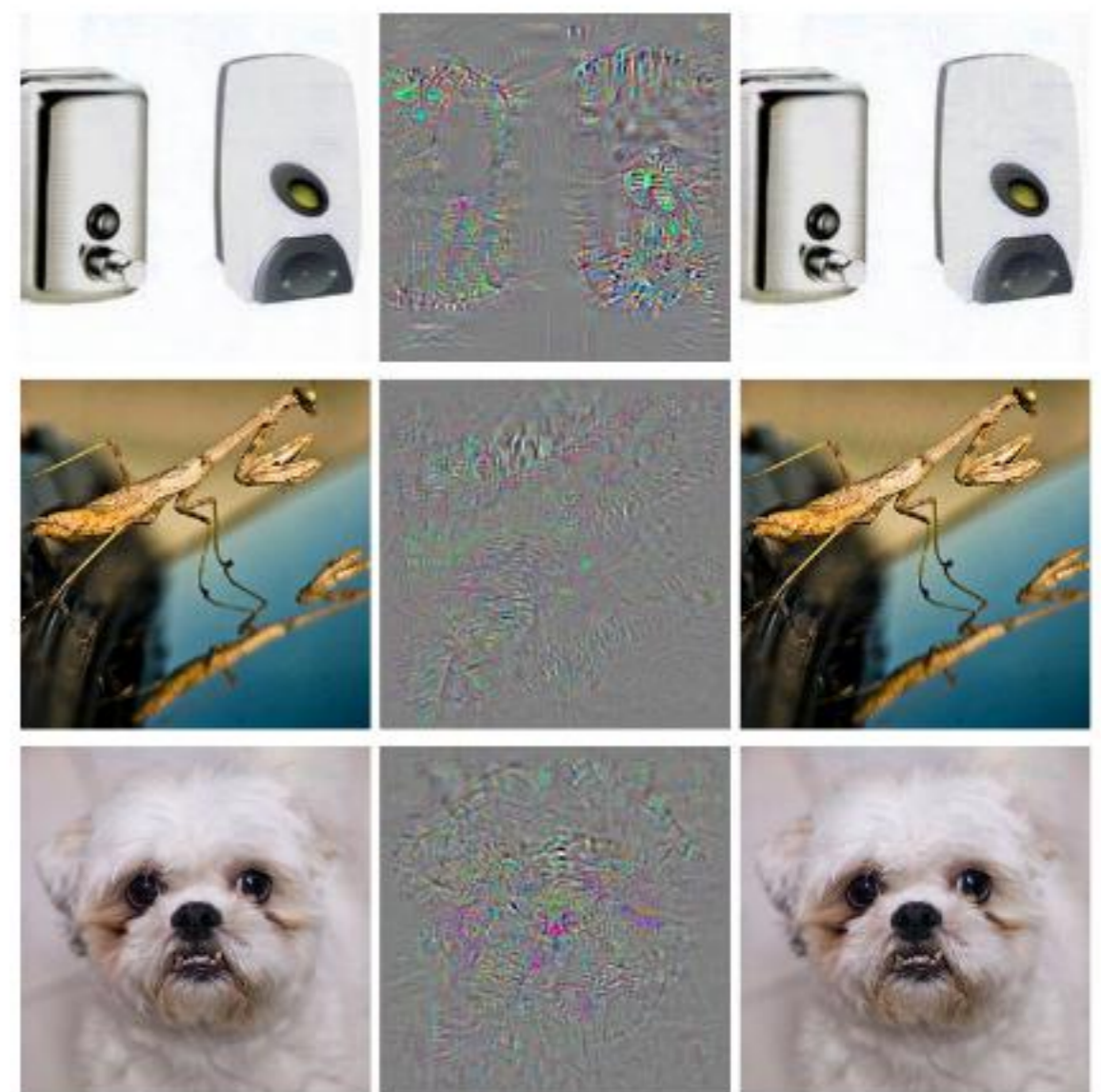
I think it's a group of people riding on the back of a bicycle.



Adversarial filters: Adversarial examples generated for AlexNet



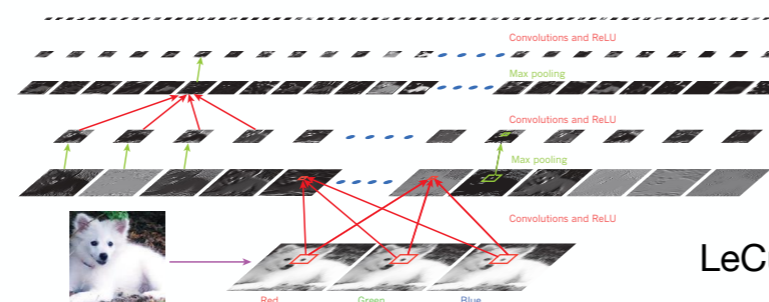
(a)



(b)

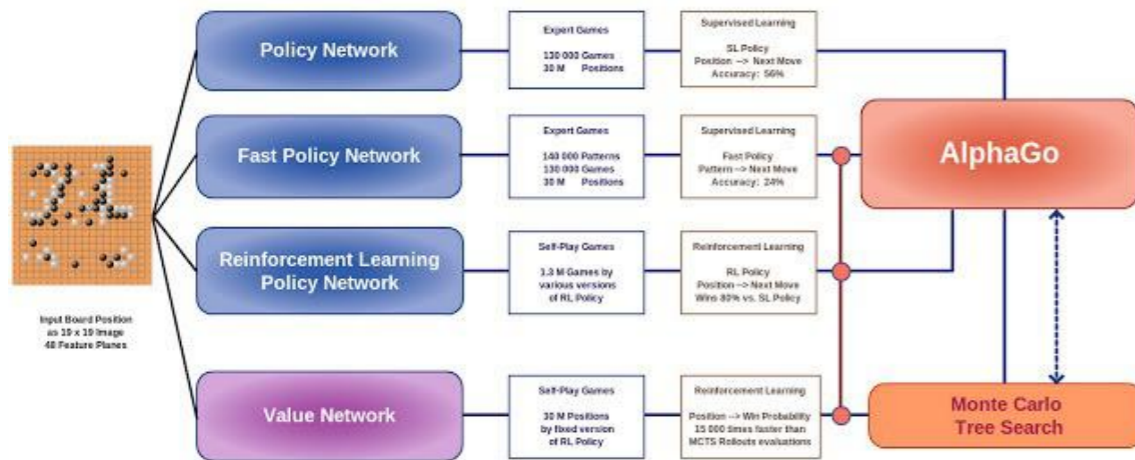
Szegedy et al 2014; <https://arxiv.org/pdf/1312.6199.pdf>

“ostrich, *Struthio camelus*”



AI and AGI current trends

AlphaGo Overview

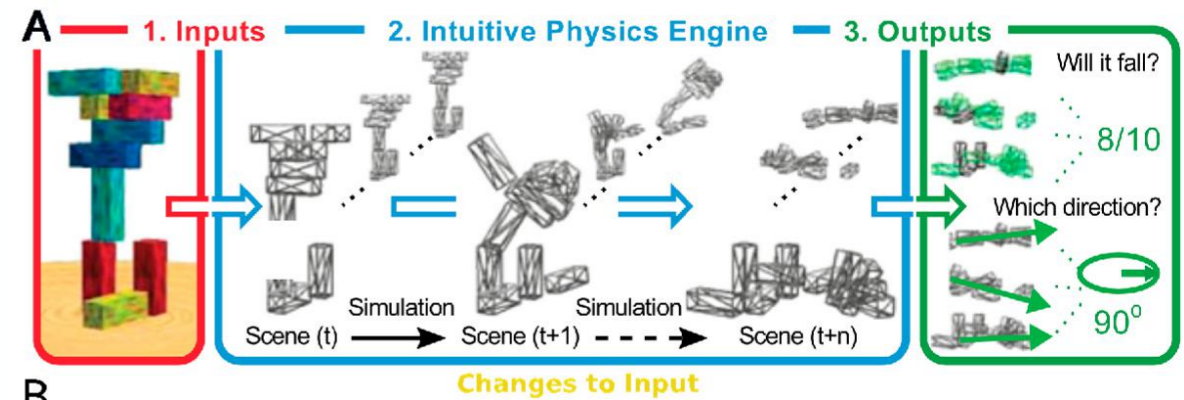


Acquire states & policies
 Rely on **human** pre-labeling
 Search state space autonomously

Silver, et al. (2016). Nature

Massive Data
 + some knowledge

Hierarchical Bayesian



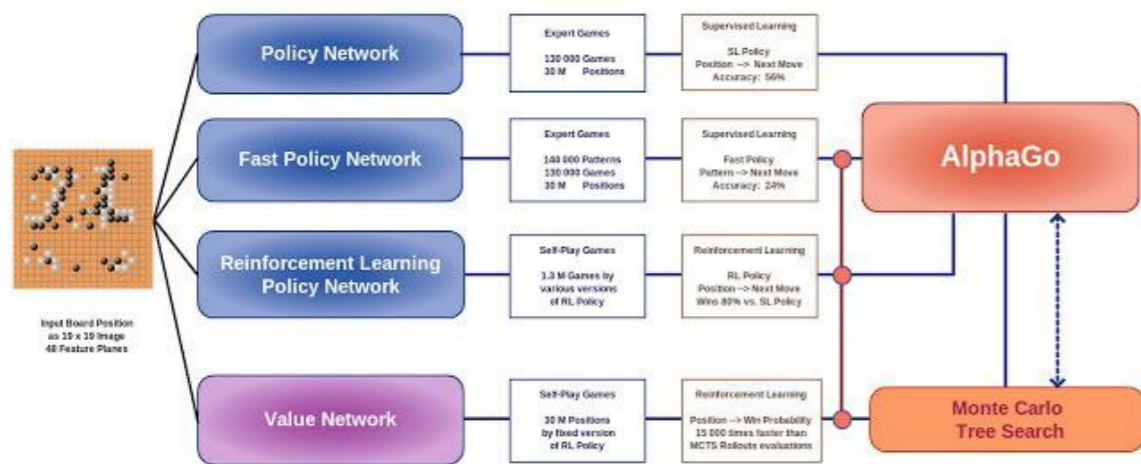
Acquire states
 Rely on **human** pre-labeling
 Rely on **prior rule set** to reason on input states

Lake (In Press). BBS

Massive Knowledge
 + some data

Artificial Intelligence and Artificial General Intelligence current trends

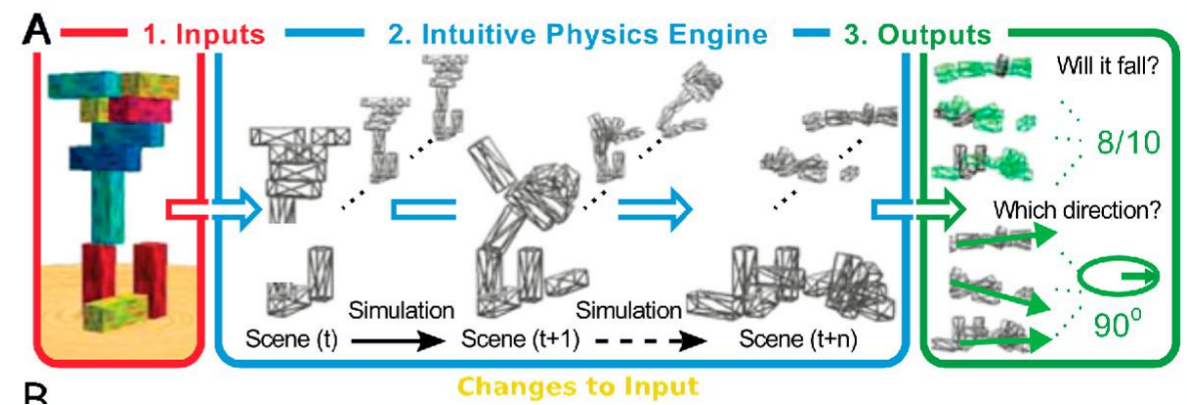
AlphaGo Overview



Acquire states & policies
 Rely on **human** pre-labeling
 Search state space autonomously

Silver, et al. (2016). Nature

Hierarchical Bayesian



Acquire states
 Rely on **human** pre-labeling
 Rely on **prior rule set** to reason on input states

Lake 2015 Science; (2018). BBS

Wish for Human level competence

Piekniewski's blog

On limits of deep learning and where to go next with AI.

[MAIN SITE](#) [BLOG HOME](#) [ABOUT](#)

AI Winter Is Well On Its Way

POSTED [MAY 28, 2018](#) BY [FILIP PIEKNIIEWSKI](#)



Deep learning has been at the forefront of the so called AI revolution for quite a few years now, and many people had believed that it is the silver bullet that will take us to the world of wonders of technological singularity (general AI). Many bets were made in 2014, 2015 and 2016 when still new boundaries were pushed, such as the Alpha Go etc.

Basic info

I'm proposing a new machine learning meta-architecture for learning forward models. The architecture is called [Predictive Vision Model \(PVM\)](#). In this blog I present my thoughts on how PVM relates to deep learning and the global AI landscape. Occasionally I'll blog about my sysadmin projects and sci-fi.

[See a short bio note here.](#)

SHARE **IN DEPTH** **COMPUTER SCIENCE**



Has artificial intelligence become alchemy?

Matthew Hutson

[+ See all authors and affiliations](#)

Science 04 May 2018:
Vol. 360, Issue 6388, pp. 478
DOI: 10.1126/science.360.6388.478

Article

[Figures & Data](#)

[Info & Metrics](#)

[eLetters](#)

[PDF](#)

You are currently viewing the summary.

[View Full Text](#)

Summary

Ali Rahimi, a researcher in artificial intelligence (AI) at Google in San Francisco, California, has charged that machine learning algorithms, in which computers learn through trial and error, have become a form of "alchemy." Researchers, he says, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another. Now, in a paper presented on 30 April at the International Conference on Learning Representations in Vancouver, Canada, Rahimi and his collaborators document examples of what they see as the alchemy problem and offer prescriptions for bolstering AI's rigor. The issue is distinct from AI's reproducibility problem, in which researchers can't replicate each other's results because of inconsistent experimental and publication practices. It also differs from the "black box" or



中南大学法国招聘会

办会时间：2018年6月10日下午3点

SHARE

IN DEPTH | COMPUTER SCIENCE



0



0

Has artificial intelligence become alchemy?

Matthew Hutson

[+ See all authors and affiliations](#)

Science 04 May 2018:

Vol. 360, Issue 6388, pp. 478

DOI: 10.1126/science.360.6388.478



THE AGE OF SURVEILLANCE CAPITALISM

**THE FIGHT FOR A
HUMAN FUTURE
AT THE NEW
FRONTIER OF POWER**

**SHOSHANA
ZUBOFF**

We will all die because of AI



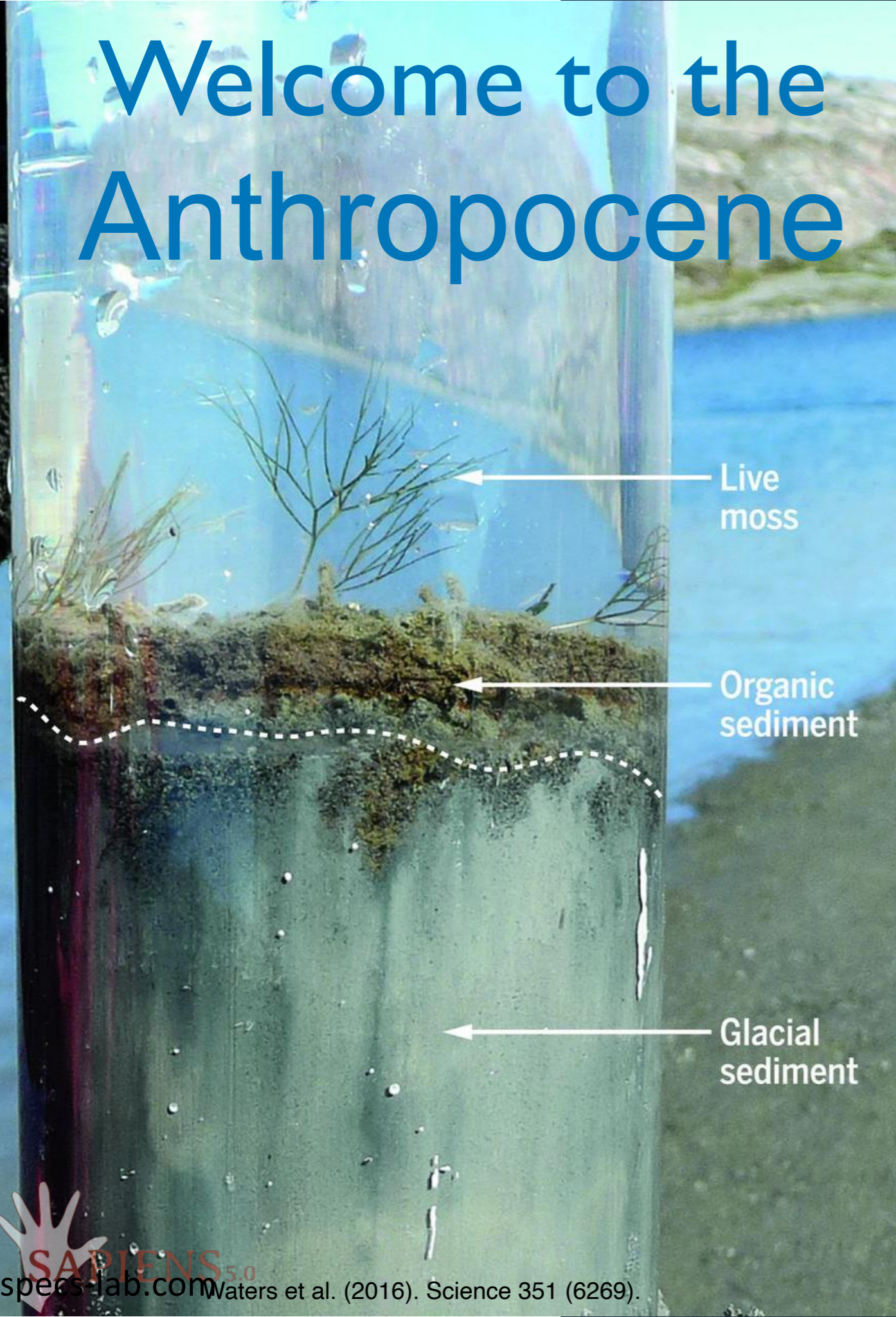
Discover the Unknown Unknowns
without external guidance

Epistemic Autonomy



Artificial Intelligence — Human Stupidity

Welcome to the Anthropocene



Live moss

Organic sediment

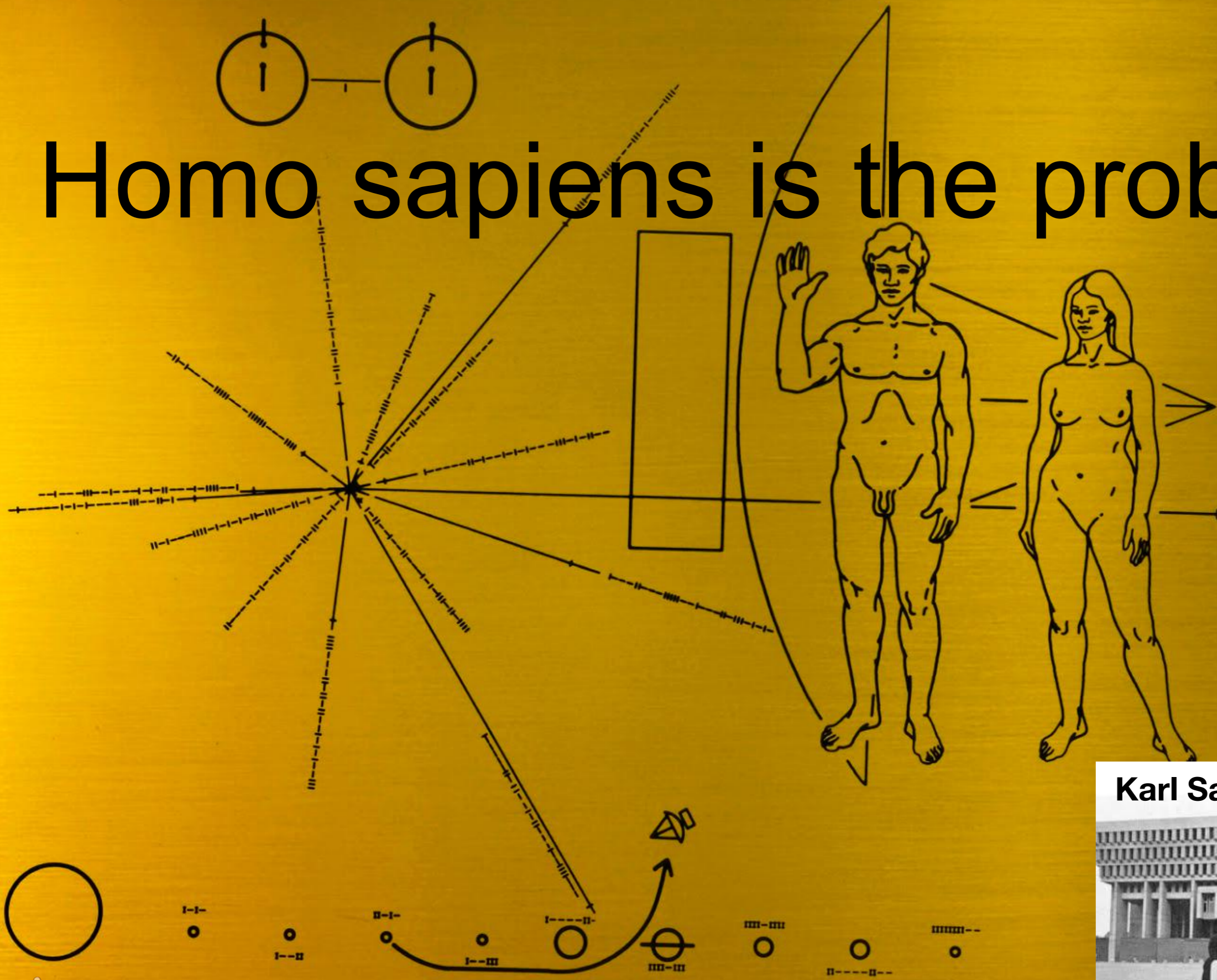
Glacial sediment



climateandsecurity.org/

1972

Homo sapiens is the problem



1972 Pioneer 1; 1973 Pioneer 2

FINALIST FOR THE PULITZER PRIZE 2011

A boldly macho book... Its thesis is simple and persuasive: The things that we do have a physical effect on our brains... What looks like logic, can argue, may be closer to madness... The internet is a distraction machine! *Sam Lashby Sunday Times*

THE SHALLOWS

How the internet is changing the way we think, read and remember



NICHOLAS CARR

 South Scarfolk College *Preparing YOU for the ochlocratic future.*

De-education classes

1976-1977

"Knowledge is power and power corrupts absolutely"

Non-experts predict that by the year 2010 knowledge and facts will be redundant and even harmful to you, your family and society as a whole.

JOB CENTRE

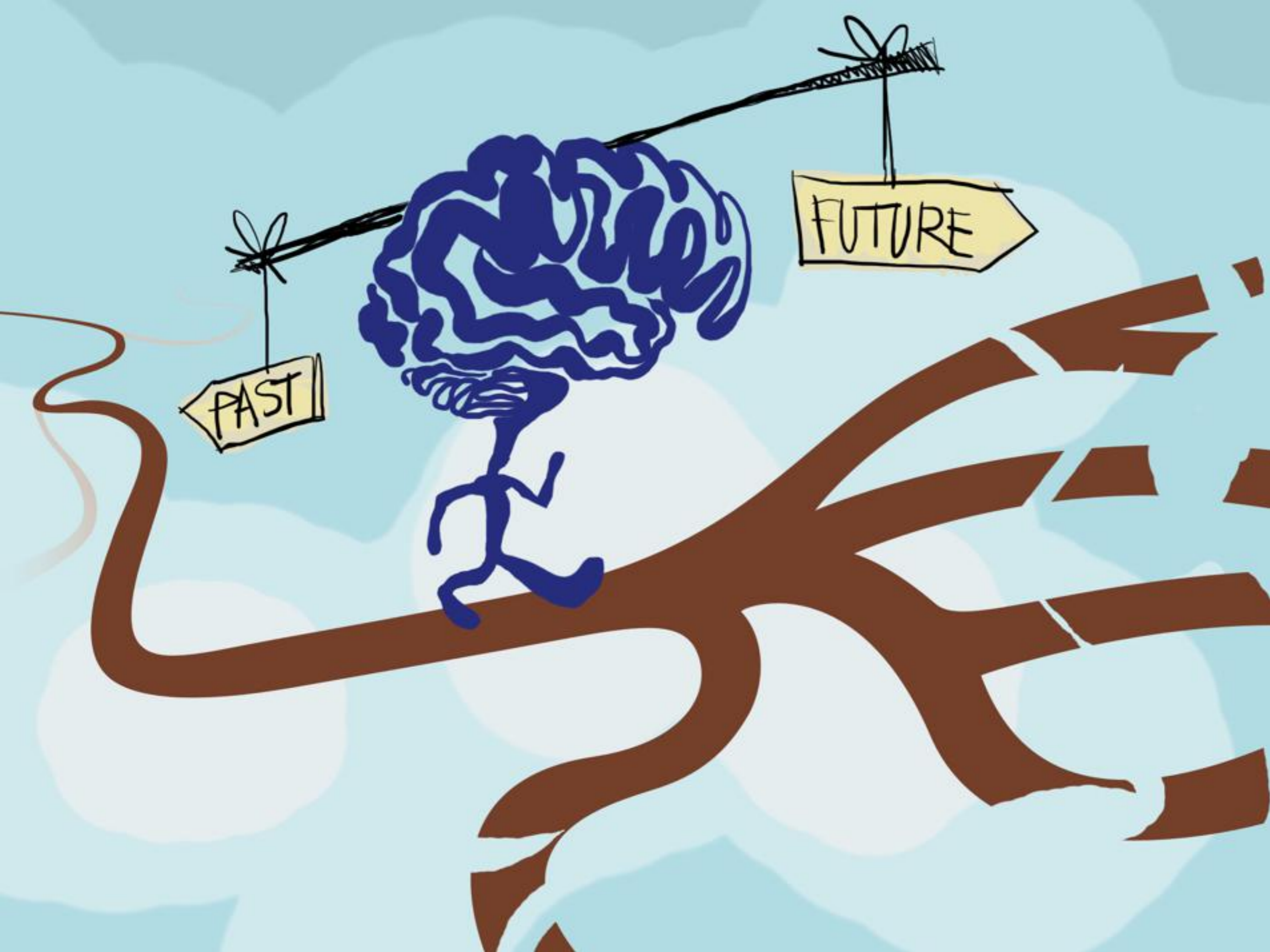
- Experts
- Academics
- Consultants
- Advisers
- Specialists
- Researchers



"I halved my knowledge and doubled the strength of my uninformed gut-feeling in only 6 MONTHS!"

Learn to unlearn

Mandatory enrolment from 13st September



FAST

FUTURE

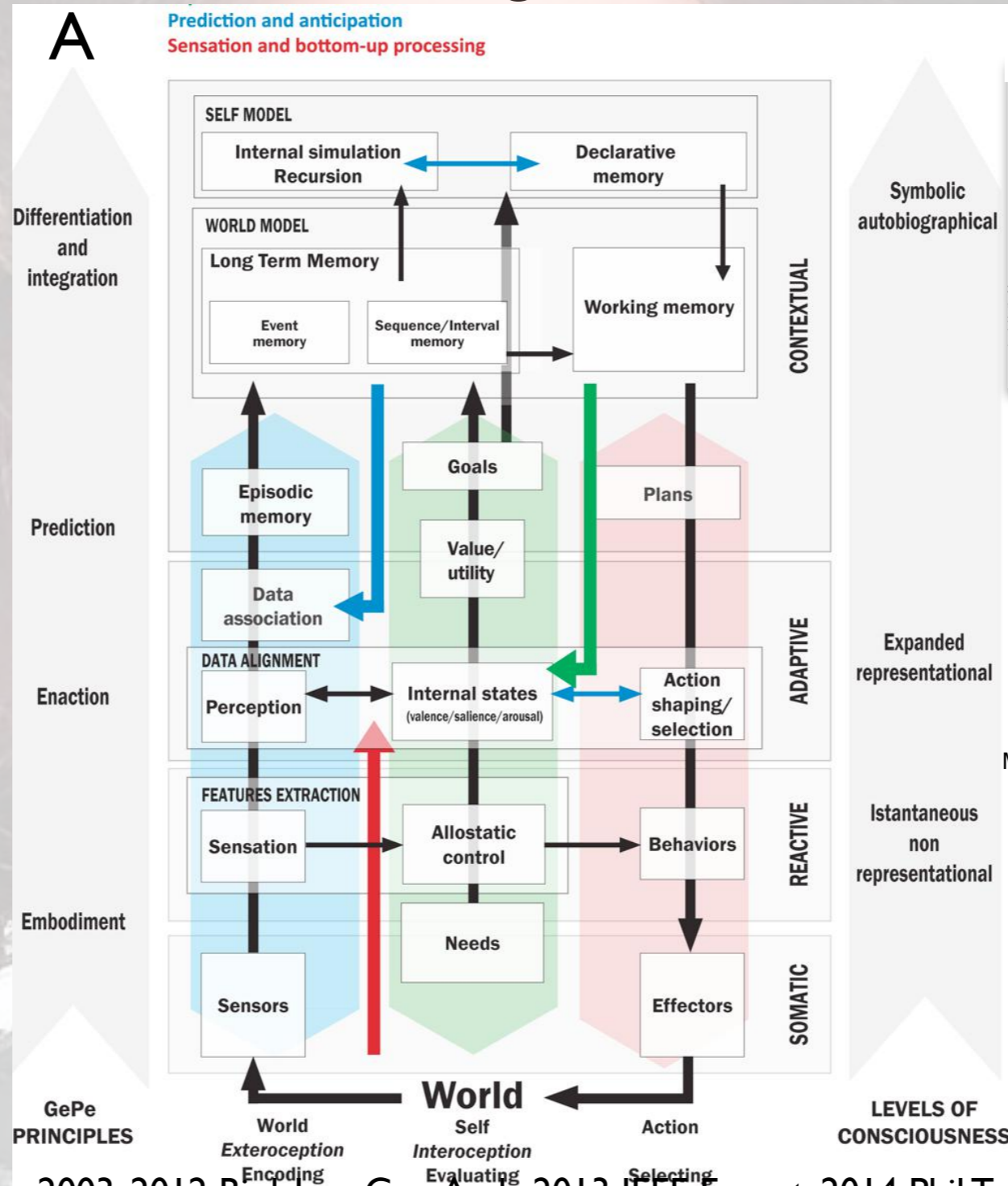
AI is mechanising Behaviorism




Autonomy, Action and Responsibility

Actus.Reus: a physical movement

Mens rea: conscious intent causing the action





Dismantling the Puppeteer

Mastering Synthetic Autonomy

Synthetic Volition