

Thinking about Thought

- Introduction
- Philosophy of Mind
- Cognitive Models
- Machine Intelligence
- Life and Organization
- Ecology
- The Brain
- Dreams and Emotions
- Language
- Modern Physics
- Consciousness

*Make it idiot proof and someone will make a better idiot
(One-liner signature file on the internet)*

Session Three: Machine Intelligence

for Piero Scaruffi's class

"Thinking about Thought"

at UC Berkeley (2014)

Roughly These Chapters of My Book “Nature of Consciousness”:

3. Machine Intelligence

5. Common Sense: Engineering The Mind

6. Connectionism And Neural Machines

Mind and Machines

Is our mind a machine?

Can we build one?

Mathematical models of the mind

Descartes: water fountains

Freud: a hydraulic system

Pavlov: a telephone switchboard

Wiener: a steam engine

Simon: a computer

The computer is the first machine that can be programmed to perform different tasks

A Brief History of Logic

Pythagoras's theorem (6th c BC): a relationship between physical quantities that is both abstract and eternal

Euclides' "Elements" (350 BC), the first system of Logic, based on just 5 axioms

Aristoteles' "Organon" (4th c BC): syllogisms

William Ockham's "Summa Totius Logicae" (1300 AD) on how people reason and learn

Francis Bacon's "Novum Organum" (1620)

Rene' Descartes' "Discours de la Methode" (1637): the analytic method over the dialectic method

Gottfried Leibniz 's "De Arte Combinatoria" (1676)

Leonhard Euler (1761) how to do symbolic logic with diagrams

Augustus De Morgan's "The Foundations of Algebra" (1844) 4

A Brief History of Logic

George Boole's "The Laws Of Thought" (1854): the laws of logic “are” the laws of thought

Propositional logic and predicate logic: true/false!

$F \vee G = G \vee F;$ $F \wedge G = G \wedge F.$ **commutativity**

$(F \vee G) \vee H = F \vee (G \vee H);$ $(F \wedge G) \wedge H = F \wedge (G \wedge H).$ **associativity**

$F \vee (G \wedge H) = (F \vee G) \wedge (F \vee H).$ **distributivity**

$F \wedge (G \vee H) = (F \wedge G) \vee (F \wedge H).$ **distributivity**

$\neg (F \vee G) = \neg F \wedge \neg G;$ $\neg (F \wedge G) = \neg F \vee \neg G.$ **De Morgan**

connectives = $\{\supset, \neg, \vee, \wedge, \equiv\}$

quantifiers = $\{\forall, \exists\}$

$(\exists x)\text{Tet}(x)$

$(\forall x)(\text{Tet}(x) \rightarrow \text{Large}(x))$

$((\forall x)(\text{Tet}(x) \rightarrow \text{Large}(x)) \rightarrow (\text{Tet}(x) \rightarrow \text{Large}(x)))$

$\text{Tet}(x) \rightarrow \text{Large}(x)$

$\text{Large}(x) \rightarrow (\exists x)\text{Large}(x)$

$\text{Tet}(x) \rightarrow (\exists x)\text{Large}(x)$

$(\exists x)\text{Tet}(x) \rightarrow (\exists x)\text{Large}(x)$

$(\exists x)\text{Large}(x)$

A Brief History of Logic

Axiomatization of Thought:

Gottlob Frege's "Foundations of Arithmetic"
(1884)

Giuseppe Peano's "Arithmetices Principia
Nova Methodo Exposita" (1889)

Bertrand Russell's "Principia Mathematica"
(1903)



The Axiomatic Method

- David Hilbert (1900)
 - Hilbert's second problem: prove the consistency of the axioms of arithmetic



The Axiomatic Method

- David Hilbert (1928)
 - Entscheidungsproblem problem: the mechanical procedure for proving mathematical theorems
 - An algorithm, not a formula
 - Mathematics = blind manipulation of symbols
 - Formal system = a set of axioms and a set of inference rules
 - Propositions and predicates
 - Deduction = exact reasoning
 - Logic emancipated from reality by dealing purely with abstractions



The Axiomatic Method

- Paradoxes
 - "I am lying"
 - The class of classes that do not belong to themselves (the barber who shaves all barbers who do not shave themselves)
 - The omnipotent god



The Axiomatic Method



- Kurt Gödel (1931)
 - The answer to Hilbert's second problem
 - Any formal system contains an "undecidable" proposition
 - A concept of truth cannot be defined within a formal system
 - Impossible to reduce logic to a mechanical procedure to prove theorems ("decision problem")

The Axiomatic Method



- Alfred Tarski (1935)
 - Definition of “truth”: a statement is true if it corresponds to reality (“correspondence theory of truth”)
 - Truth is defined in terms of physical concepts
 - Logic grounded back into reality

The Axiomatic Method



- Alfred Tarski (1935)
 - Base meaning on truth, semantics on logic (truth-conditional semantics)
 - “Interpretation” and “model” of a theory (“model-theoretic” semantics)
 - Theory = a set of formulas.
 - Interpretation of a theory = a function that assigns a meaning (a reference in the real world) to each formula
 - Model for a theory = interpretation that satisfies all formulas of the theory
 - The universe of physical objects is a model for physics

The Axiomatic Method



- Alfred Tarski (1935)
 - Build models of the world which yield interpretations of sentences in that world
 - Meaning of a proposition=set of situations in which it is true
 - All semantic concepts are defined in terms of truth
 - Meaning grounded in truth
 - Truth can only be relative to something
 - “Meta-theory”

The Axiomatic Method

- Alan Turing (1936)
 - Hilbert's challenge (1928): an algorithm capable of solving all the mathematical problems
 - Turing Machine (1936): a machine whose behavior is determined by a sequence of symbols and whose behavior determines the sequence of symbols
 - A universal Turing machine (UTM) is a Turing machine that can simulate an arbitrary Turing machine



The Axiomatic Method

- Alan Turing (1936)
 - Computation = the formal manipulation of symbols through the application of formal rules
 - Turing machine = a machine that is capable of performing any type of computation
 - Turing machine = the algorithm that Hilbert was looking for
 - Hilbert's program reduced to manipulation of symbols
 - Problem solving = symbol processing



The Axiomatic Method



- The Turing machine:
 - ...an infinite tape (an unlimited memory capacity)
 - ... marked out into squares, on each of which a symbol can be printed...
 - The machine can read or write one symbol at a time
 - At any moment there is one symbol in the machine, the scanned symbol
 - The machine can alter the scanned symbol based on that symbol and on a table of instructions
 - The machine can also move the tape back and forth

7. Detailed description of the universal machine.

A table is given below of the behaviour of this universal machine. The m -configurations of which the machine is capable are all those occurring in the first and last columns of the table, together with all those which occur when we write out the unabbreviated tables of those which appear in the table in the form of m -functions. *E.g.*, $e_1(anf)$ appears in the table and is an m -function. Its unabbreviated table is (see p. 239)

$e_1(anf)$	$\left\{ \begin{array}{l} o \\ \text{not } o \end{array} \right.$	$\left\{ \begin{array}{l} R \\ L \end{array} \right.$	$\left\{ \begin{array}{l} e_1(anf) \\ e(anf) \end{array} \right.$
$e_2(anf)$	$\left\{ \begin{array}{l} \text{Any} \\ \text{None} \end{array} \right.$	$\left\{ \begin{array}{l} R, E, R \\ \text{None} \end{array} \right.$	$\left\{ \begin{array}{l} e_1(anf) \\ anf \end{array} \right.$

Consequently $e_1(anf)$ is an m -configuration of U .

When U is ready to start work the tape running through it bears on it the symbol o on an F -square and again o on the next F -square; after this, on F -squares only, comes the S.D. of the machine followed by a double colon "::" (a single symbol, on an F -square). The S.D. consists of a number of instructions, separated by semi-colons.

Each instruction consists of five consecutive parts

(i) " D " followed by a sequence of letters " A ". This describes the relevant m -configuration.

(ii) " D " followed by a sequence of letters " C ". This describes the scanned symbol.

(iii) " D " followed by another sequence of letters " C ". This describes the symbol into which the scanned symbol is to be changed.

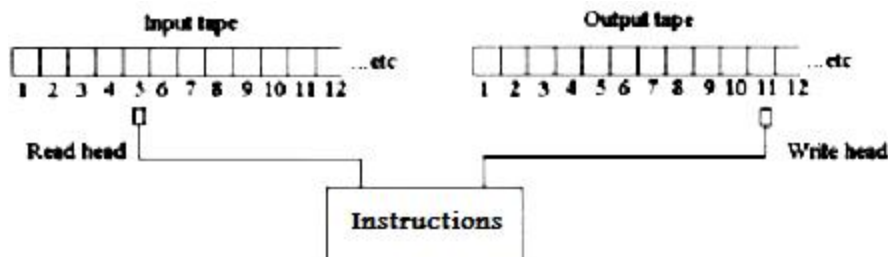
(iv) " L ", " R ", or " N ", describing whether the machine is to move to left, right, or not at all.

(v) " D " followed by a sequence of letters " A ". This describes the final m -configuration.

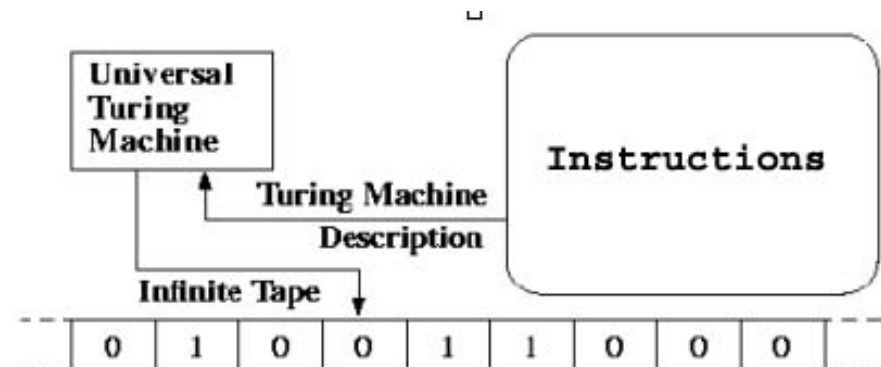
The machine U is to be capable of printing " A ", " C ", " D ", " o ", " 1 ", " u ", " y ", " w ", " x ", " y ", " z ". The S.D. is formed from " $;$ ", " A ", " C ", " D ", " L ", " R ", " N ".

The Axiomatic Method

- Alan Turing (1936)
 - Universal Turing Machine: a Turing machine that is able to simulate any other Turing machine
 - The universal Turing machine reads the description of the specific Turing machine to be simulated



Turing Machine



The Axiomatic Method

- Turing machines in nature: the ribosome, which translates RNA into proteins
 - Genetic alphabet: nucleotides ("bases"): A, C, G, U
 - The bases are combined in groups of 3 to form "codons"
 - RNA is composed of a string of nucleotides ("bases") according to certain rules
 - There are special carrier molecules ("tRNA") that are attached to specific aminoacids (proteins)
 - The start codon encodes the aminoacid Methionine
 - A codon is matched with a specific tRNA
 - The new aminoacid is attached to the protein
 - The tape then advances 3 bases to the next codon, and the process repeats
 - The protein keeps growing
 - When the "stop" codon is encountered, the ribosome dissociates from the mRNA

The Axiomatic Method

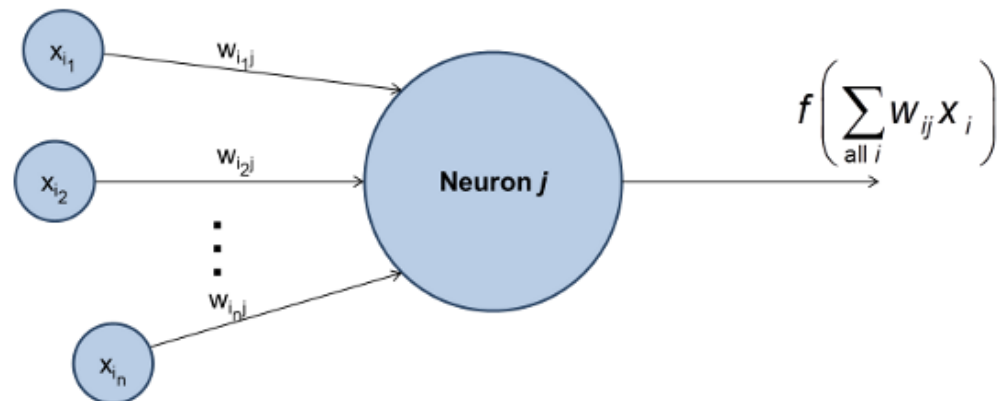
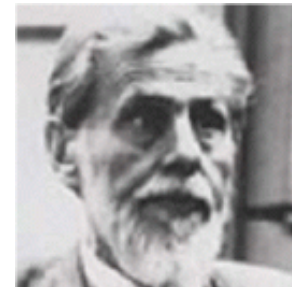
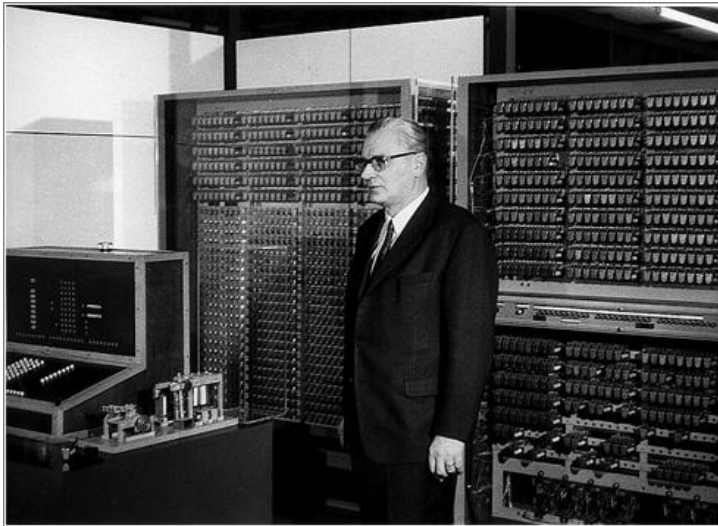


- Alan Turing (1936)
 - Computable functions = “recursive” functions
 - Recursion, the Lambda-calculus, and the Turing machine are equivalent
 - Each predicate is defined by a function, each function is defined by an algorithm

Computer Science

1941: Konrad Zuse's Z3 programmable electronic computer

1943: Warren McCulloch's and Walter Pitts' binary neuron

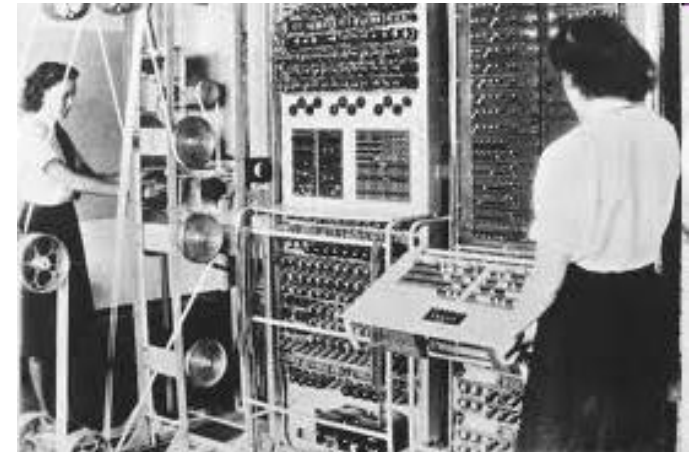


Computer Science

- World War II:
 - Breaking the Enigma code (Bombe)
 - Turing worked at Bletchley Park where the Colossus was built but it was not a universal Turing machine (not general purpose)



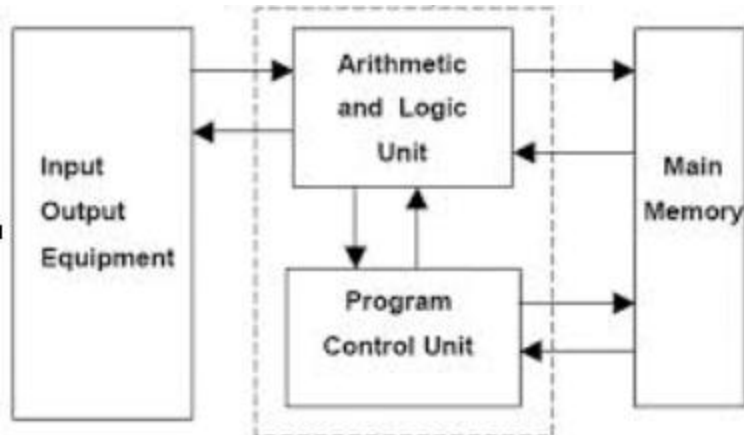
Replica of the Bombe



Computer Science

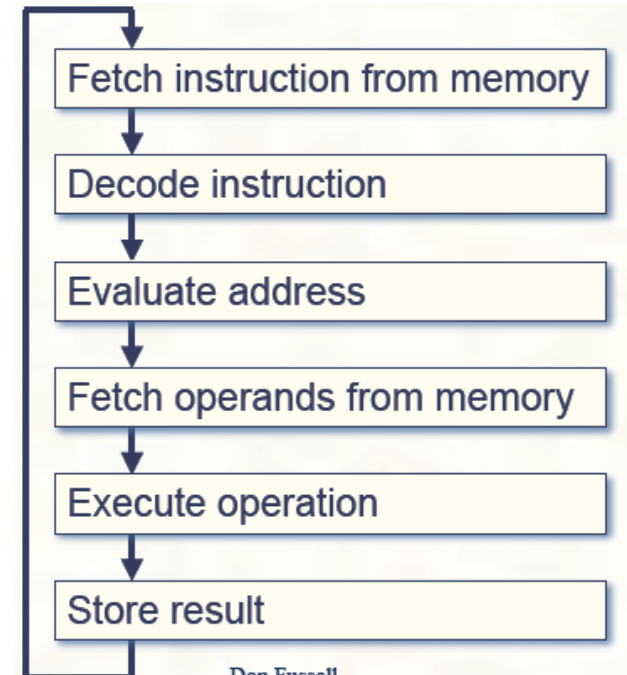
1945: John Von Neumann's computer architecture

- Separation of instructions and data (although both are sequences of 0s and 1s)
- Sequential processing



Control unit:

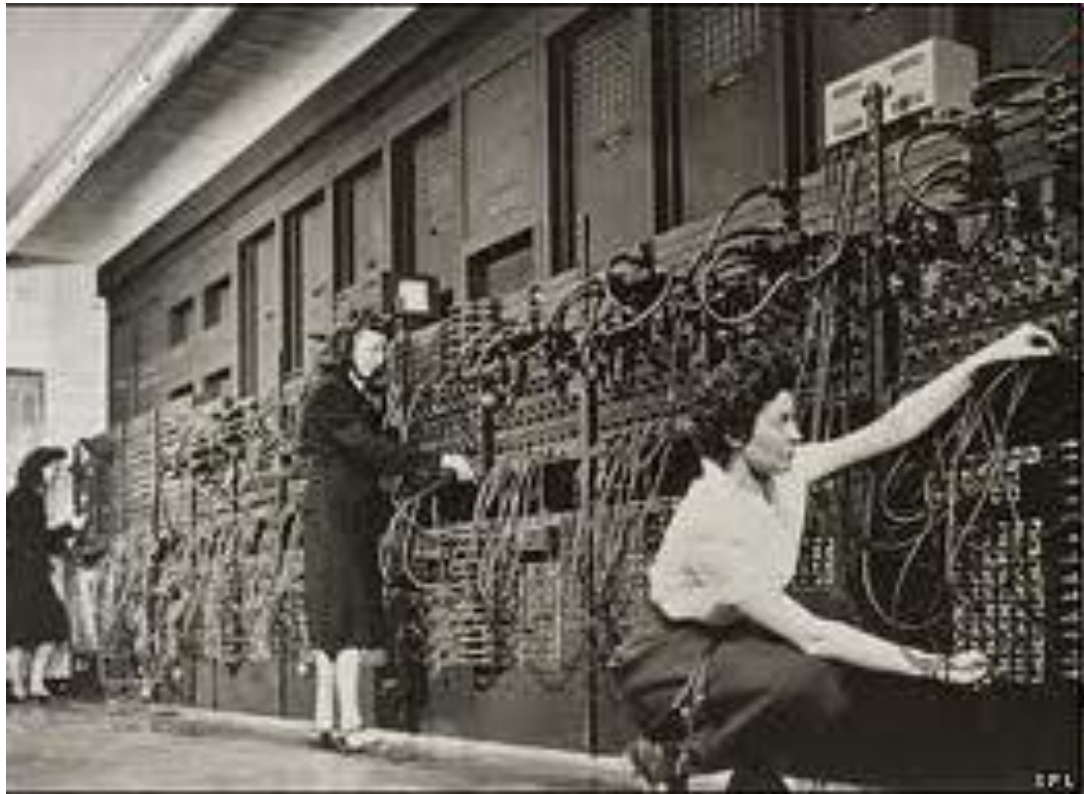
- reads an instruction from memory
- interprets/executes the instruction
- signals the other components what to do



Don Fussell

Computer Science

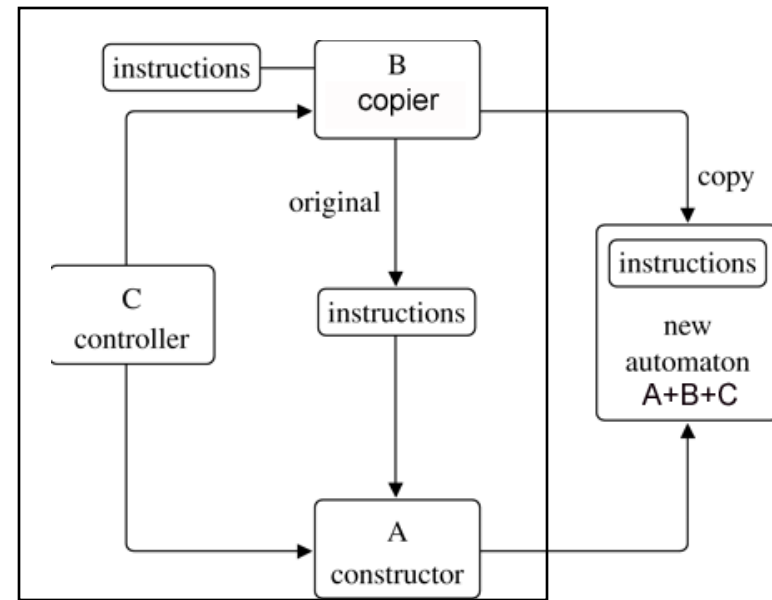
- The first non-military computer, ENIAC, by John Mauchly and Presper Eckert (1946)



Computer Science

1947: John Von Neumann's self-reproducing automata

1. A universal constructor A, which can construct another automaton according to instruction I.
2. A copier B, which can make a copy of the instruction tape I.
3. A controller C, which combines A and B and allows A to construct a new automaton according to I, B to copy instructions from I and attach them to the newly created automaton
4. An instruction tape describing how to construct automaton $A+B+C$



Cybernetics

1943: Beginning of Cybernetics

1946: The first Macy Conference on Cybernetics

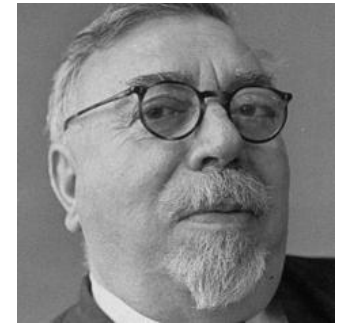
- John von Neumann (computer science),
- Norbert Wiener, Walter Pitts (mathematics),
- Rafael Lorente de No, Ralph Gerard, etc (neurology),
- Warren McCulloch (neuropsychiatry),
- Gregory Bateson, Margaret Mead (anthropology),
- Heinrich Kluever, Lawrence Kubie, etc (psychology)
- Filmer Northrop (philosophy),
- Lawrence Frank, Paul Lazarsfeld (sociology)



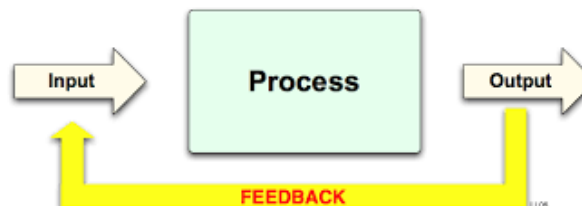
Cybernetics

Norbert Wiener (1947)

- Bridge between machines and nature, between "artificial" systems and natural systems
- Feedback, by sending back the output as input, helps control the work of the machine
- A control system is realized by a loop of action and feedback
- A control system is capable of achieving a "goal", is capable of "purposeful" behavior
- Living organisms are control systems



VS



Cybernetics

- Nicholas Bernstein (1920s): self-regulatory character of the human nervous system
- Walter Cannon (1930s): Feedback is crucial for "homeostasis", the process by which an organism tends to compensate variations in the environment in order to maintain its internal stability
- Message
- Noise
- Information
- Entropy = a measure of disorder = a measure of the lack of information



Cybernetics

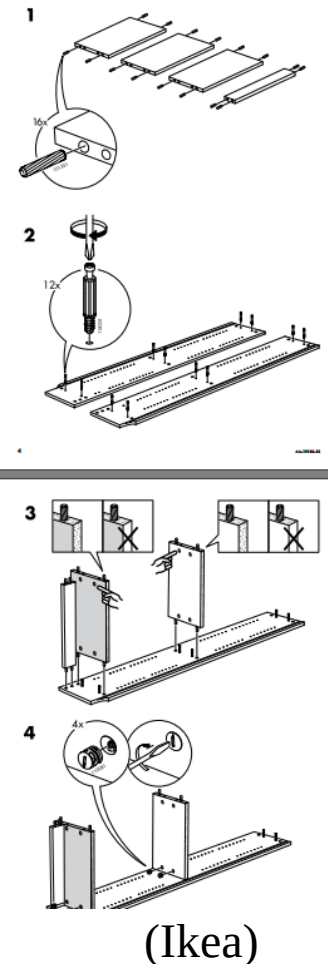


Norbert Wiener (1947)

- Information is the opposite of entropy:
 - *"The amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization"*
 - *"The quantity that we here define as amount of information is the negative of the quantity usually defined as entropy in similar situations."*
 - *"The processes which lose information are... closely analogous to the processes which gain entropy. "*

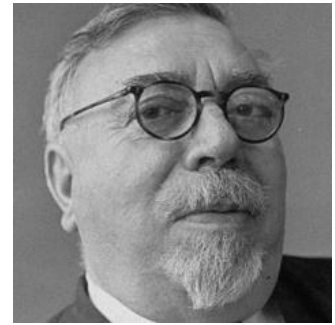
Cybernetics

- Paradigm shift from the world of continuous laws to the discrete world of algorithms
- The effect of an algorithm is to turn time's continuum into a sequence of discrete quanta, and, correspondingly, to turn an analog instrument into a digital instrument
- A watch is the digital equivalent of a sundial: the sundial marks the time in a continuous way, the watch advances by units.

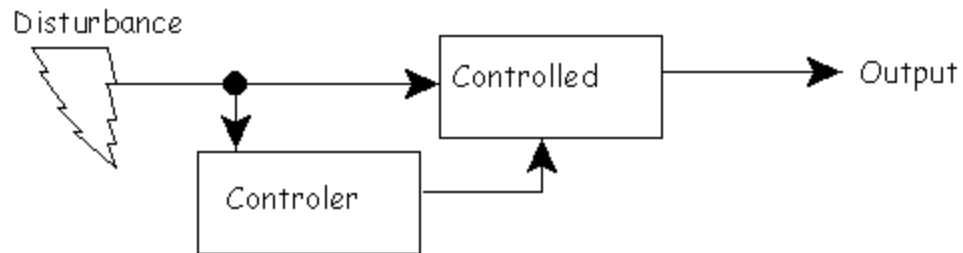


Cybernetics

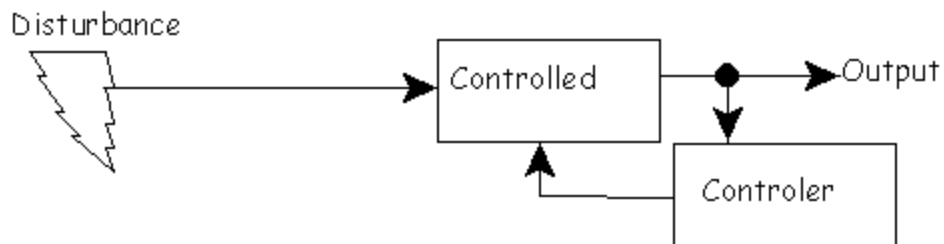
Norbert Wiener (1947)



Feed Forward



Feedback



Cybernetics

Cybernetics

- An analog instrument can be precise, and there is no limit to its precision.
- A digital instrument can only be approximate, its limit being the smallest magnitude it can measure



Titan Octane 9308KM01

Cybernetics



Ross Ashby (1952)

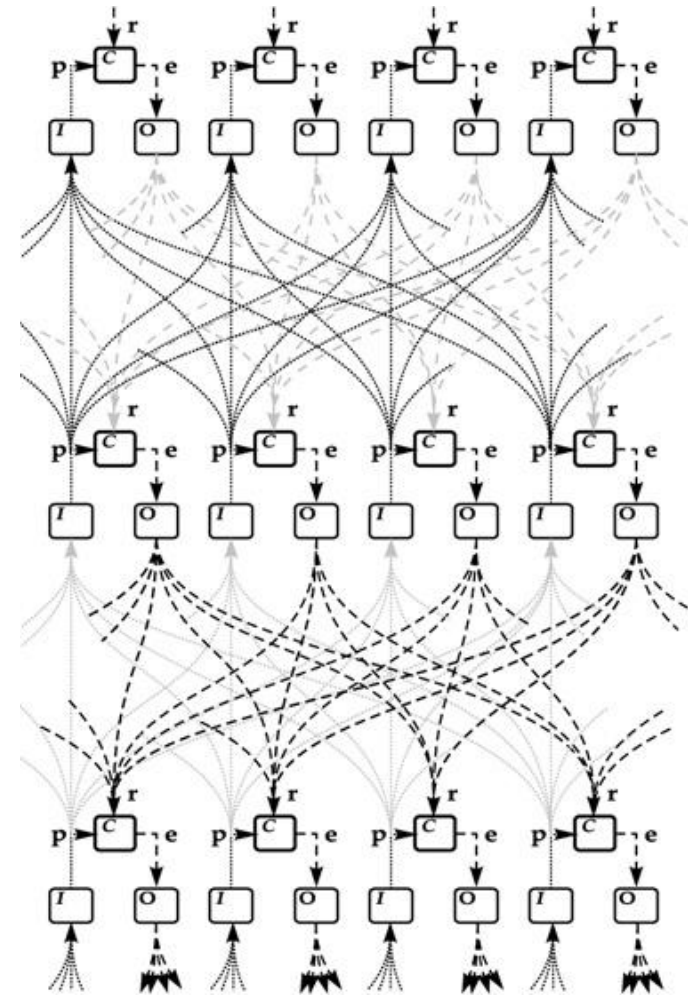
- Both machines and living beings tend to change in order to compensate variations in the environment, so that the combined system is stable
- The "functioning" of both living beings and machines depends on feedback processes
- The system self-organizes
- In any isolated system, life and intelligence inevitably develop

Cybernetics



William Powers (1973)

- Living organisms and some machines are made of hierarchies of control systems
- Behavior is a backward chain of behaviors: walking up the hierarchy one finds out why the system is doing what it is doing (e.g., it is keeping the temperature at such a level because the engine is running at such a speed because... and so forth).
- The hierarchy is a hierarchy of goals (goals that have to be achieved in order to achieve other goals in order to achieve other goals in order to...)



Cybernetics



William Powers (1973)

- Instinctive behavior is the result of the interaction between control systems that have internal goals
- In a sense, there is no learning: there is just the blind functioning of a network of control systems.
- In another sense, that "is" precisely what we call "learning": a control system at work
- A hierarchy of control systems can create the illusion of learning and of intelligence

Information Theory

Claude Shannon (1949)

- The entropy of a question is related to the probability assigned to all the possible answers to that question
- The amount of information given the probabilities of its component symbols

$$H = - \sum p(x) \log p(x)$$

- The amount of information is equal to entropy.



Information Theory

Claude Shannon (1949)

- Information

- *"Information is a measure of one's freedom of choice in selecting a message. The greater this freedom of choice, the greater the information, the greater is the uncertainty that the message actually selected is some particular one."* (Weaver)



Information Theory

Claude Shannon (1949)

- Information and entropy
 - Shannon's "information" is the opposite of Wiener's "information".
 - Shannon defines information as chaos (entropy), Wiener and Brillouin define information as order (negentropy)
 - Shannon's entropy of information is Brillouin's negentropy



Information Theory

Claude Shannon (1949)

– No interest in the meaning of the message:

- *"The fundamental problem of communication is that of reproducing at one point exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that they are selected from a set of possible meanings."*



Information Theory

Entropy

- Sadi Carnot:
 - Steam engines cannot exceed a specific maximum efficiency
 - There are no reversible processes in nature
- “*I propose to name the quantity S the entropy of the system, after the Greek word [τροπή trope], the transformation*” (Rudolf Clausius, 1865)
- Clausius: the ratio between absorbed energy and absolute temperature



$$\Delta S = \frac{\Delta Q}{T}$$

$$dS \geq 0$$

Information Theory

Entropy

- Ludwig Boltzmann: the number of molecular degrees of freedom
- Boltzmann's entropy is nothing more than Shannon's entropy applied to equiprobable microstates.
- Entropy can be understood as “missing” information
- Entropy measures the amount of disorder in a physical system; i.e. entropy measures *the lack of information* about the structure of the system



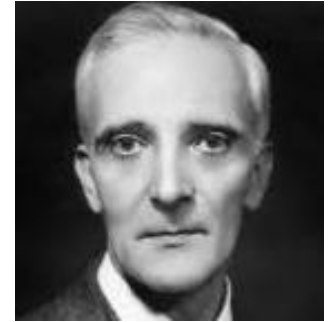
$$S \propto \log \Omega$$

$$\text{Entropy} = - \text{Information}$$

Information Theory

Entropy

- Leon Brillouin's negentropy principle of information (1953)
 - New information can only be obtained at the expense of the negentropy of some other system.

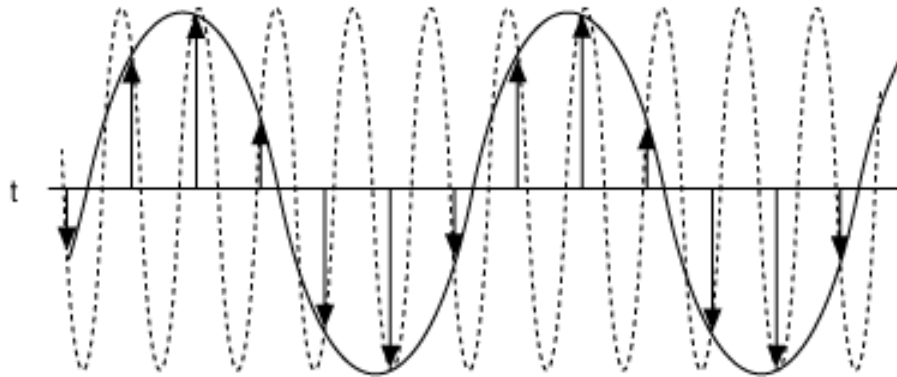


Information Theory



Vladimir Kotelnikov (1933)

- The “Nyquist–Shannon” sampling theorem: how to convert continuous signals (analog domain) into discrete signals (digital domain) and back without losing fidelity



Information Theory



Andrei Kolmogorov (1963)

- Algorithmic information theory
- Complexity = quantity of information
- Complexity of a system = shortest possible description of it = the shortest algorithm that can simulate it = the size of the shortest program that computes it
- Randomness: a random sequence is one that cannot be compressed any further
- Random number: a number that cannot be computed (that cannot be generated by any program)

Prehistory of Artificial Intelligence

Summary

- David Hilbert (1928)
- "I am lying"
- The omnipotent god
- Kurt Goedel (1931)
- Alfred Tarski (1935)
- Alan Turing (1936)
- Norbert Wiener (1947)
- Claude Shannon and Warren Weaver (1949)
- Entropy = a measure of disorder = a measure of the lack of information
- Complexity

History of Artificial Intelligence

1950: Alan Turing's "Computing Machinery and Intelligence" (the "Turing Test")

A. M. Turing (1950) Computing Machinery and Intelligence. *Mind* 49: 433-460.



COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

History of Artificial Intelligence



The Turing Test (1950)

- A machine can be said to be “intelligent” if it behaves exactly like a human being
- Hide a human in a room and a machine in another room and type them questions: if you cannot find out which one is which based on their answers, then the machine is intelligent

VOL. LIX. No. 236.]

[October, 1950]

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

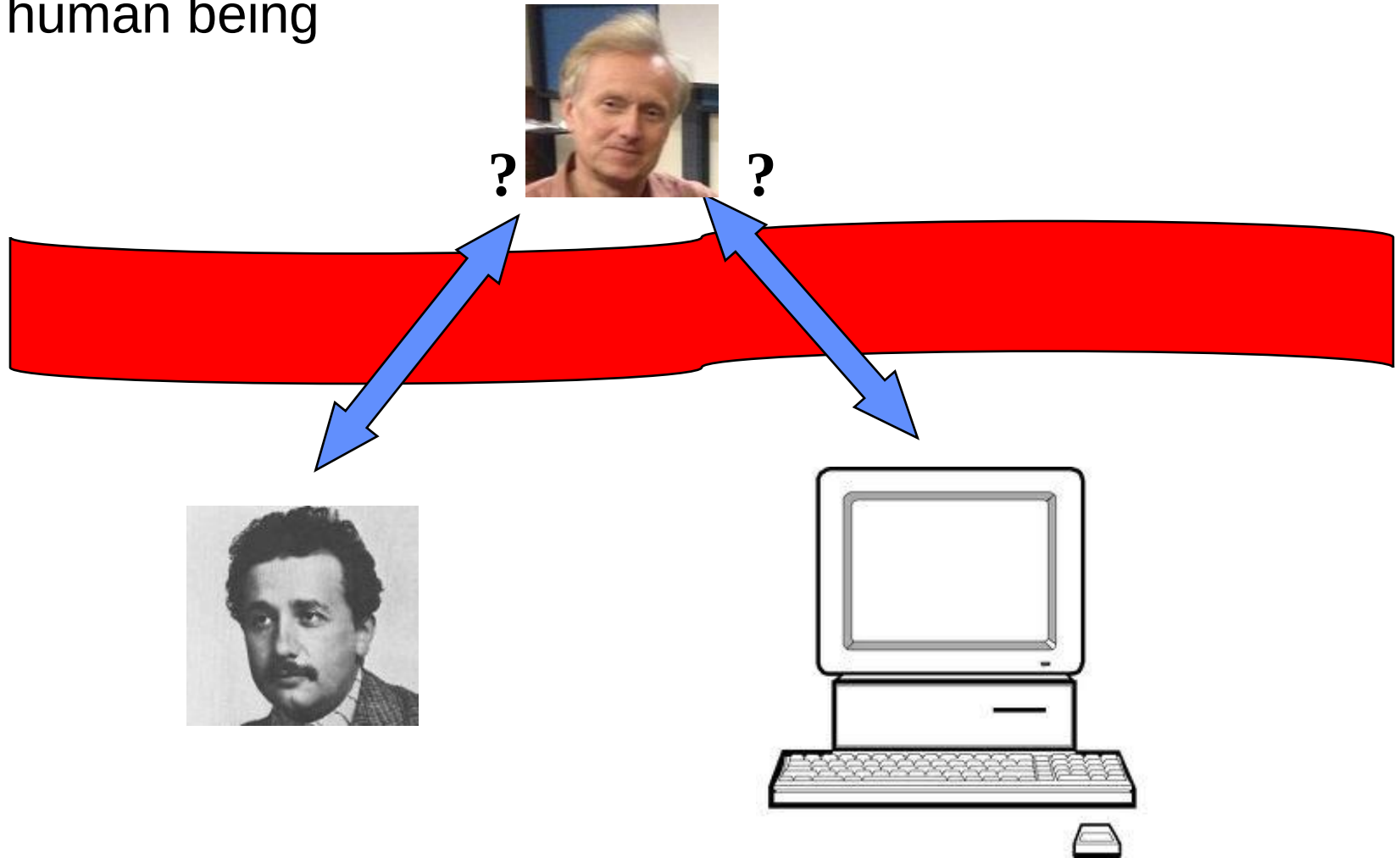
I.—COMPUTING MACHINERY AND
INTELLIGENCE

By A. M. TURING



History of Artificial Intelligence

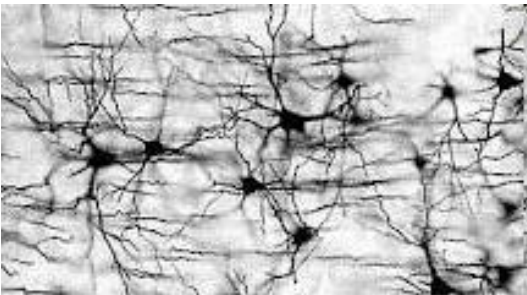
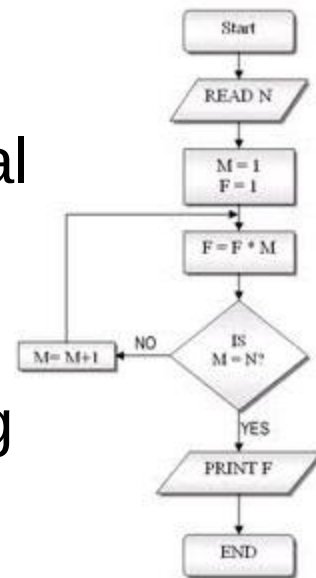
The “Turing point”: a computer can be said to be intelligent if its answers are indistinguishable from the answers of a human being



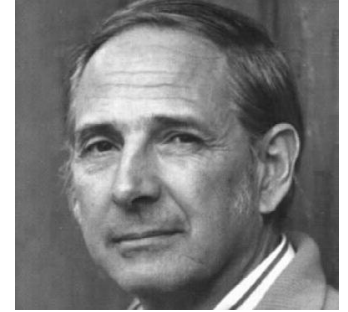
History of Artificial Intelligence

The fundamental critique to the Turing Test

- The computer cannot (qualitatively) do what the human brain does because the brain
 - does parallel processing rather than sequential processing
 - uses pattern matching rather than binary logic
 - is a connectionist network rather than a Turing machine



History of Artificial Intelligence



The Turing Test

- John Searle's Chinese room (1980)
 - Whatever a computer is computing, the computer does not "know" that it is computing it
 - A computer does not know what it is doing, therefore "that" is not what it is doing
 - Objection: The room + the machine "knows"

History of Artificial Intelligence

The Turing Test

- Hubert Dreyfus (1972):
 - Experience vs knowledge
 - Meaning is contextual
 - Novice to expert
 - Minds do not use a theory about the everyday world
 - Know-how vs know that
- Terry Winograd
 - Intelligent systems act, don't think.
 - People are “thrown” in the real world



History of Artificial Intelligence



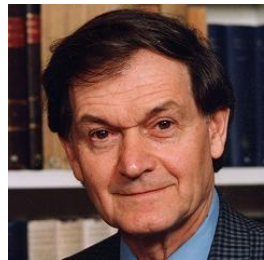
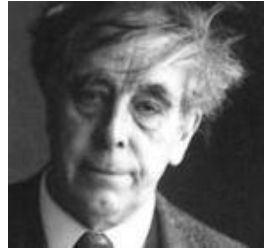
The Turing Test

- Rodney Brooks (1986)
 - Situated reasoning
 - Intelligence cannot be separated from the body.
 - Intelligence is not only a process of the brain, it is embodied in the physical world
 - Cognition is grounded in the physical interactions with the world
 - There is no need for a central representation of the world
 - Objection: Brooks' robots can't do math

History of Artificial Intelligence

The Turing Test

- John Randolph Lucas (1961) & Roger Penrose
 - Goedel's limit: Every formal system ($>$ Arithmetic) contains a statement that cannot be proved
 - Some logical operations are not computable, nonetheless the human mind can treat them (at least to prove that they are not computable)
 - The human mind is superior to a computing machine



History of Artificial Intelligence

The Turing Test

- John Randolph Lucas (1961) & Roger Penrose
 - Objection: a computer can observe the failure of “another” computer’s formal system
 - Goedel’s theorem is about the limitation of the human mind: a machine that escapes Goedel’s theorem can exist and can be discovered by humans, but not built by humans

History of Artificial Intelligence

The Turing Test

- What is measured: intelligence, cognition, brain, mind, or consciousness?
- What is measured: one machine, ..., all machines?
- What is intelligence? What is a brain? What is a mind? What is life?
- Who is the observer? Who is the judge?
- What is the instrument (instrument = observer)?
- What if a human fails the Turing test?

History of Artificial Intelligence

The ultimate Turing Test

- Build a machine that reproduces my brain, neuron by neuron, synapses by synapses
- Will that machine behave exactly like me?
- If yes, is that machine “me”?

History of Artificial Intelligence

1954: Demonstration of a machine-translation system by Leon Dostert's team at Georgetown University and Cuthbert Hurd's team at IBM



1956: Dartmouth conference on Artificial Intelligence



Artificial Intelligence (1956): the discipline of building machines that are as humans

AI Magazine Volume 27 Number 4 (2006) (© AAAI)

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

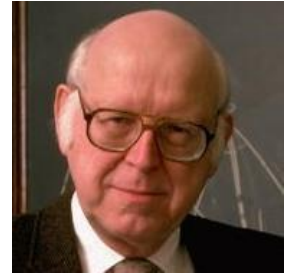
August 31, 1955

*John McCarthy, Marvin L. Minsky,
Nathaniel Rochester,
and Claude E. Shannon*

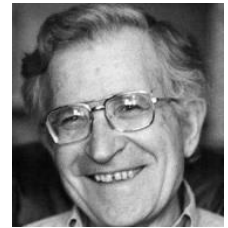
History of Artificial Intelligence

1956: Allen Newell and Herbert Simon demonstrate the "Logic Theorist", the first A.I. program, that uses "heuristics" (rules of thumb) and proves 38 of the 52 theorems in Whitehead's and Russell's "Principia Mathematica"

1957: "General Problem Solver" (1957): a generalization of the Logic Theorist but now a model of human cognition



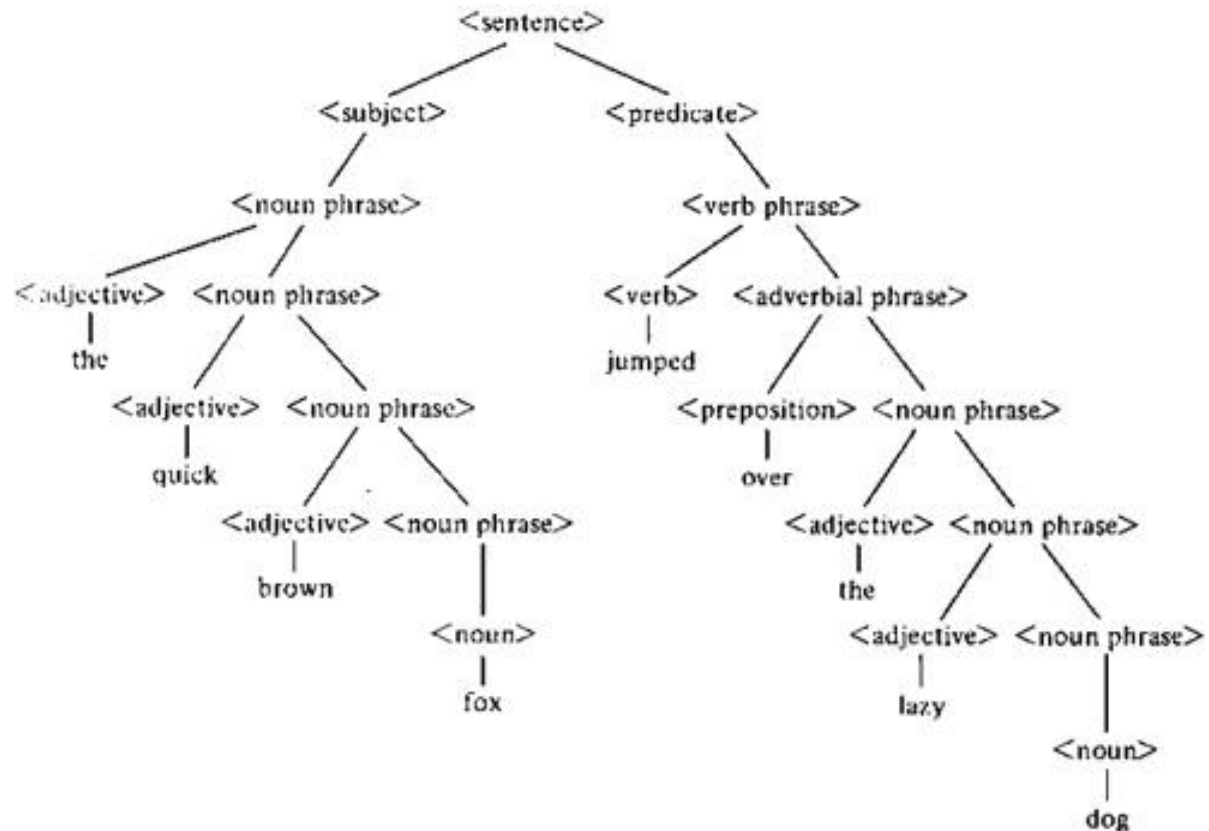
History of Artificial Intelligence



1957: Noam Chomsky's "Syntactic Structures"

- (1) S → NP + VP
- (2) VP → Verb + NP
- (3) NP → Det + N
- (4) Verb → Aux + V
- (5) Det → *the, a, ...*
- (6) N → *man, ball, ...*
- (7) Aux → *will, can, ...*
- (8) V → *hit, see, ...*

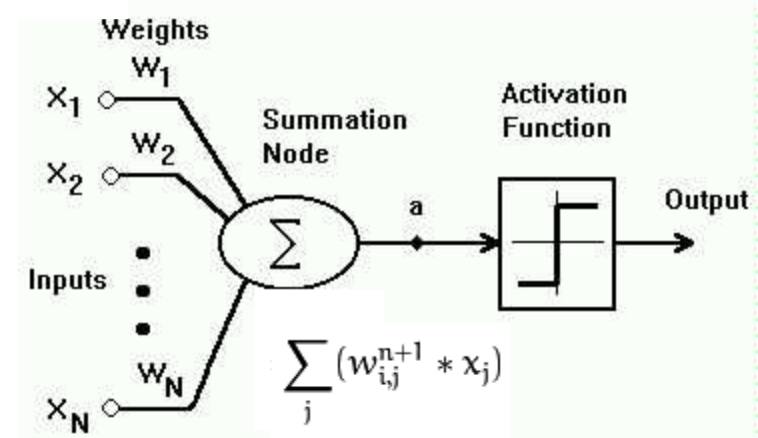
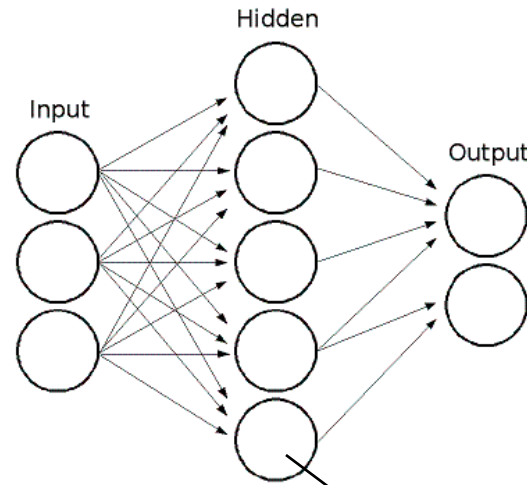
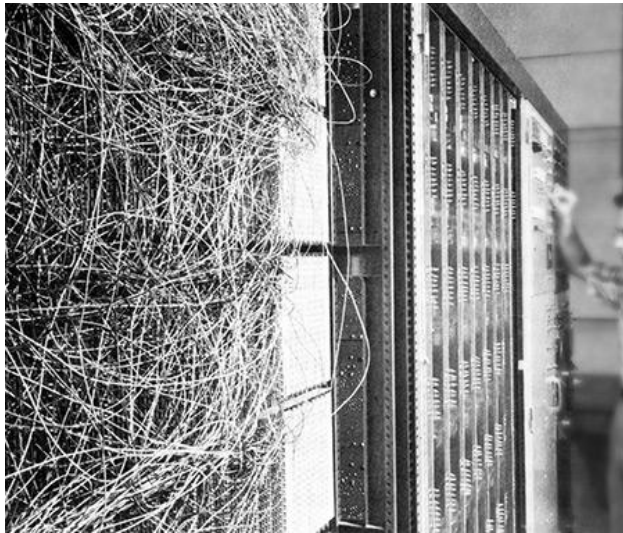
- (1) S → A + B
- (2) B → C + D
- (3) A → a + b
- (4) C → c + e + f
- (5) D → d + g + h



S stands for Sentence, NP for Noun Phrase, VP for Verb Phrase, Det for Determiner, Aux for Auxiliary (verb), N for Noun, and V for Verb stem

History of Artificial Intelligence

1957: Frank Rosenblatt's Perceptron, the first artificial neural network



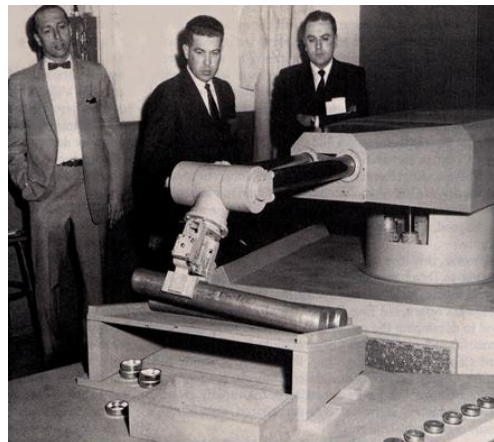
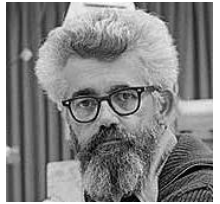
History of Artificial Intelligence

1959: John McCarthy's "Programs with Common Sense" (1949) focuses on knowledge representation

1959: Arthur Samuel's Checkers, the world's first self-learning program

1960: Hilary Putnam's Computational Functionalism (see chapter on "Philosophy of Mind")

1962: Joseph Engelberger deploys the industrial robot Unimate at General Motors



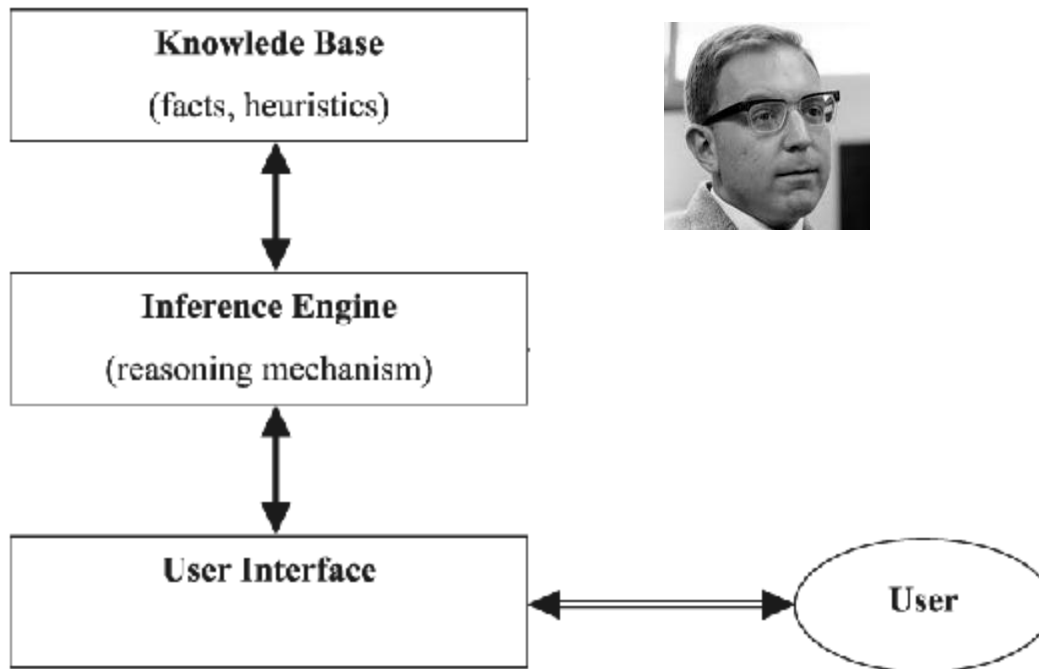
History of Artificial Intelligence



1963 Irving John Good speculates about "ultraintelligent machines" (the "singularity")

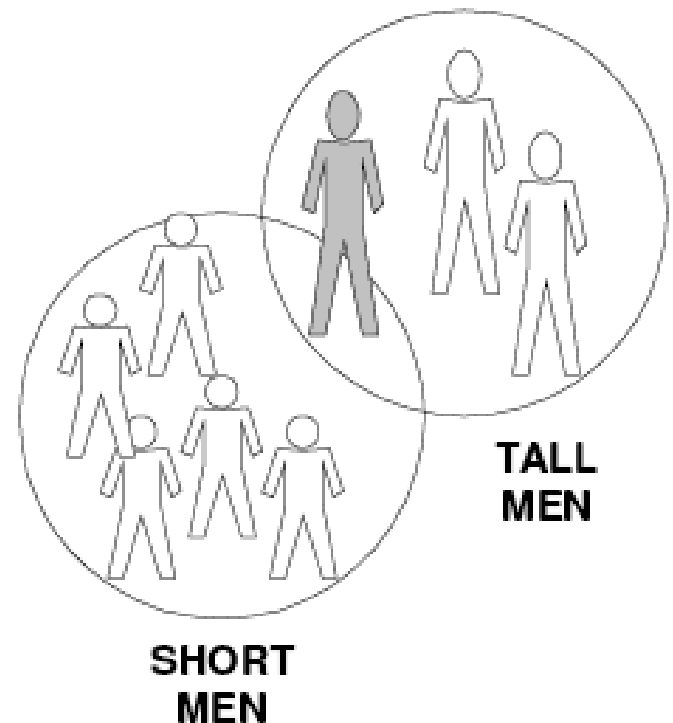
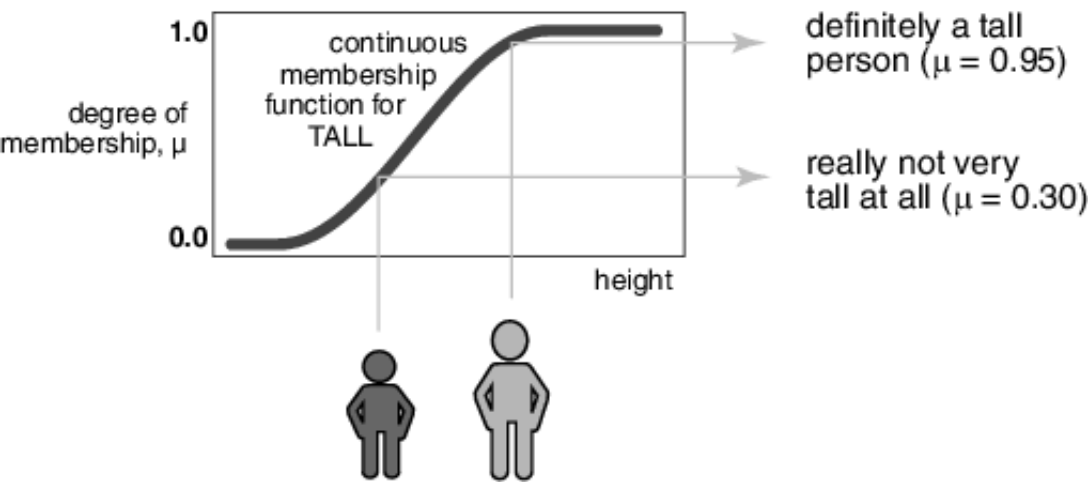
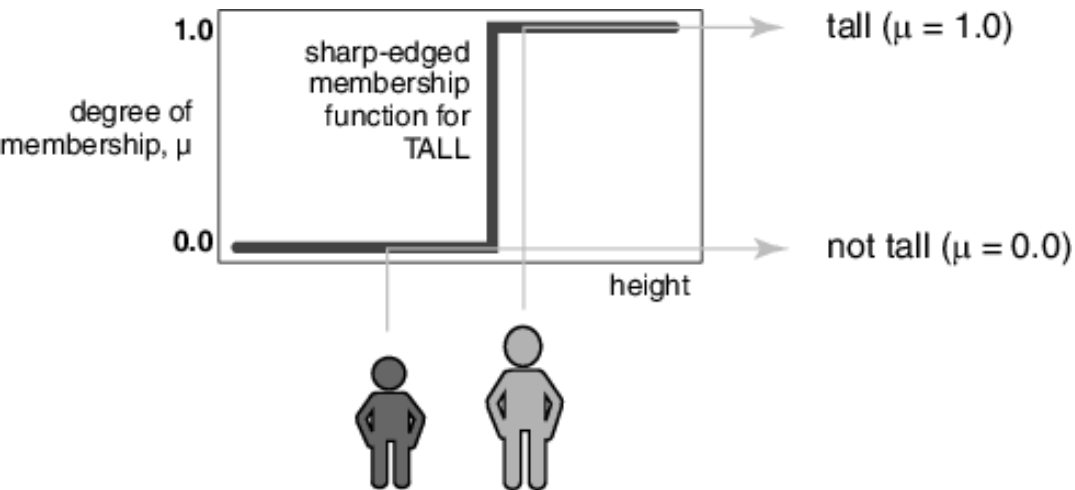
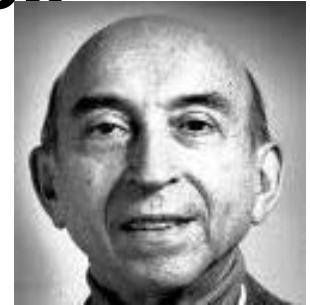
1964: IBM's "Shoebox" for speech recognition

1965: Ed Feigenbaum's Dendral expert system: domain-specific knowledge



History of Artificial Intelligence

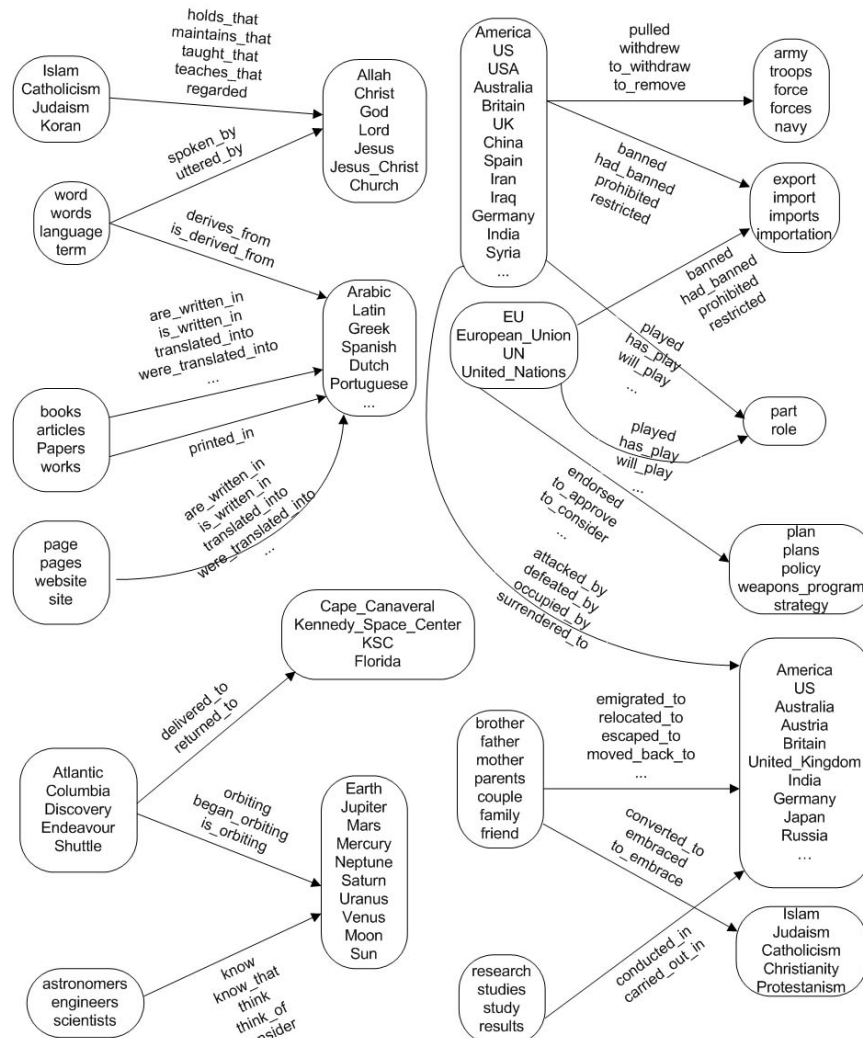
1965: Lotfi Zadeh's Fuzzy Logic



History of Artificial Intelligence



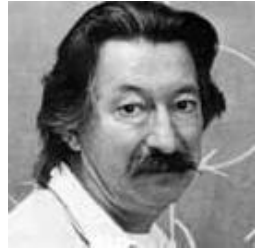
1966: Ross Quillian's semantic networks



Extracting Semantic Networks from Text via Relational Clustering
(Stanley Kok and Pedro Domingos . 2008)

History of Artificial Intelligence

1966: Joe Weizenbaum's Eliza



```
=====
EEEEEEEE L      IIIIII ZZZZZZZ      AAA
E      L      I      Z
E      L      I      Z
EEEEEE L      I      Z
E      L      I      Z
E      L      I      Z
EEEEEEEE LLLLLLLL IIIIII ZZZZZZ      A      A
=====
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA. WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE... !
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...
=====
```

1968: Peter Toma founds Systran to commercialize machine-translation systems

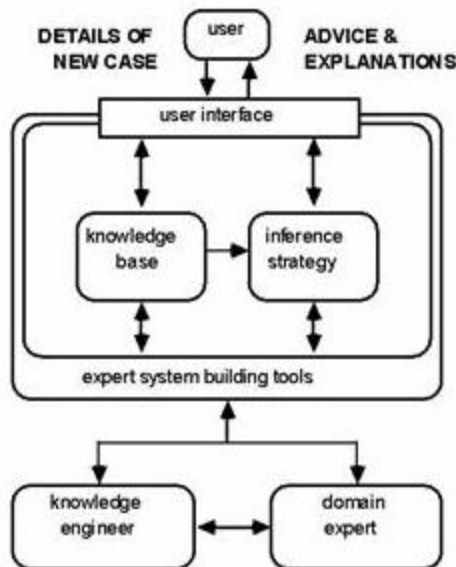


History of Artificial Intelligence

1969: Marvin Minsky & Samuel Papert's
"Perceptrons" kill neural networks

1969: Stanford Research Institute's Shakey the
Robot

1972: Bruce Buchanan's MYCIN



- a knowledge base
- a patient database
- a consultation/explanation program
- a knowledge acquisition program

Knowledge is organised as a series of IF THEN rules

History of Artificial Intelligence

1972: Terry Winograd's Shrdlu



Person: PICK UP A BIG RED BLOCK.
Computer: OK. (does it)
Person: GRASP THE PYRAMID.
Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
Person: FIND A BLOCK WHICH IS TALLER THAN THE ONE YOU ARE HOLDING AND PUT IT INTO THE BOX.
Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.
Computer: OK. (does it)
Person: WHAT DOES THE BOX CONTAIN?
Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.
Person: WHAT IS THE PYRAMID SUPPORTED BY?
Computer: THE BOX.
Person: HOW MANY BLOCKS ARE NOT IN THE BOX?
Computer: FOUR OF THEM.
Person: IS AT LEAST ONE OF THEM NARROWER THAN THE ONE WHICH I TOLD YOU TO PICK UP?
Computer: YES, THE RED CUBE.
Person: IS IT SUPPORTED?
Computer: YES, BY THE TABLE.
Person: CAN THE TABLE PICK UP BLOCKS?
Computer: NO.
Person: CAN A PYRAMID BE SUPPORTED BY A BLOCK?
Computer: YES.
Person: CAN A PYRAMID SUPPORT A PYRAMID?
Computer: I DON'T KNOW.
Person: STACK UP TWO PYRAMIDS.
Computer: I CAN'T.
Person: WILL YOU PLEASE STACK UP BOTH OF THE RED BLOCKS AND EITHER A GREEN CUBE OR A PYRAMID?
Computer: OK. (does it)
Person: WHICH CUBE IS SITTING ON THE TABLE?

History of Artificial Intelligence

1972: Hubert Dreyfus's "What Computers Can't Do"

1974: Marvin Minsky's Frame (see chapter on "Cognition")

1975: Roger Schank's Script (see chapter on "Cognition")

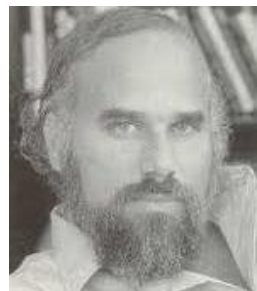
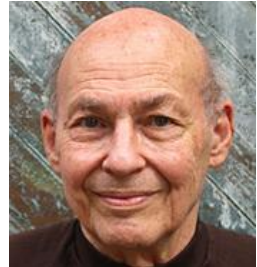
1975: John Holland's Genetic Algorithms

1976: Doug Lenat's AM

1979: Cordell Green's system for automatic programming

1979: Drew McDermott's non-monotonic logic

1979: David Marr's theory of vision

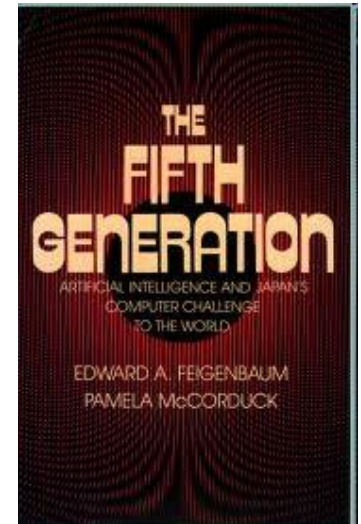
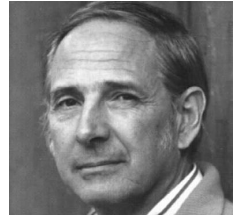


History of Artificial Intelligence

1980: John Searle's "Chinese Room"

1980: Intellicorp, the first major start-up for Artificial Intelligence

1982: Japan's Fifth Generation Computer Systems project



History of Artificial Intelligence

1982: John Hopfield describes a new generation of neural networks, based on a simulation of annealing



1983: Geoffrey Hinton's and Terry Sejnowski's Boltzmann machine for unsupervised learning



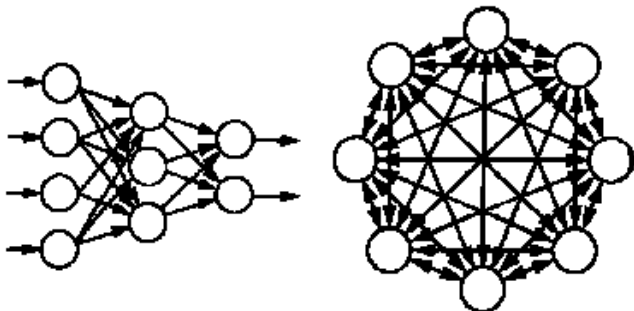
1986: Paul Smolensky's Restricted Boltzmann machine

1986: David Rumelhart's "Parallel Distributed Processing"



Rummelhart network

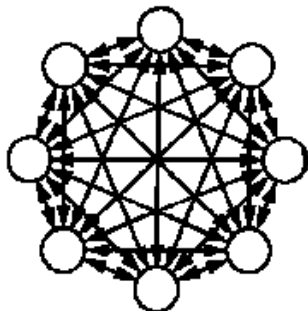
Neurons arranged in layers, each neuron linked to neurons of the neighboring layers, but no links within the same layer
Requires training with supervision



RUMMELHART

Hopfield networks

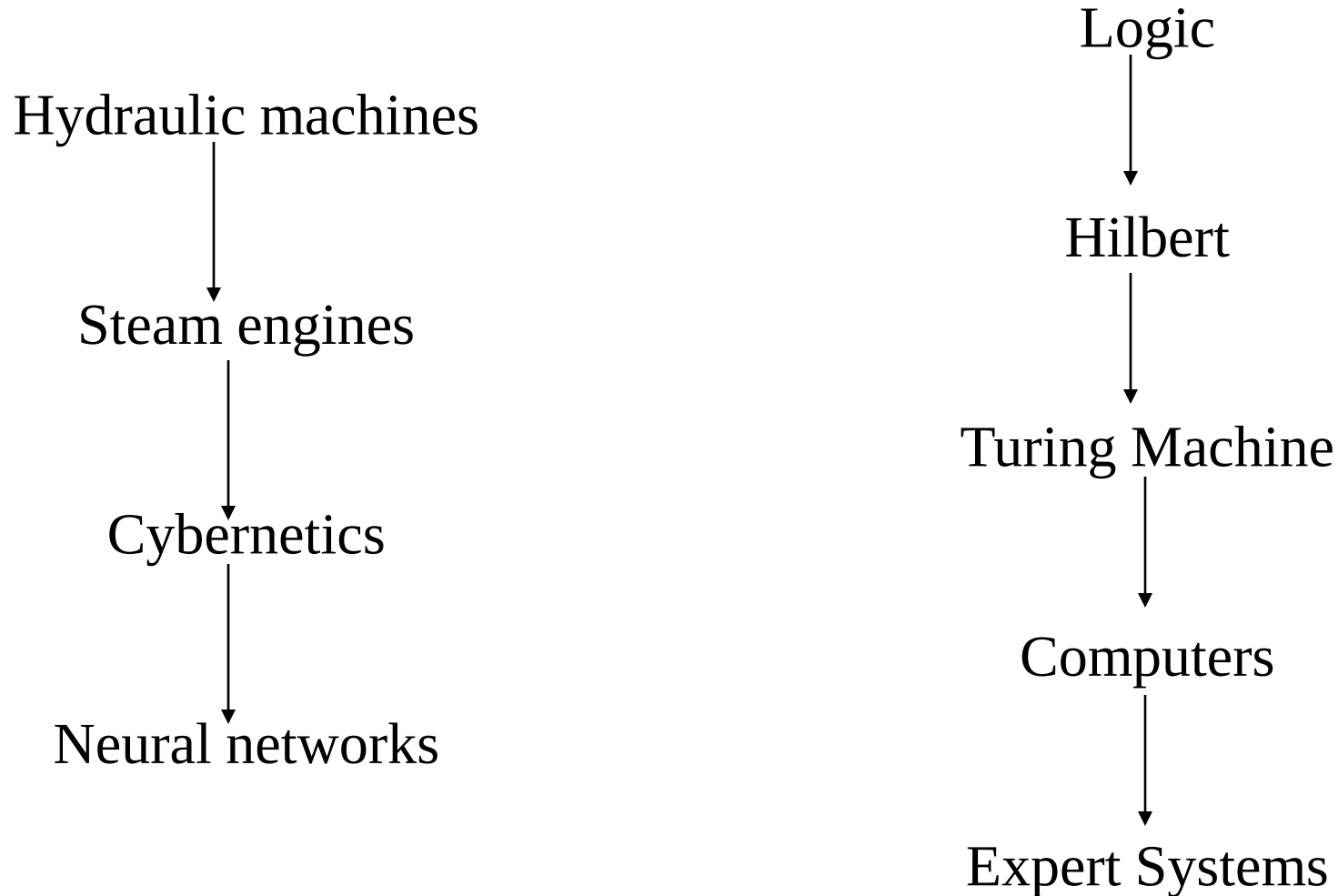
Multidirectional data flow
Total integration between input and output data
All neurons are linked between themselves
Trained with or without supervision



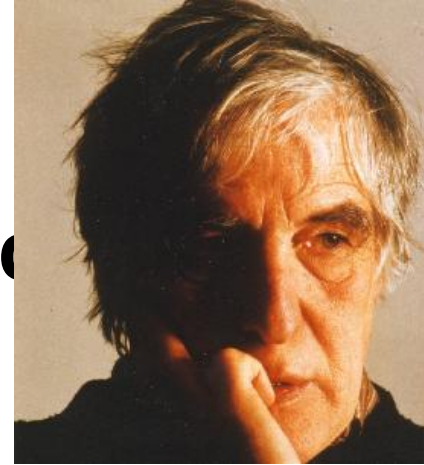
HOPFIELD

History of Artificial Intelligence

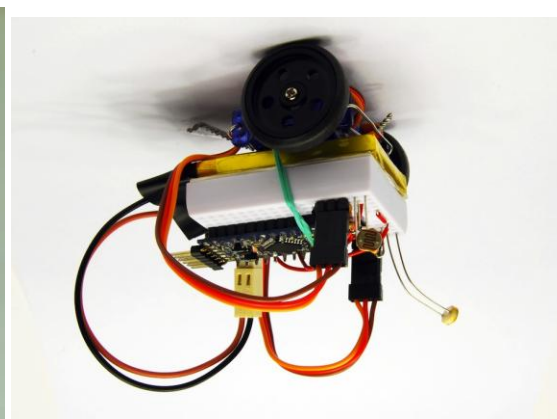
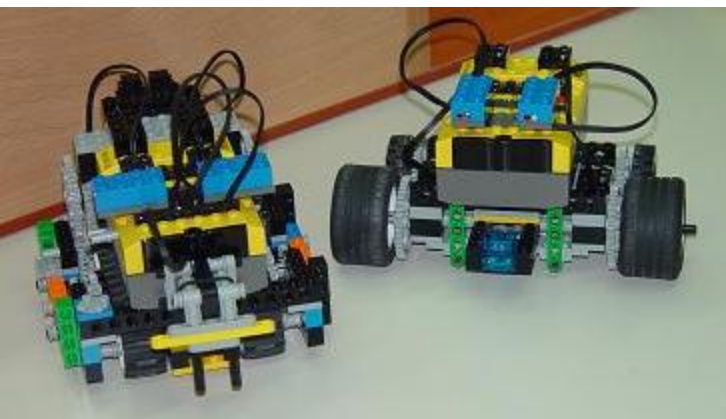
Genealogy of Intelligent Machines



History of Artificial Intelligence



- Valentino Breitenberg's "vehicles" (1984)
 - Vehicle 1: a motor and a sensor
 - Vehicle 2: two motors and two sensors
 - Increase little by little the circuitry, and these vehicles seem to acquire not only new skills, but also a personality.



History of Artificial Intelligence

1985: Judea Pearl's "Bayesian Networks"

1987: Chris Langton coins the term "Artificial Life"

1987: Rodney Brooks' robots

1990: Carver Mead describes a neuromorphic processor

1992: Thomas Ray develops "Tierra", a virtual world



History of Artificial Intelligence

1997: IBM's "Deep Blue" chess machine beats the world's chess champion, Garry Kasparov



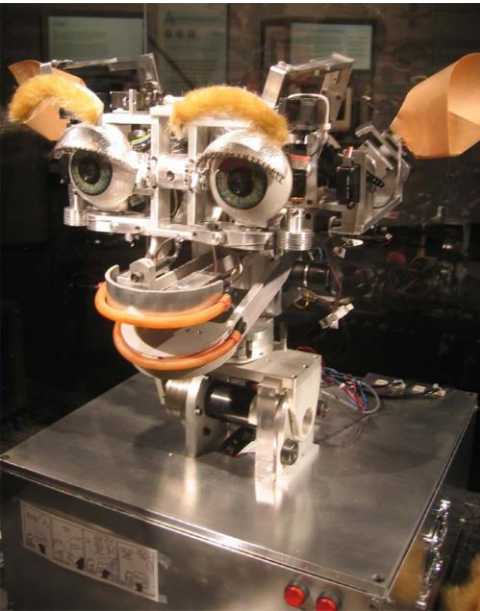
2011: IBM's Watson debuts on a tv show



History of Artificial Intelligence

2000: Cynthia Breazeal's emotional robot, "Kismet"

2003: Hiroshi Ishiguro's Actroid, a young woman



History of Artificial Intelligence

2004: Mark Tilden's biomorphic robot Robosapien

2005: Honda's humanoid robot "Asimo"



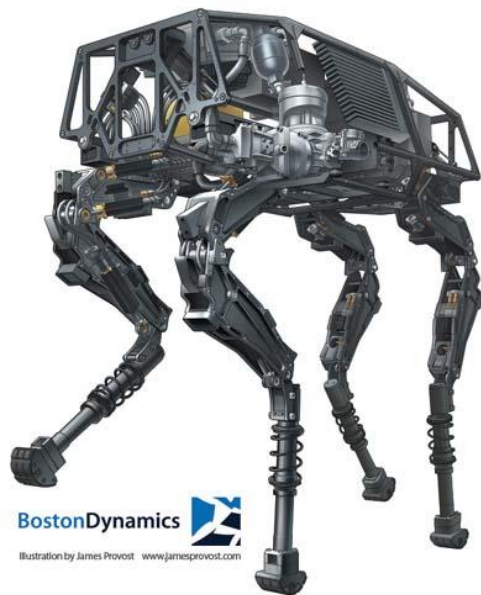
History of Artificial Intelligence

2005: Boston Dynamics' quadruped robot "BigDog"

2010: Lola Canamero's Nao, a robot that can show its emotions

2011: Osamu Hasegawa's SOINN-based robot that learns functions it was not programmed to do

2012: Rodney Brooks' hand programmable robot "Baxter"



History of Artificial Intelligence

- Rod Brooks/ Rethink Robotics (2012)
 - Vision to locate and grasp objects
 - Can be taught to perform new tasks by moving its arms in the desired sequence



History of Artificial Intelligence

Deep Learning

1998: Yann LeCun 's second generation Convolutional Neural Networks

2006: Geoffrey Hinton's Deep Belief Networks

2007: Yeshua Bengio's Stacked Auto-Encoders



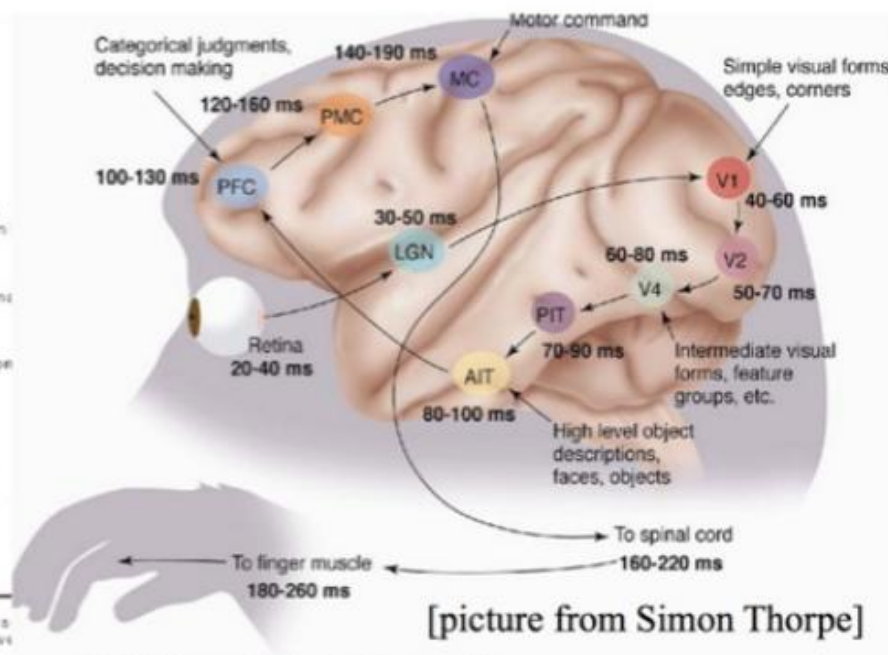
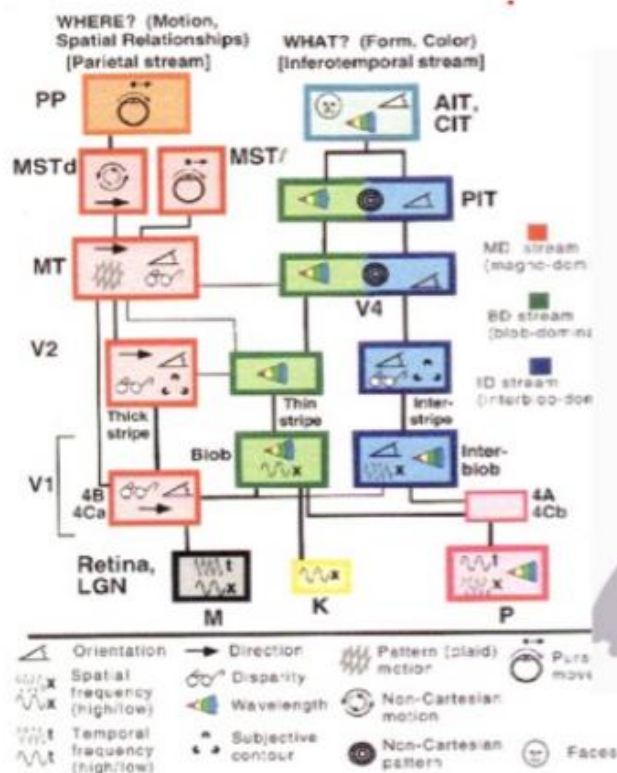
Deep Learning →

Error rates on the ILSVRC-2012 competition	
• Krizhevsky et. al.	• 16.4%
much bigger gap	
• University of Tokyo	• 26.1%
• Oxford University Vision Group	• 26.9%
• INRIA + XRCE	• 27.0%
• University of Amsterdam	• 29.5%



History of Artificial Intelligence

Deep Learning mimics the workings of the brain: the audiovisual cortex works in multiple hierarchical stages



[Gallant & Van Essen]

[picture from Simon Thorpe]

History of Artificial Intelligence

2014: Vladimir Veselov's and Eugene Demchenko's program Eugene Goostman, which simulates a 13-year-old Ukrainian boy, passes the Turing test at the Royal Society in London

2014: Li Fei-Fei's computer vision algorithm that can describe photos ("Deep Visual-Semantic Alignments for Generating Image Descriptions", 2014)

2014: Alex Graves, Greg Wayne and Ivo Danihelka publish a paper on "Neural Turing Machines"

2014: Jason Weston, Sumit Chopra and Antoine Bordes publish a paper on "Memory Networks"

2014: Microsoft demonstrates a real-time spoken language translation system

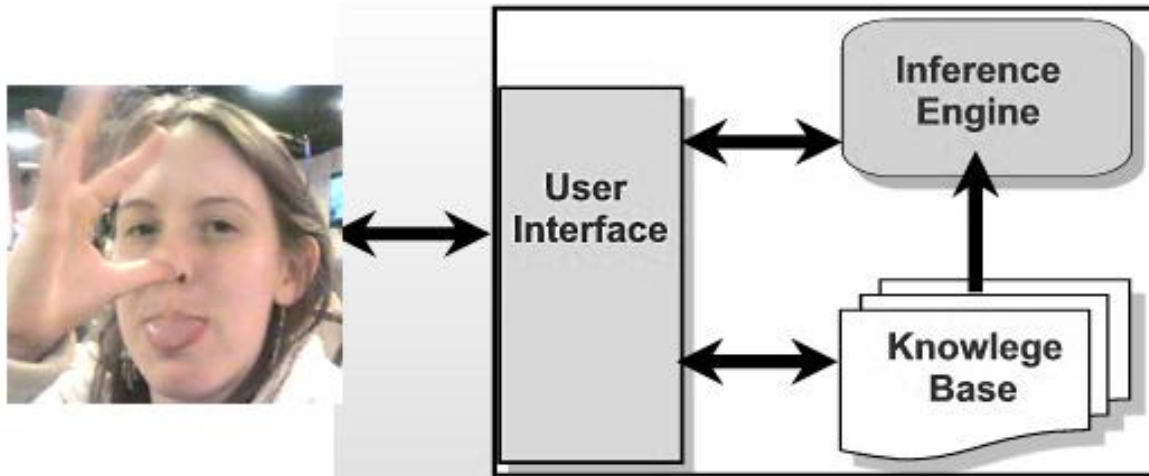
2014: *GoogLeNet*, a 22 layers convolutional network, wins Imagenet 2014, almost doubling the mean average precision of the previous year's winner

Break

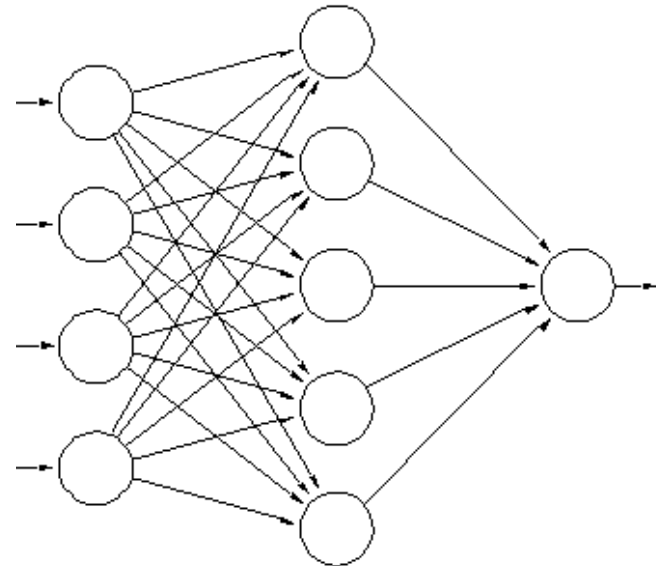


*"There is more stupidity than hydrogen in the universe,
and it has a longer shelf life"*
(Frank Zappa)

Artificial Intelligence



VS



Knowledge-based Systems

“General problem solver”:
the program capable of
solving all problems



Intelligence =
reasoning about
knowledge



Domain knowledge
and domain experts



Knowledge
Representation



Knowledge-based
systems
(expert systems)

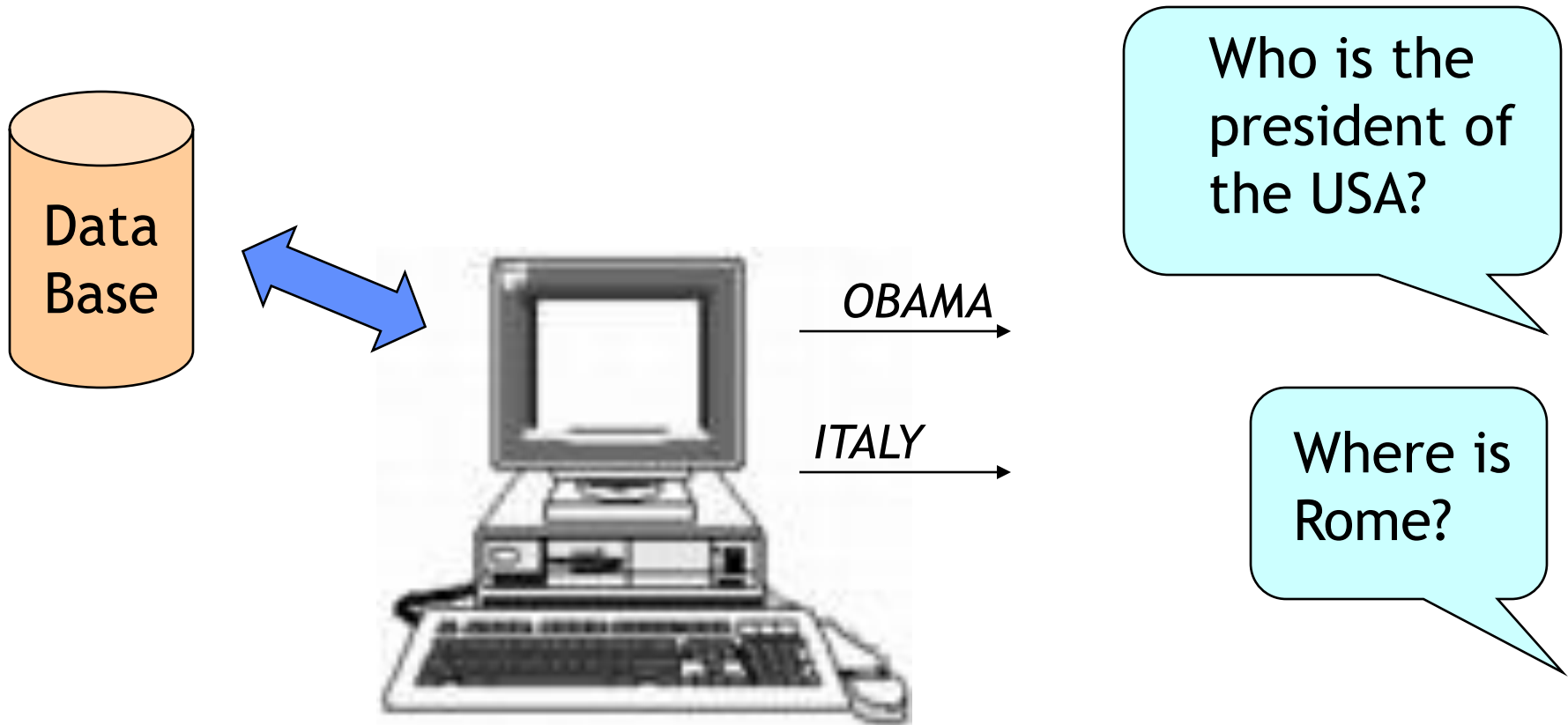


Knowledge
Engineering

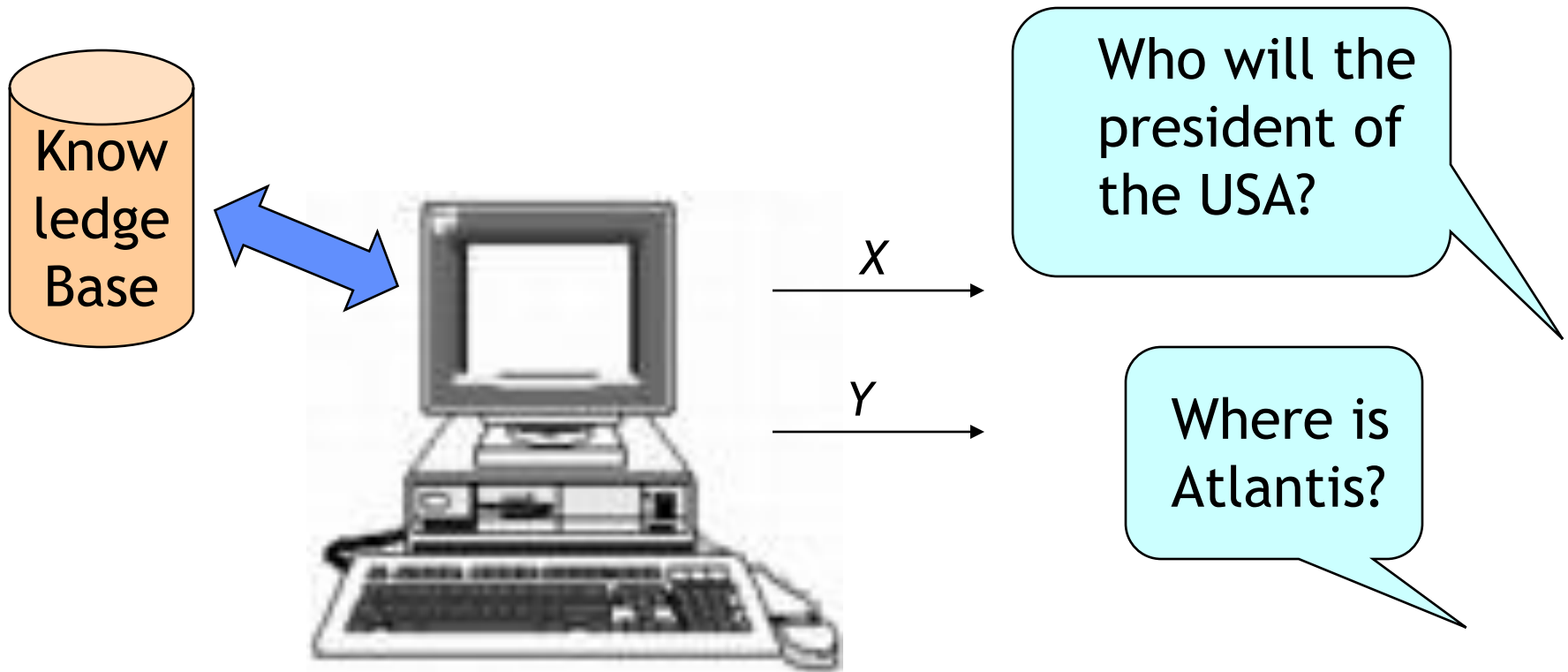
Knowledge-based Systems

- Knowledge representation
 - Predicates
 - Production rules
 - Semantic networks
 - Frames
- Inference engine: reasoning mechanisms
- Common sense & heuristics
- Uncertainty
- Learning

Information-based System



Knowledge-based System



Information Processing VS Knowledge Processing

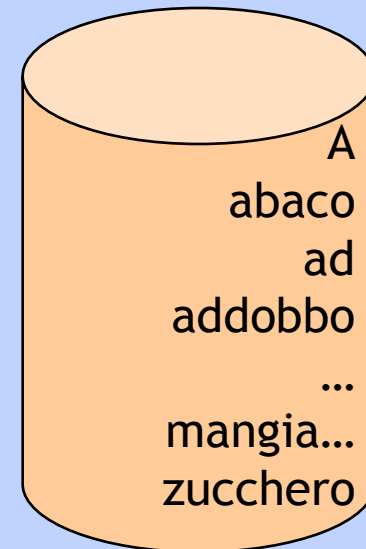
Puzzle: What is the Italian word “PNAATEI”?

1)

AAEINPT
AAEINTP
AAEITPN
...

$7! = 5000$

PIANETA



No
know-
ledge
of Italian

2)

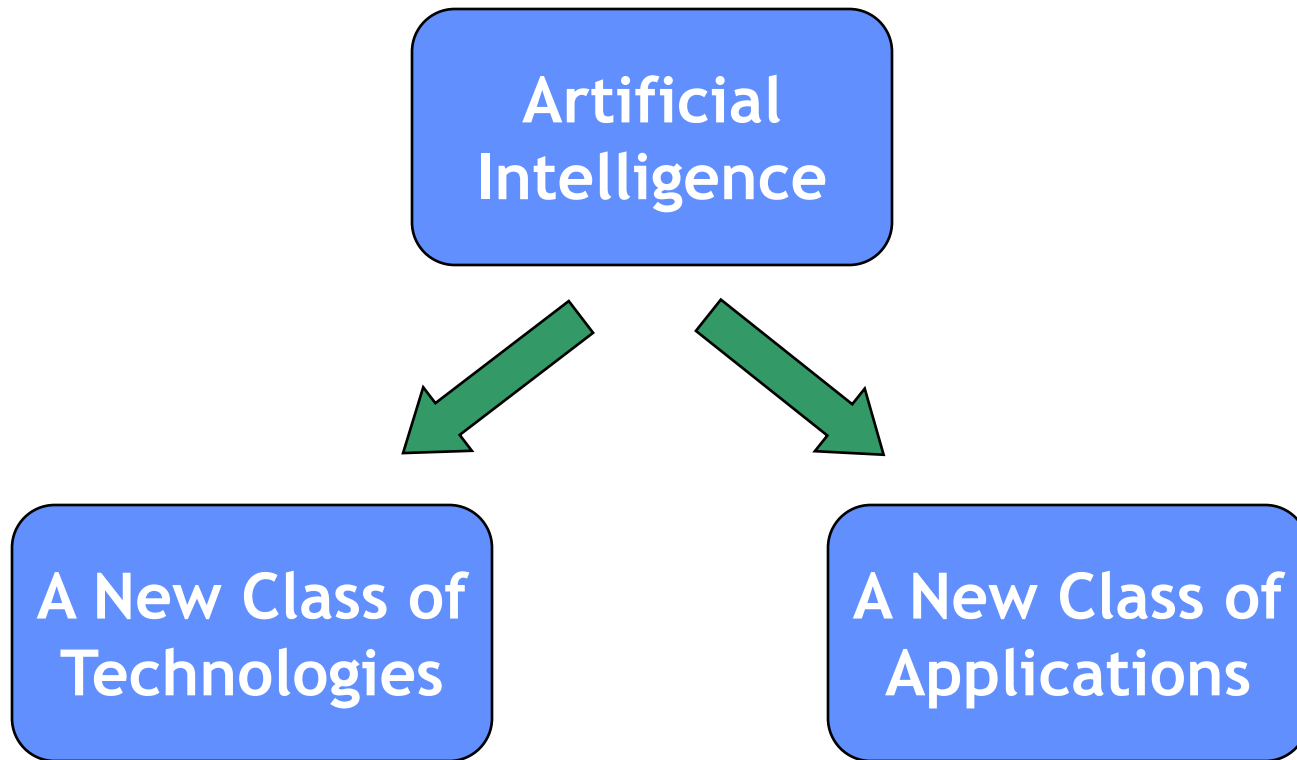
PANIETA
NAPIETA
TANIEPA
PENIATA
...

Less than 100

PIANETA



Artificial Intelligence



Artificial Intelligence

A New Class of Applications

Expert Tasks

The algorithm does not exist
A medical encyclopedia is not equivalent to a physician

Heuristics

There is an algorithm but it is “useless”
Don’t touch boiling water

Uncertainty

The algorithm is not possible
Italy will win the next world cup

“Complex” Problem Solving

The algorithm is too complicated
Design a cruise ship

Artificial Intelligence

A New Class of
Applications

Expert
Tasks

Heuristics

Uncertainty

“Complex”
Problem Solving

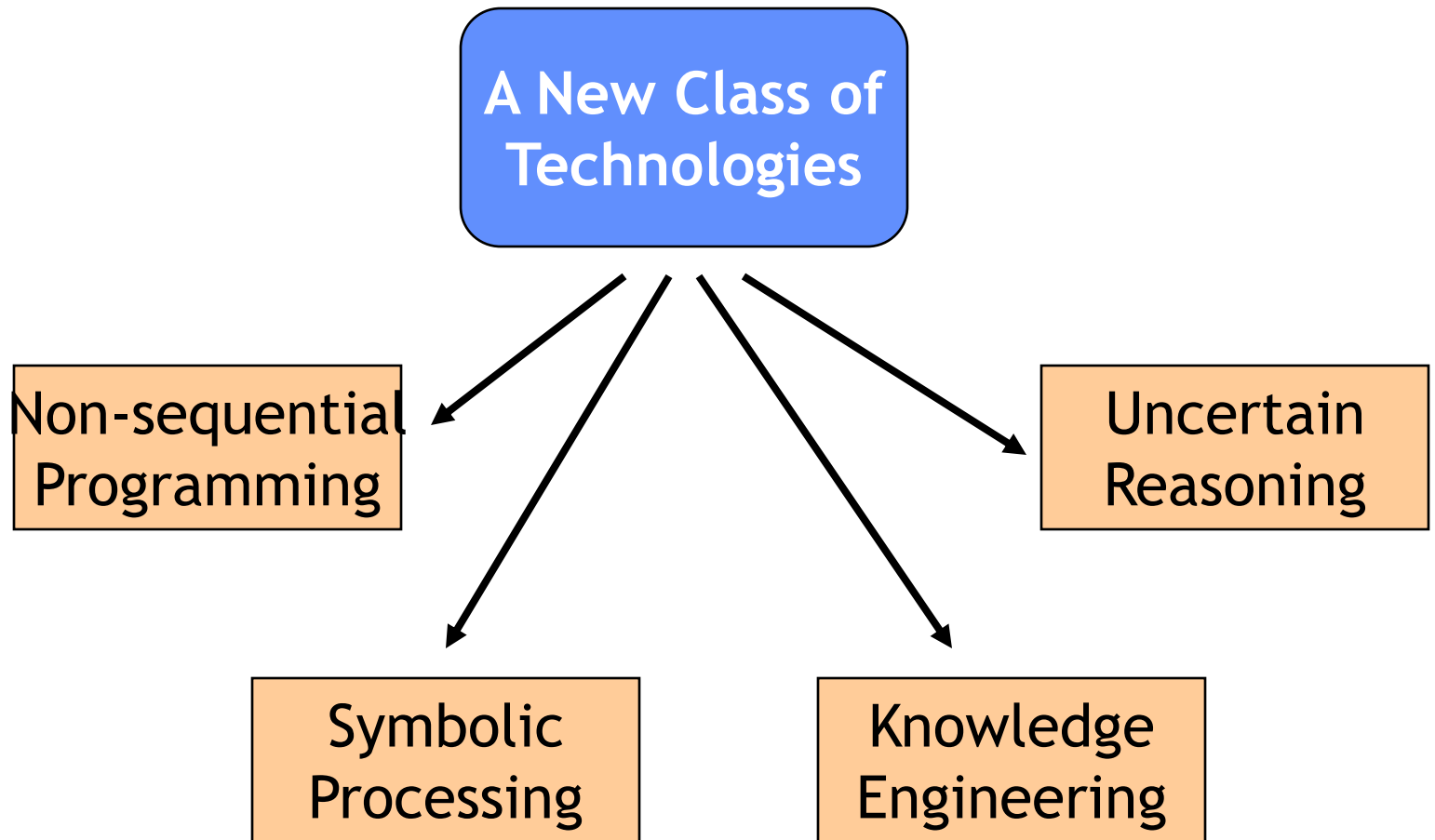
Expert
Systems

Natural
Language

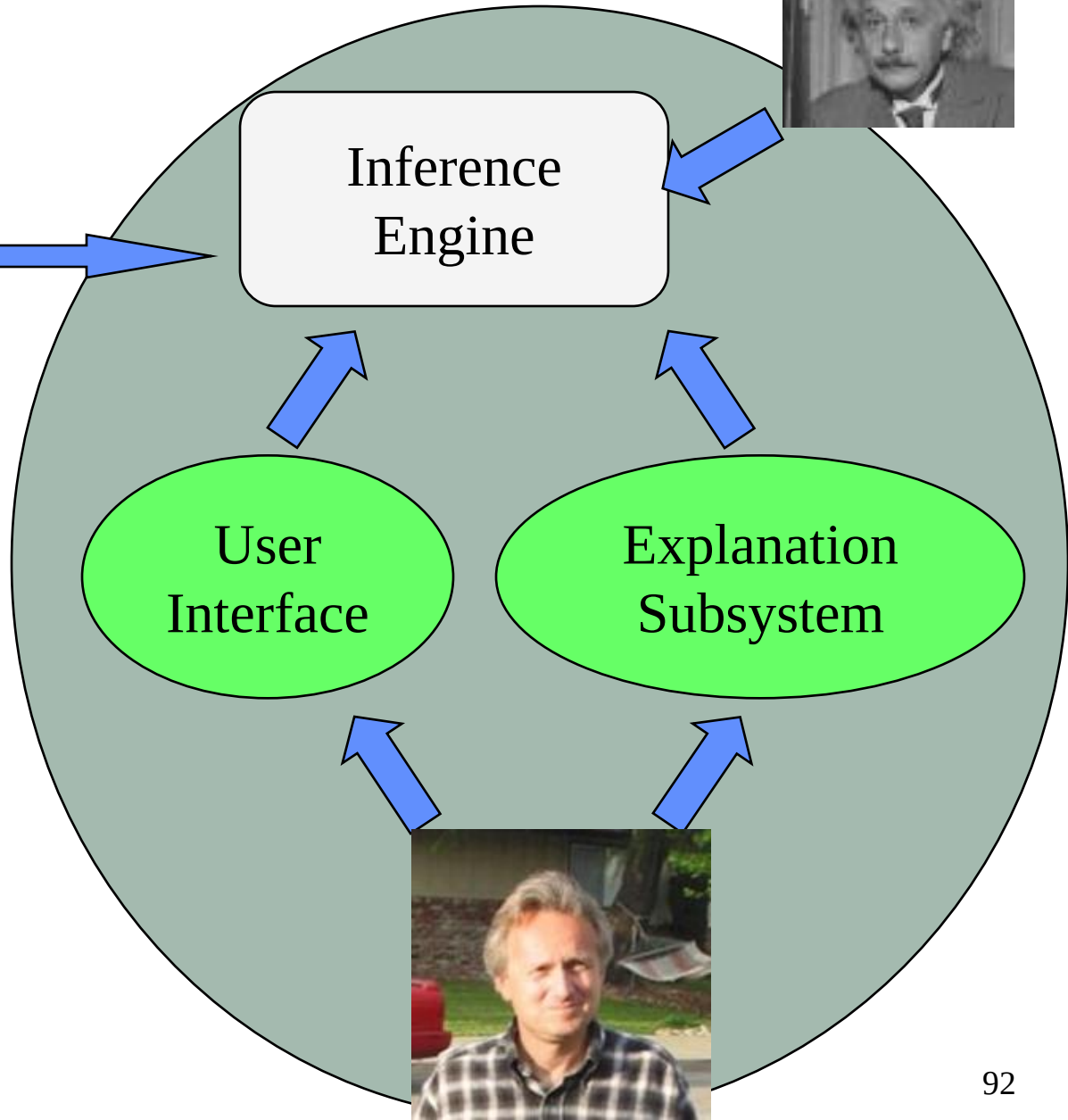
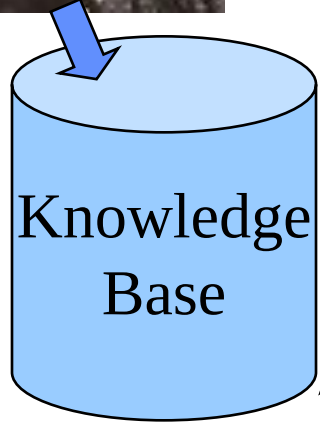
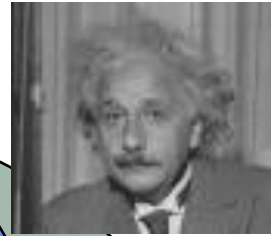
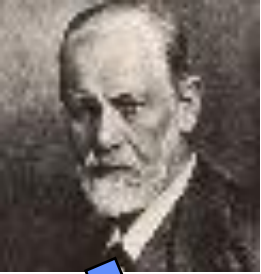
Vision
Speech



Artificial Intelligence



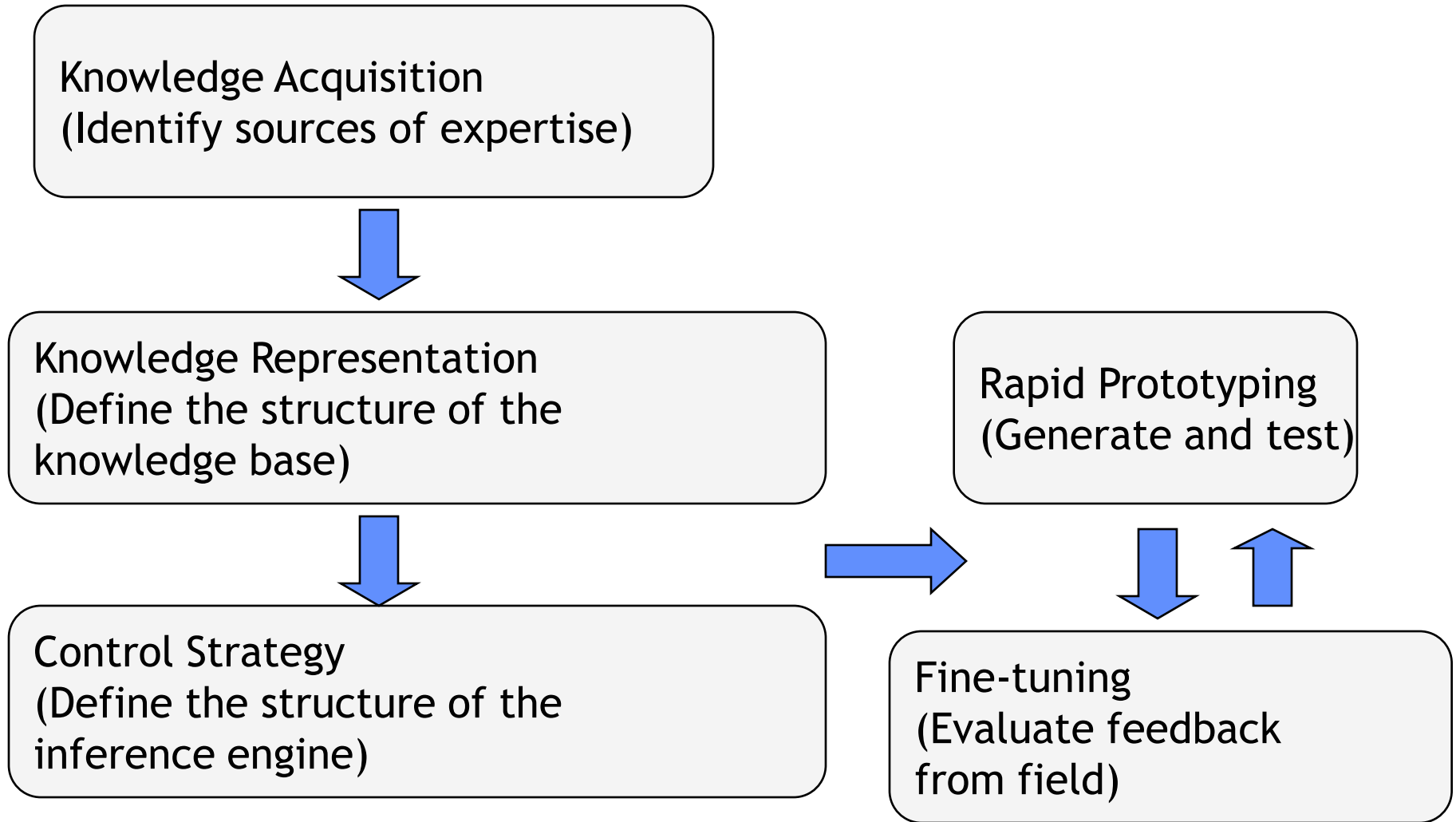
Expert system



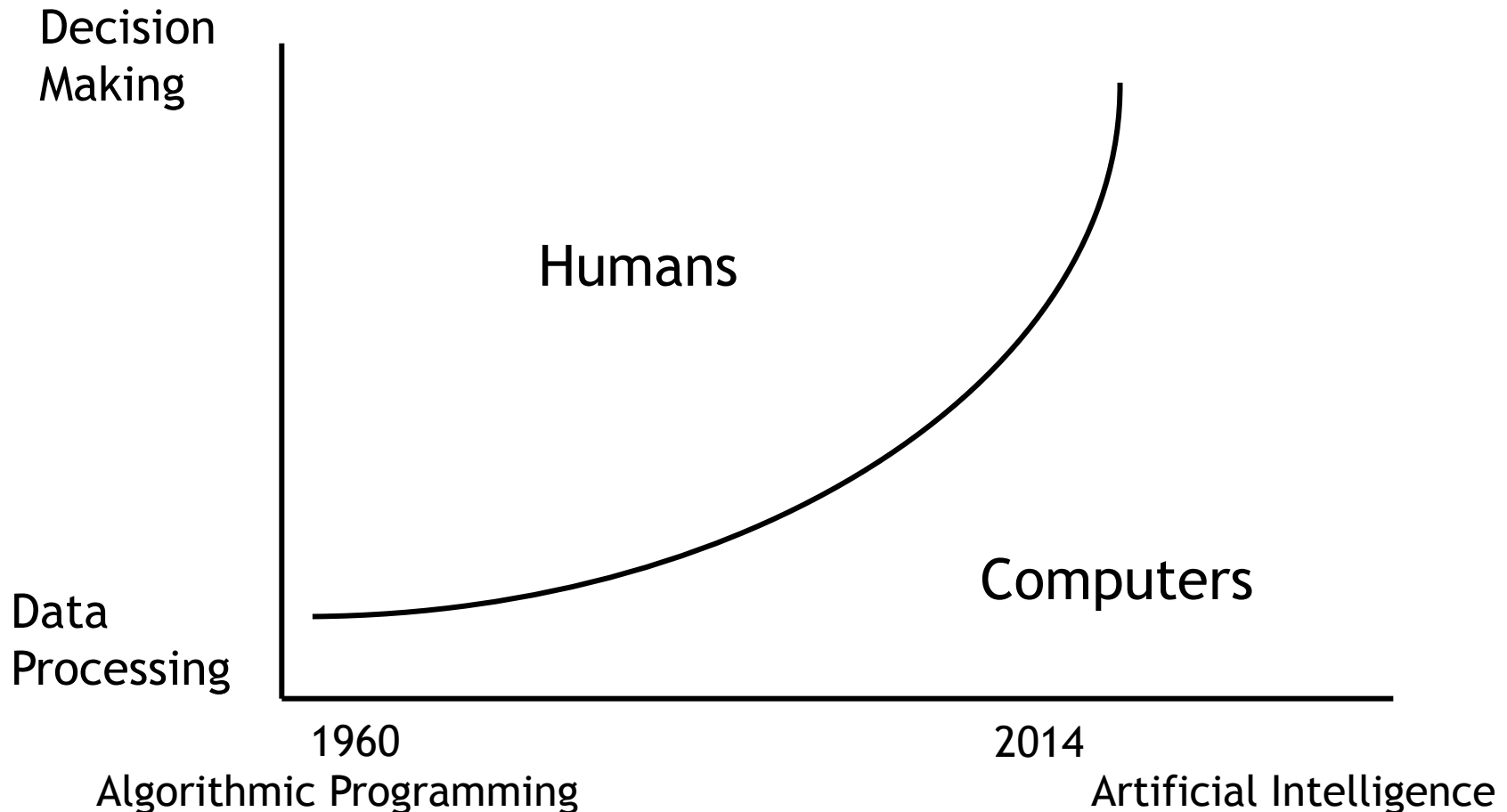
Expert Systems

- **Protagonists**
 - A.I. Scientist
 - Domain Expert
 - Knowledge Engineer
 - End-user
- **provide**
 - A.I. Technology
 - Knowledge
 - Knowledge representation
 - Feedback

Expert System Manufacturing Cycle



The Evolution of Computers in the Information Age



2000s

- Computational power becomes a distraction
 - Translation
 - Search
 - Voice recognitionbased on statistical analysis, not “intelligence”
- Emphasis on guided machine learning, in most cases probabilistic analysis of cases
- “Best Guess AI”

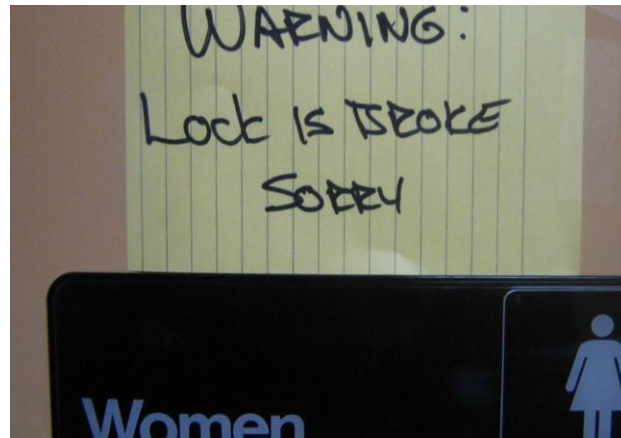
Common Sense



*Small minds are concerned with the extraordinary,
great minds with the ordinary"*
(Blaise Pascal)

Common Sense

- In which cases it is perfectly logical to find a woman in the men's restrooms?



TOILET
OUT OF
ORDER
PLEASE
USE
FLOOR
BELOW



Toilet
Out of Order
Please use Sink



Common Sense

- How easy is it to build a robot that can make these trivial inferences?

Common Sense

- What is wrong about this picture?



Communication tower
on top of Mt Diablo (2015)

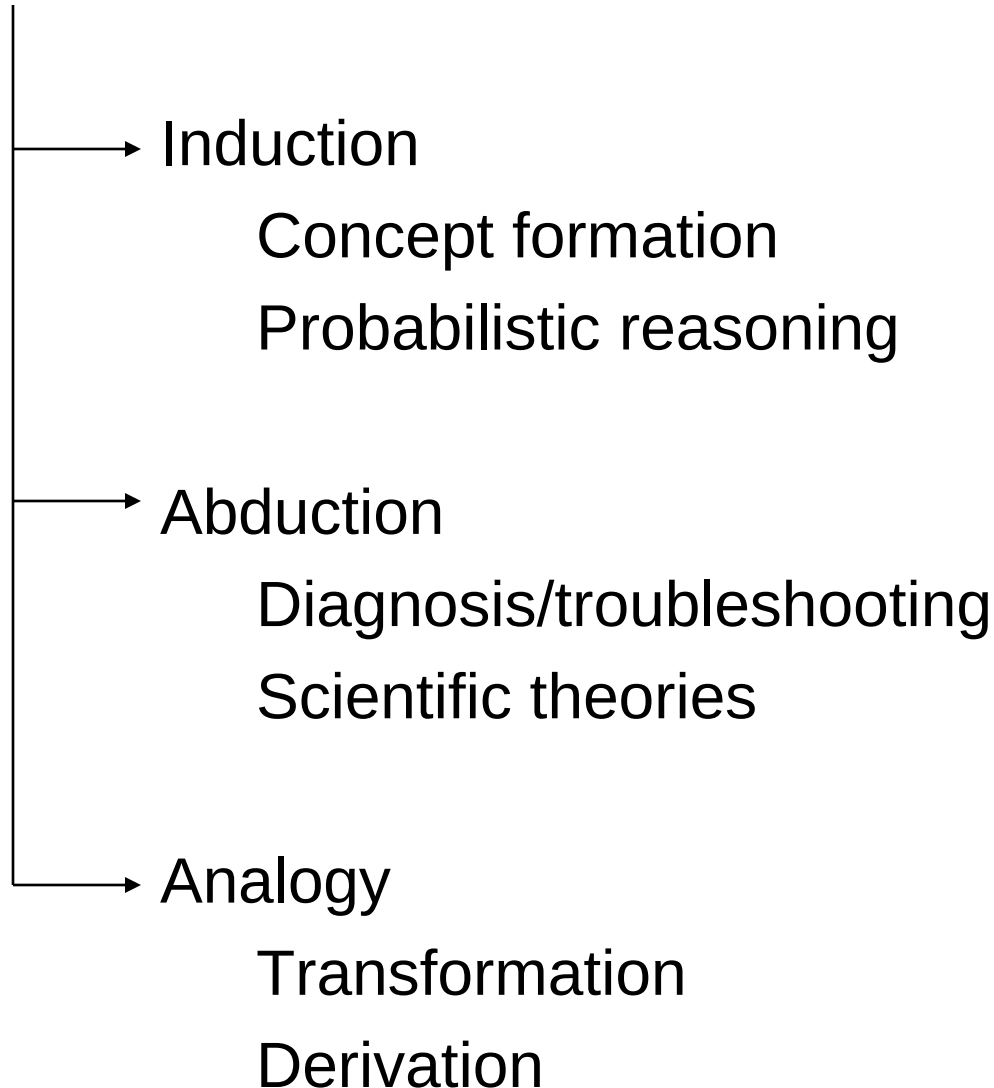


Common Sense

- Deduction is a method of exact inference (classical logic)
 - All Greeks are humans and Socrates is a Greek, therefore Socrates is a human
- Induction infers generalizations from a set of events (science)
 - Water boils at 100 degrees
- Abduction infers plausible causes of an effect (medicine)
 - You have the symptoms of a flue

Common Sense

Types of inference



Common Sense

- Classical Logic is inadequate for ordinary life
- Intuitionism (Luitzen Brouwer, 1925)
 - “My name is Piero Scaruffi or $1=2$ ”
 - “Every unicorn is an eagle”
 - Only “constructable” objects are legitimate
- Frederick and Barbara Hayes-Roth (1985): opportunistic reasoning
 - Reasoning is a cooperative process carried out by a community of agents, each specialized in processing a type of knowledge

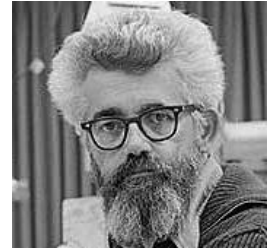


Common Sense

- Multi-valued Logic (Jan Lukasiewicz, 1920)
 - Trinary logic: adds “possible” to “true” and “false”
 - Or any number of truth values
 - A logic with more than “true” and “false” is not as “exact” as classical Logic, but it has a higher expressive power
- Plausible reasoning
 - Quick, efficient response to problems when an exact solution is not necessary
- Non- monotonic Logic
 - Second thoughts: inferences are made provisionally and can be withdrawn at any time



Common Sense



The Frame Problem

- Classical logic deducts all that is possible from all that is available
- In the real world the amount of information that is available is infinite
- It is not possible to represent what does “not” change in the universe as a result of an action (“ramification problem”)
- Infinite things change, because one can go into greater and greater detail of description
- The number of preconditions to the execution of any action is also infinite, as the number of things that can go wrong is infinite (“qualification problem”)

Common Sense

Uncertainty

“Maybe i will go shopping”

“I almost won the game”

“This cherry is red”

“piero is an idiot”

Probability

Probability measures "how often" an event occurs

But we interpret probability as “belief”

Glenn Shafer’s and Stuart Dempster’s “Theory of Evidence” (1968)

Common sense



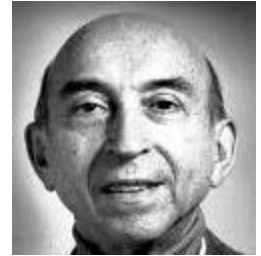
Principle of Incompatibility (Pierre Duhem)

The certainty that a proposition is true decreases with any increase of its precision

The power of a vague assertion rests in its being vague (“I am not tall”)

A very precise assertion is almost never certain (“I am 1.71cm tall”)

Common Sense



Fuzzy Logic

Not just zero and one, true and false

Things can belong to more than one category, and
they can even belong to opposite categories, and
that they can belong to a category only partially

The degree of “membership” can assume any value
between zero and one

Common Sense



The world of objects

Pat Hayes' measure space (1978)

- measure space for people's height: the set of natural numbers from 100 (cms) to 200 (cms)
- measure space for driving speed: the set of numbers from 0 to 160 (km/h)
- measure space for a shirt's size: small, medium, large, very large

John McCarthy's "Situation Calculus" (1963)

Qualitative reasoning (Benjamin Kuipers; Johan DeKleer; Kenneth Forbus)

- Qualitative descriptions capture the essential aspects of structure, function and behavior, e.g. "landmark" values

Common Sense

Heuristics

- Knowledge that humans tend to share in a natural way: rain is wet, lions are dangerous, most politicians are crooks, carpets get stained...
- Rules of thumbs

György Polya (1940s): “Heuretics” - the nature, power and behavior of heuristics: where it comes from, how it becomes so convincing, how it changes



Common Sense



Douglas Lenat (1990):

- A global ontology of common knowledge and a set of first principles (or reasoning methods) to work with it
- Units of knowledge for common sense are units of "reality by consensus"
- Principle of economy of communications: minimize the acts of communication and maximize the information that is transmitted.

Connectionism

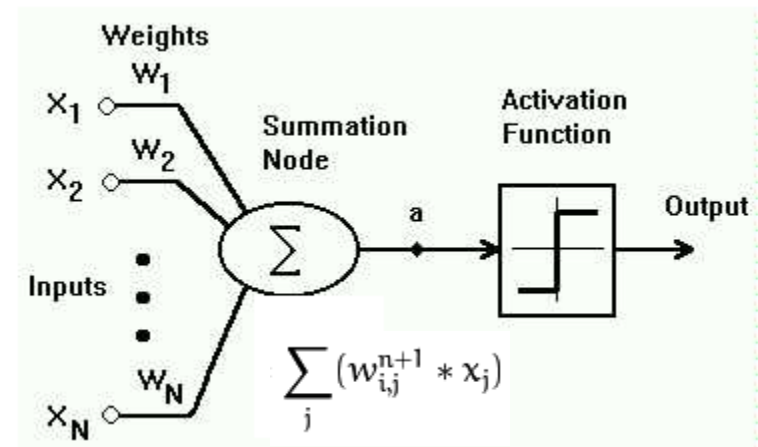
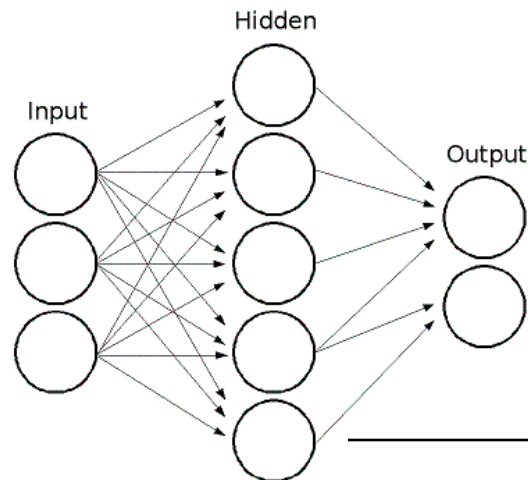
"Politics is the art of looking for trouble, finding it everywhere, diagnosing it incorrectly and applying the wrong remedies"
(Groucho Marx)

Connectionism

A neural network is a set of interconnected neurons (simple processing units)

Each neuron receives signals from other neurons and sends an output to other neurons

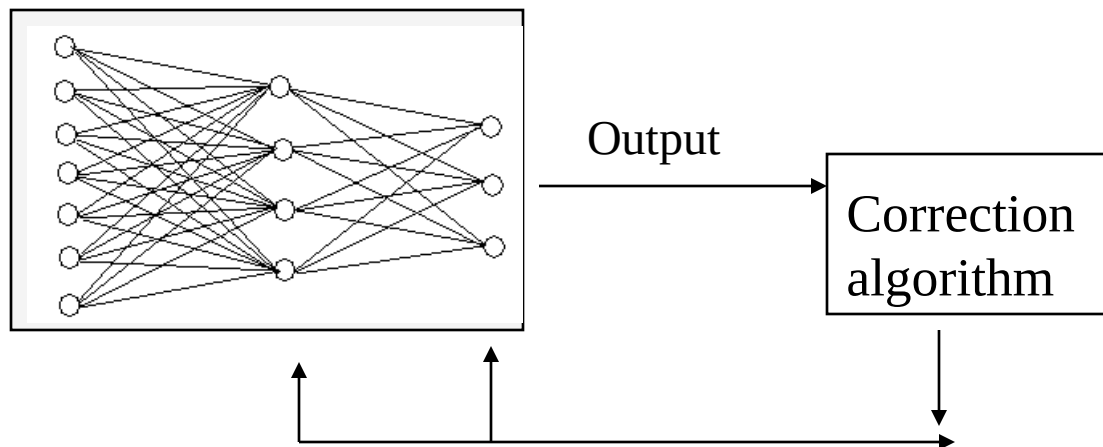
The signals are “amplified” by the “strength” of the connection



Connectionism

The strength of the connection changes over time according to a feedback mechanism (output desired minus actual output)

The net can be “trained”



Connectionism

- Distributed memory
- Nonsequential programming
- Fault-tolerance
- Recognition
- Learning

Connectionism

Where are we?

Largest neural network:

- 1,000,000,000,000 neurons

Worm's brain:

- 1,000 neurons

But the worm's brain outperforms neural networks on most tasks

Human brain:

- 100,000,000,000 neurons
- 200,000,000,000,000 connections

More Than Intelligence

- Summary
 - Common Sense
 - Deduction/ Induction/ Abduction
 - Plausible Reasoning
 - The Frame Problem
 - Uncertainty, Probability, Fuzzy Logic
 - Neural Networks
 - Fault-tolerant
 - Recognition tasks
 - Learning

Artificial Intelligence

- Notes...
 - How many people can fly and land upside down on a ceiling? Don't underestimate the brain of a fly.
 - Computers don't grow up. Humans do.

Artificial Life

- 1947: John Von Neumann's self-replicating and evolving systems
- 1962: first computer viruses
- 1970: John Conway: The Game of Life
- 1975: John Holland's genetic algorithms
- 1976: Richard Dawkins: the meme
- 1987: Chris Langton: "Artificial Life"
- 1988: Robert Morris: the first computer virus on the Arpanet
- 1991: Thomas Ray: "Evolution and optimization of digital organisms"

Emergent Computation

- Alan Turing's reaction-diffusion theory of pattern formation
- Von Neumann's cellular automata
 - Self-reproducing patterns in a simplified two-dimensional world
 - A Turing-type machine that can reproduce itself could be simulated by using a 29-state cell component



1. Machine P (parent) reads the blueprint and makes a copy of itself, machine C (child).
 2. Machine P now puts its blueprint in the photocopier, making a copy of the blueprint.
 3. Machine P hands the copy of the blueprint to machine C.
- The blueprint is both **active instructions** and **passive data**.

Emergent Computation

- Turing proved that there exists a universal computing machine
- Von Neumann proved that
 - There exists a universal computing machine which, given a description of an automaton, will construct a copy of it
 - There exists a universal computing machine which, given a description of a universal computing machine, will construct a copy of it
 - There exists a universal computing machine which, given a description of itself, will construct a copy of itself

Artificial Life

- John Holland's genetic algorithms (1975)
- Genetic algorithms as "search algorithms based on the mechanics of natural selection and natural genetics"
 - “Reproduction” (copies chromosomes according to a fitness function)
 - “Crossover” (that switches segments of two chromosomes)
 - “Mutation”
 - Etc
- Thomas Ray’s “Tierra” (1992)



Artificial Life

- Possible solutions "evolve" in that domain until they fit the problem.
- Solutions evolve in populations according to a set of "genetic" algorithms that mimic biological evolution.
- Each generation of solutions, as obtained by applying those algorithms to the previous generation, is better "adapted" to the problem at hand.

Artificial Life

- Frank Tipler: one has no way to tell a computer simulation of the real world from the real world, as long as one is inside the simulation
- The distinction between reality and simulation is fictitious
- Artificial life replaces the "problem solver" of artificial intelligence with an evolving population of problem solvers
- The "intelligence" required to solve a problem is not in an individual, it is in an entire population and its successive generations

Virtual Reality

- Morton Heilig: the Sensorama (1962)
- Rainer Werner Fassbinder: "Welt am Draht/ World on a Wire" movie (1973)
- William Gibson: the cyberspace (1982)
- John Badham: "War Games" movie (1983)
- Bruce Bethke: "Cyberpunk" (1983)
- NASA Ames: Virtual Planetary Exploration Workstation (1984)
- Lucasfilm: online game "Habitat" (1985) that introduces "avatars"
- Neal Stephenson: the metaverse (1992)
- Philip Rosedale: Second Life (2003)

Virtual Reality

- David Deutsch (1997)
 - The technology of virtual reality (the ability of a computer to simulate a world) is the very technology of life
 - Genes embody knowledge about their ecological niche
 - An organism is a virtual-reality rendering of the genes

Summary of the summaries

- David Hilbert (1928)
- The Turing Machine (1936)
- Cybernetics (1947)
- Information Theory (1949)
- Artificial Intelligence: Symbol Processing
- Artificial Intelligence: Connectionism
- Artificial Intelligence: Robots
- Common Sense
- Plausible Reasoning
- Recognition tasks
- Artificial Life

Summary of the summaries

- Which one is the robot?



Bibliography

- Dreyfus, Hubert & Dreyfus Stuart: Mind Over Machine (Free Press, 1985)
- Hofstadter, Douglas: Goedel Escher Bach (Vintage, 1980)
- Holland, John: Emergence (Basic, 1998)
- Kosko Bart: Neural Networks And Fuzzy Systems (Prentice Hall, 1992)
- Langton, Christopher: Artificial Life (Mit Press, 1995)
- Rumelhart David & Mcclelland James: Parallel Distributed Processing Vol. 1 (Mit Press, 1986)
- Russell, Stuart Jonathan & Norvig Peter: Artificial Intelligence (Prentice Hall, 1995)
- Wiener, Norbert: Cybernetics (John Wiley, 1948)

Machine Intelligence

"The person who says it cannot be done should not interrupt the person doing it" (Chinese proverb)

Piero Scaruffi
www.scaruffi.com