

Some Foundations / (Very Brief) History of A.I.

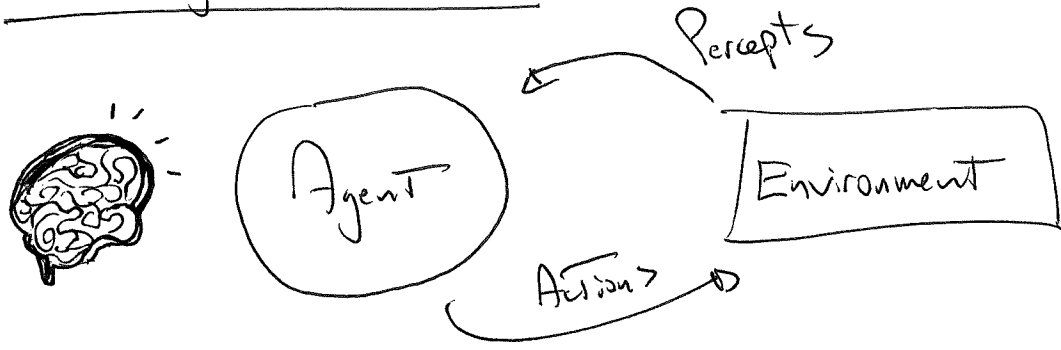
"Numberless are the world's wonders - but none more wonderful than mankind" - Sophocles, Antigone

What is AI

Behavior (Acting "humanly")

Rationality (Acting & Thinking "rationally")

Rational Agent Model for AI



A Rational Agent "behaves as well as possible."

"Strong AI"

Machines are actually "Thinking"

vs. "Weak AI"

Machines act as if they were intelligent

(*) Most (but not all) practitioners believe weak AI is (sometimes trivially) achievable - but disagree about the feasibility of strong AI.

(*) Few mainstream researchers believe that anything significant hinges on the outcome of this debate.

(ELIZA: 1969 (MIT))

(2)

Turing Test (1950)

= Please read this paper
(linked on website)

(Total Turing Test)

(*) Again, few researchers believe Turing Test is crucial to the future development of AI - however many fascinating & canonical issues are raised by Turing in this paper on the nature of AI.

Q: What is intelligence - is it a well-defined concept?

(*) 1955 study found that in 19/20 "expert" disciplines, an elementary mathematical model (e.g. Regression, Naive Bayes) outperformed human practitioner!

Aspects of AI & AI systems:

NLP: Natural Language Processing

Knowledge Representation: Store/Retrieve what is known

Automated Reasoning: Use stored information for inference/deduction.

Machine Learning: Detect/extrapolate patterns.

Agent's Representation of the World: Environment-knowledge representation.

Computer Vision

Robotics

(sensors, actuators)

"Total Turing Test"

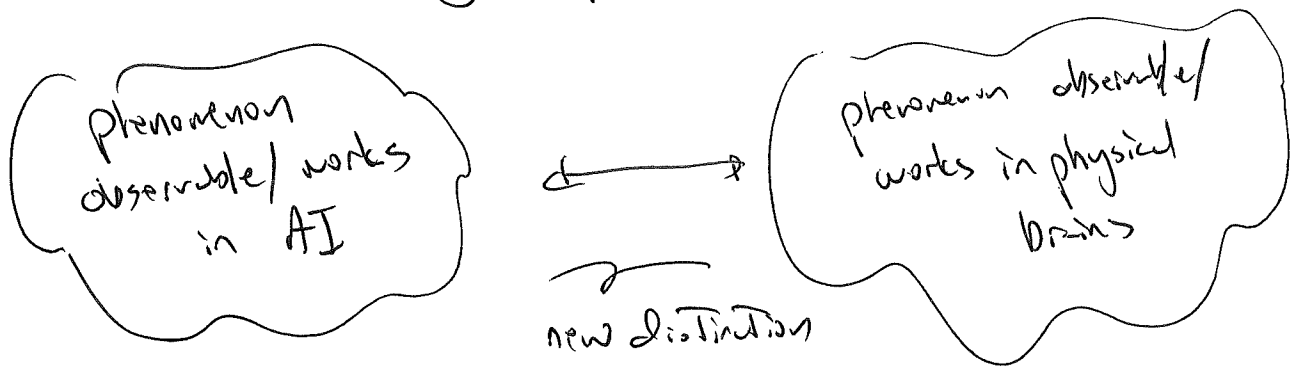
Related Disciplines:

Neuroscience: Direct study of nervous system (Brain function, etc)

Psychology: Emphasis on Behaviorism (Skinner) -
difficult to directly test.

Cognitive Science: "AI + Psych." (use Testable Theories)

(*) Most researchers today accept the basic distinction:



(*) Nevertheless, many important biological models have inspired, fruitfully, models in AI (e.g. computer vision).

Still though, rarely (to date) do researchers begin with the premise: "let's build a biological brain".

Why? Because we still don't know how a biological brain works! Consciousness, for example is still a mystery.

Q: Is computation enough? Not quite.

Human Brain:
10¹¹ neurons; 10¹⁴ synapses
~ 10¹ - 10⁵ connections/neuron

≠

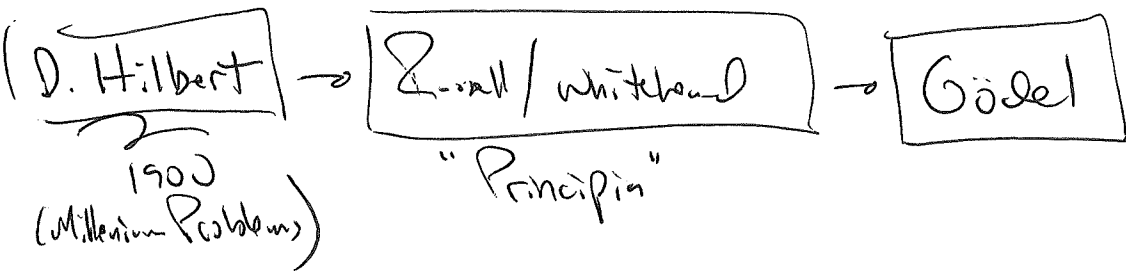
Super Computer:
10⁹ CPUs; 10¹² Transistors
10¹⁹ bits of RAM

Rational Thinking

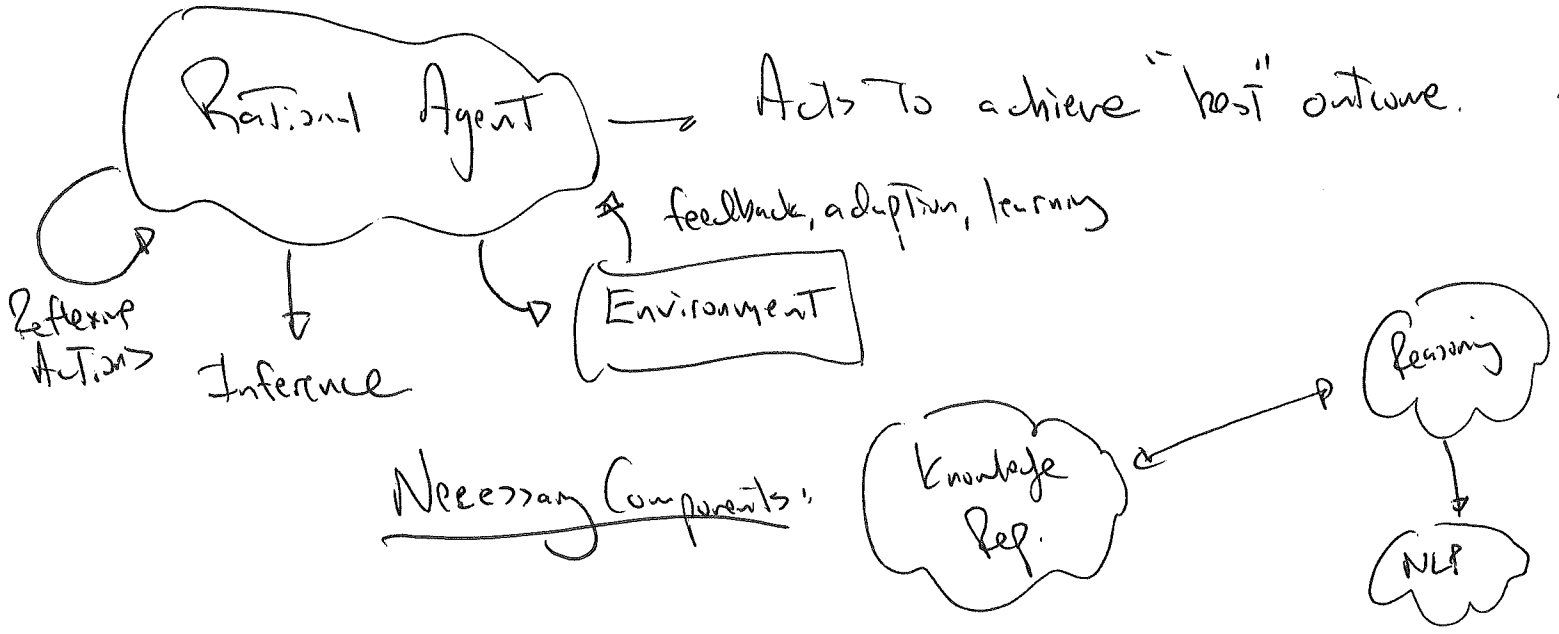
(*) Early codification of "right thinking" - Aristotle (syllogisms).

Logic: Is there a set (of ^{finite?} discoverable) laws that govern the operation of the mind?

Issues with Logician-AI methodology: Difficult to translate all real-world problems into symbolic form; tractability; some argue incompleteness (Gödel) renders Logician approach to Mathematics/AI futile.



A More Complete Rational Agent?



"Perfect Rationality" - a good starting point for AI.

A (very brief) History of AI

Myth of Perfection, Faust

What is the nature of knowledge, where does it come from?
(Epistemology) Plato - Meno Dialogue & a priori knowledge.

Induction vs. Deduction (Reasoning)

Q's (i) Can we simply learn a huge list of inference rules?

(i) Does a look-up Table constitute intelligence?

(ii) How to proceed from knowledge to action?

Hobbes: "Reasoning is like Numerical Computation" (1651) (Leviathan)

Leonardo: Designs for mechanical calculator (first conceived?) (ISC)
(IBM built a replica in 1968)

Leibniz: "Step Reckoner": could perform all (1694)
4 arithmetic ops.

Also conceived of computation machine operating on concepts!

B. Pascal "Pascaline" (1652): Arithmetic Mechanical Calculator.

Descartes: Cogito (Meditations, 1641)

Rationalists Dualism: Part of Mind/Spirit outside body & external world

Materialism: Brain constitutes the Mind.

Sources & Nature of Knowledge (Epistemology)

Empiricism: Bacon (16C), Locke, Hume
(Aristotle)

Inductive Reasoning: General rules acquired by repeated exposure/association.
Can we "prove" induction?

Paradoxes of Induction: Black Swan

Utilitarianism Mill: Ethics: Maximize/Quantity of utility.
19C

Phenomenology: Husserl → Heidegger: attempts to square subjective experiences w/ Rationalism.
20C

Logical Positivism Russell → Wittgenstein → Carnap
20C
(Verificationism): only meaningful problems are those solvable by logical analysis (against metaphysics).

(1820s)

C. Babbage

Difference Engine (healy: computes

polynomial coeff's from Newton's Divided Differences
(^{classical} interpolation)).

Analytical Engine:

(1837)

(AE)

Proposed general-purpose computer;
integrating loops, memory, logic unit (Turing-complete machine).

Ada Lovelace:

First programmer? Wrote programs for AE;
speculated about creative abilities of AI (class/music).

More Recent AI History

(Turing: Bombe - early 1940s)

Gestation 1943-1955

McCulloch & Pitts

(1943): Artificial Neurons

Perception Learning Algorithm

(Hebb 1949
Rosenblatt 1962)

→ Showed any computable function can be computed by some network of connected neurons.

M. Minsky & Edmonds

(1950, MIT)

First ANN

(SARC: 6000 vacuum tubes)

(8)

Van Neumann on Minsky's PhD committee: Is Neural Network research relevant? "If not now, some day."

Turing (1950): "Computing Machinery & Intelligence": introduces Turing test, GAs, ML, Reinforcement Learning.

1956 Dartmouth Workshop on AI

Win attendees:

John McCarthy: founded Stanford AI Lab, invented Lisp, Advice Taker (use logic).

Minsky, Shannon, Newell, Simon.

Claim to have solved mind/body problem!

LT: Logic Theorist: Proved many theorems in Principia (some were skater)
Paper was rejected. unintelligible.

Next 20 years: AI research dominated by people from workshop.

1952-1969: Early Enthusiasm / Great Expectations

GPS: (Newell & Simon): General Problem Solver: embodied "Thinking humanly" approach.

Geometry Theorem Prover (IBM 1959)

A. Samuel (1952) Checkers AI game (best programmer)

Minsky (MIT): supervised students in Microworlds (e.g. block world) 9
problems.

H. Simon: "There are now in the world machines that think,
(1957) that learn & that create."

1966-1973 Microworld problems turn out not to be
indicators of general intelligence.

Searle: Only "symbolic" manipulation.

Ex. During US-Soviet Space Race, US government
funded research in machine translation.

Famous mistranslation: "The spirit is willing, but the flesh is weak" \rightarrow Translated
"The vodka is good but the meat is rotten."

1966 All funding halted for this project!

Other reasons for AI pessimism: Intractability

Q: "How to scale up" AI models?
Microworlds \rightarrow exhaustive search.

"Combinatorial Explosion"
P vs. NP

(Before formal Theory of computation it was thought that
simply faster hardware/more memory was sufficient.)


Theorem Provers also failed for larger scale problems.

Minsky: Perceptrons : Highly Contradictory : showed
(1969)

That for 3-layered perceptron, cannot compute some logical predicates (using "local neurons").

* Note: This result does not apply in general to multilayer feed-forward NNs.

* Most AI funding dries up - focus switches to "symbolic" systems.

2 AI "Winters" follow: 1974-1980 (I) Is winter coming?
 1987-1997 (II)

(*) Some Industry successes w/ AI (1979-1980)

Alternative to general-purpose search ("weak methods")?

Idea: Encode Domain Knowledge

Expert Systems MYCIN → Recommended Arti.-Diatrics.
(1970s: Stanford) "Correctness vs. Symbolic Models" vs. NN Training feasible

1986-Present Return of ANNs
(1969) Backprop Alg. "Reinvented"; Big Data + Parallel computing

J Pearl (1988) Prob. Reasoning in Intelligent Systems

11

(*) led to new acceptance of probability & decision theory in AI
(Bayes Nets; efficient representation of uncertainty).
decision making

Some More Recent Benchmarks

Y. Lecun (1990) Handwritten Digit Recognition
(CNNs & Deep learning)

Deep Blue (1997) : IBM, beat Kasparov 3.5 TO 2.5.

Deep Learning (2012): AlexNet : CNN used to achieve
top 5 error rate of 15.4% (on ImageNet)
22,000 categories

Adversarial Learning (2014): Goodfellows.

Watson (2011): (NLP + efficient search)

Playing Atari w/ Deep Reinforcement Learning (2013): Ms. Pac-Man, etc.

Alpha GO (2015):
(Google DeepMind)

Final Considerations

Is human-level (AGI) possible? → (singularity)
"Artificial General Intelligence"

Is deep learning the answer? (Still a "microworld"? See Adversarial Nets)

Supervised vs. unsupervised learning: Turns out more/"better" data might trump effectiveness of an algorithm!

"knowledge bottleneck" automate learning process, "bootstrap" new patterns.

(Lovelace) Is AI creativity achievable? Check out Deep Dream! (among other applications)