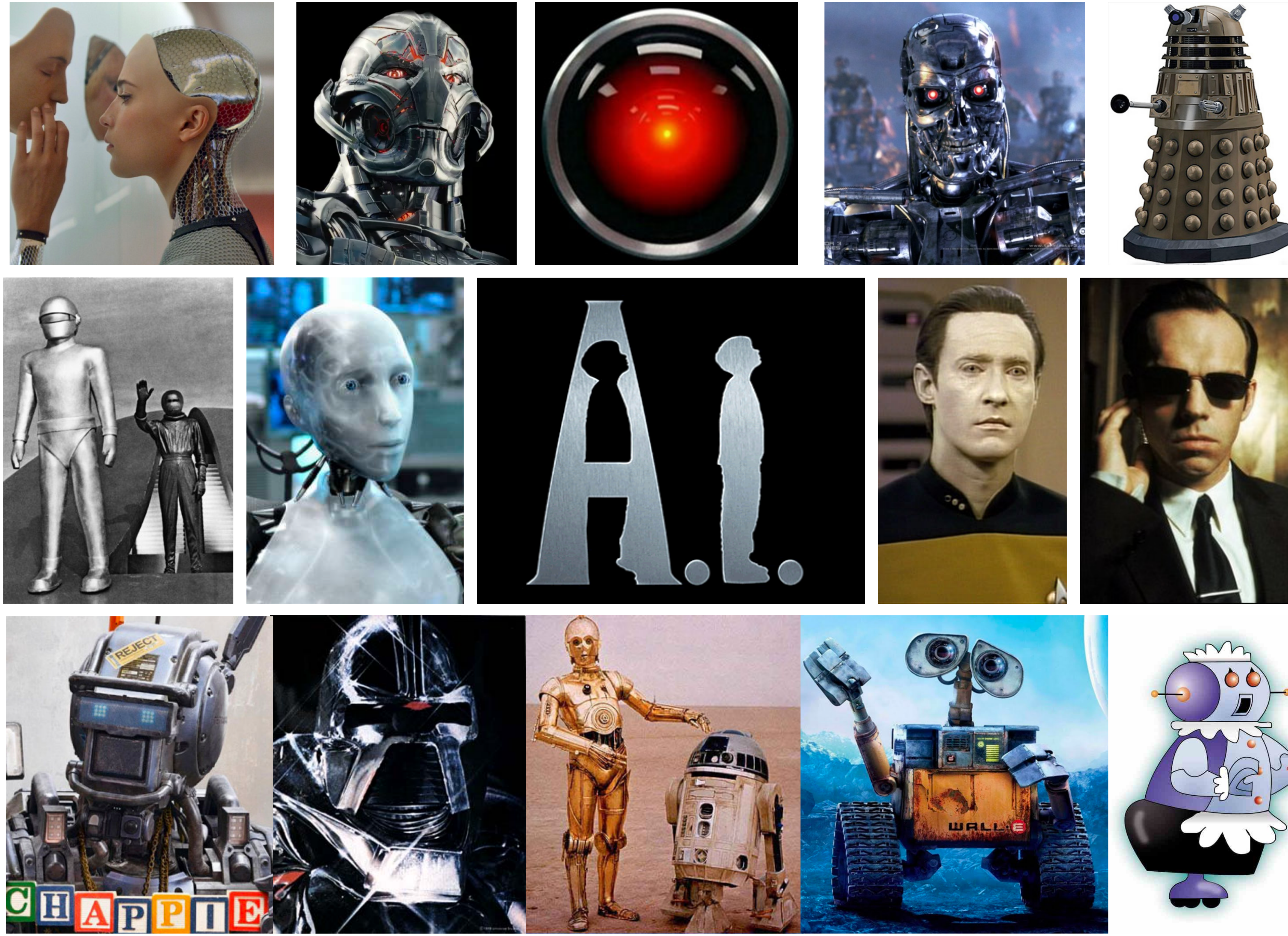
Intelligent Agents

CPSC 470 – Artificial Intelligence

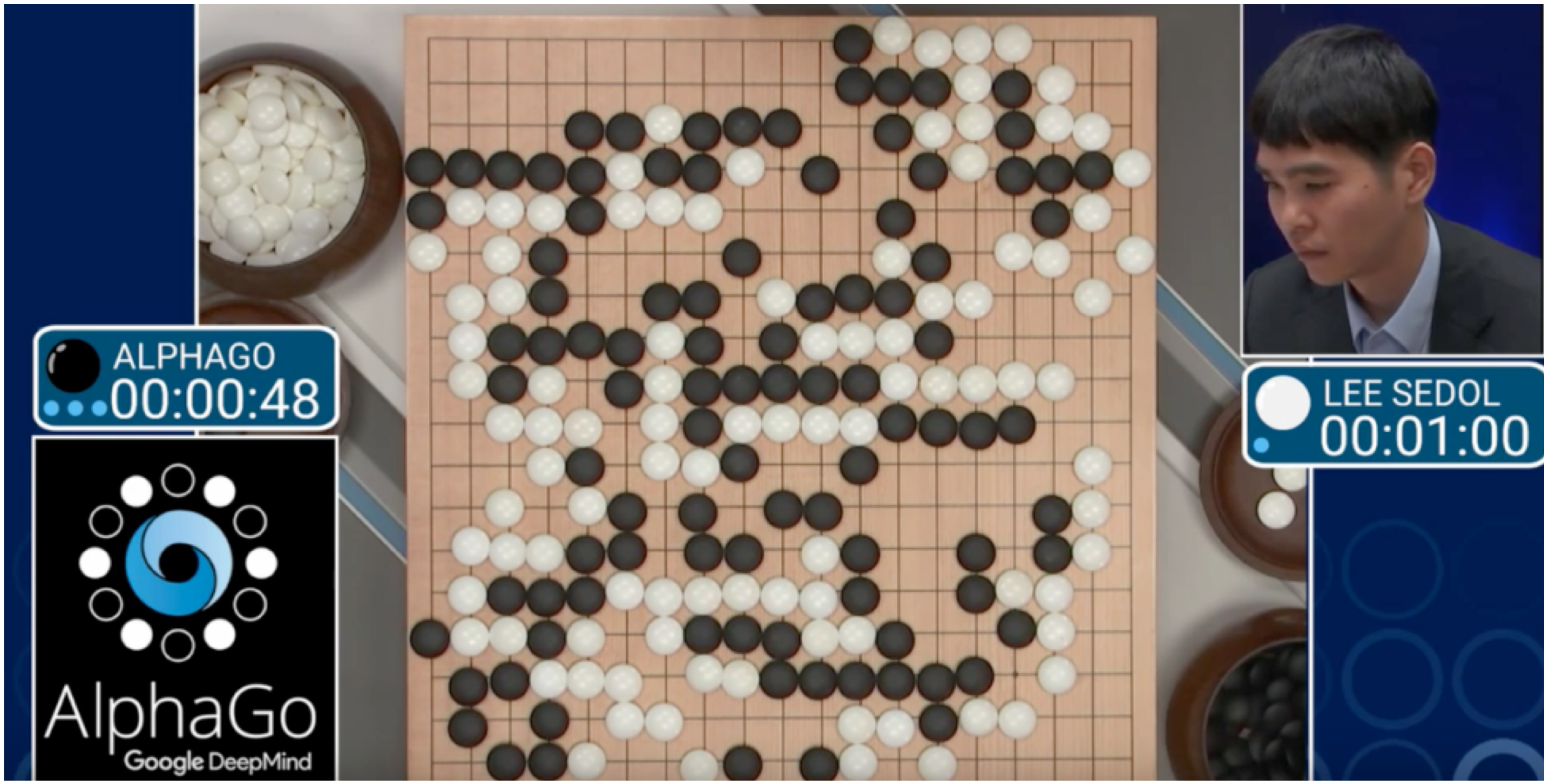
Brian Scassellati

Today’s Objectives

- Why hasn’t AI fulfilled all of our expectations and promises?
  - Historical overview
  - Unifying Formalism
- How can we understand this shortcoming?


A.I.
REJECT
CHAPPIE
WALL·E

Google's AlphaGo and AlphaZero



- 2016: AlphaGo beats Lee Sedol 4-1
- 2017: AlphaGo beats Ke Jie (#1 ranked human player
- 2018: Self-trained AlphaZero beats AlphaGo 100-0

Silver, David, et al. "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." *Science* 362.6419 (2018): 1140-1144.

Boston Dynamics' Atlas

Source: Boston Dynamics. https://www.youtube.com/watch?v=WcbGRBPkrps

Facebook's Deep Face

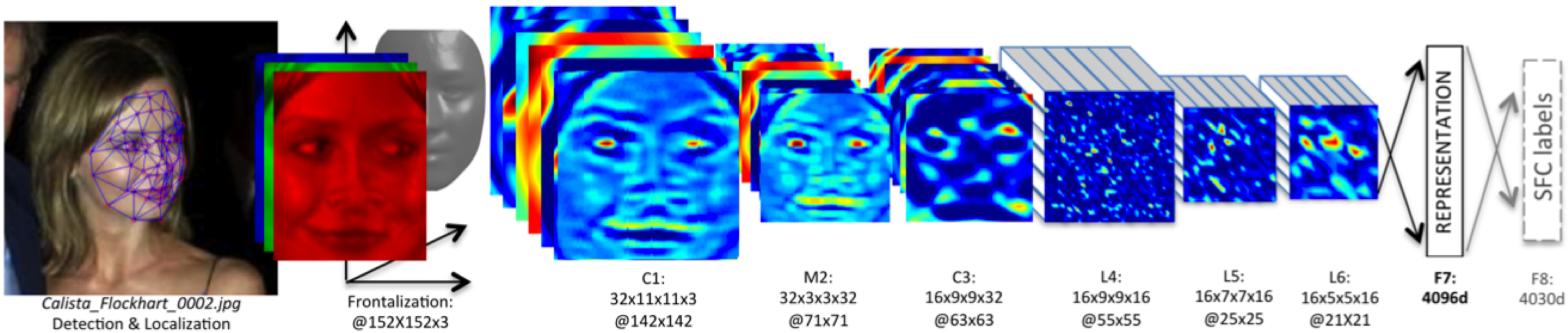


Figure 2. **Outline of the *DeepFace* architecture.** A front-end of a single convolution-pooling-convolution filtering on the rectified input, followed by three locally-connected layers and two fully-connected layers. Colors illustrate feature maps produced at each layer. The net includes more than 120 million parameters, where more than 95% come from the local and fully connected layers.

Face Recognition: Detect → Align → Represent → Classify

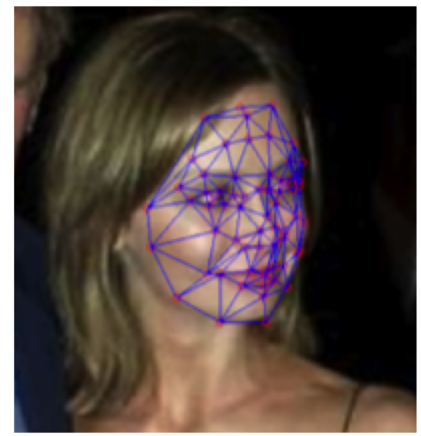
- Labeled Faces in the Wild dataset with 4M images
- Deep Face: 97.35% accuracy with 120M parameters
- Human performance: 97.5% accuracy

Taigman,Yang, Ranzato, & Wolf (Facebook), CVPR 2014

Why don't we have intelligent machines?

+

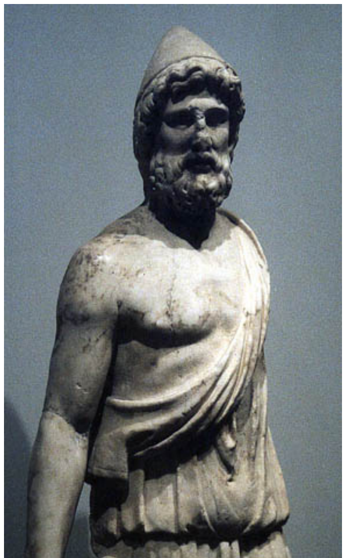
+



= ?

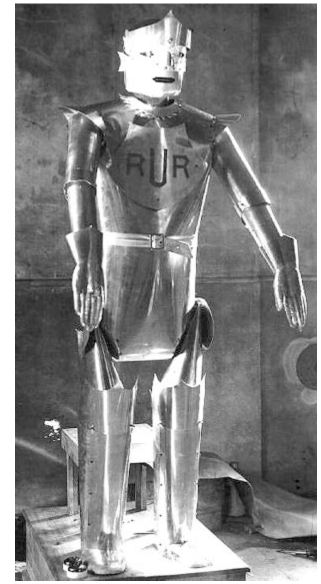
To answer this, we need a bit of history….

Pre-history



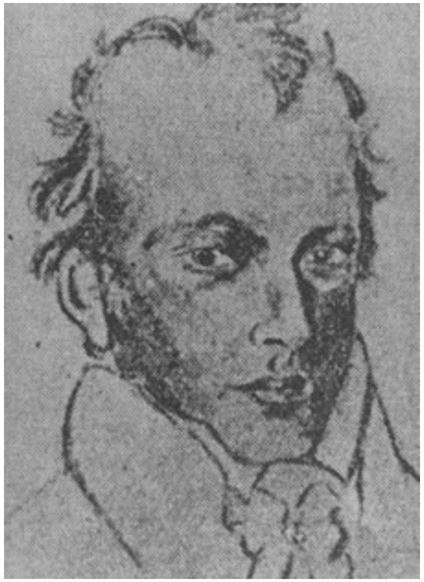
Vulcain. IIe siècle après J.-C. Photo © Maicar Förlag – GML

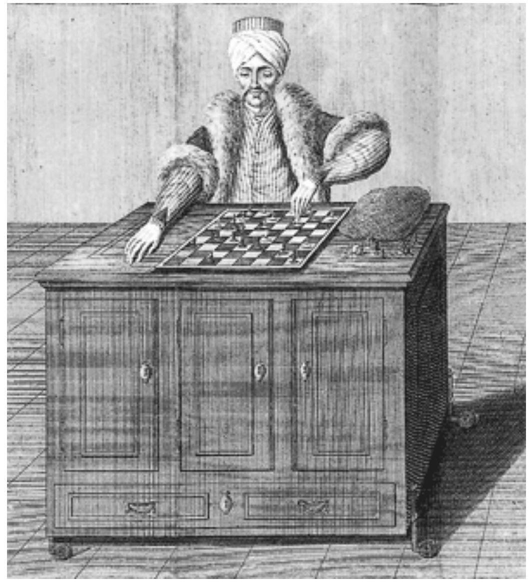
Talos. Contemporary. Photo © Maicar Förlag – GML.

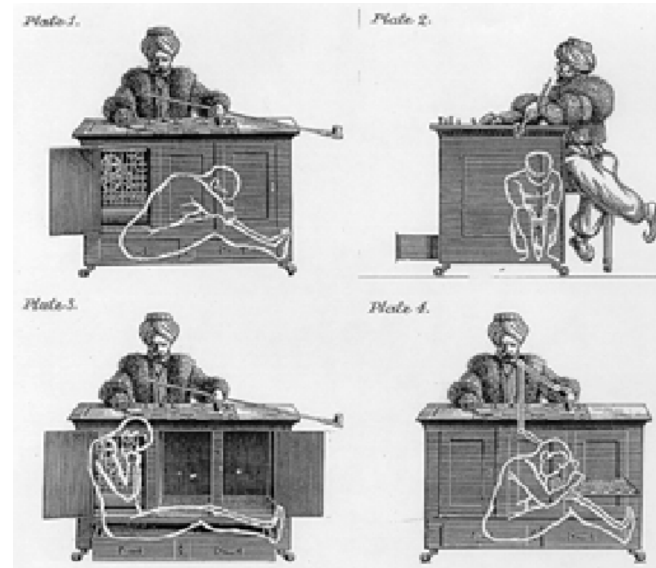


- Hephaestus and Talos in Greek and Cretan myth
- Golem of Jewish folklore
- Karel Capek’s *Rossum’s Universal Robots (1920)*

Automata

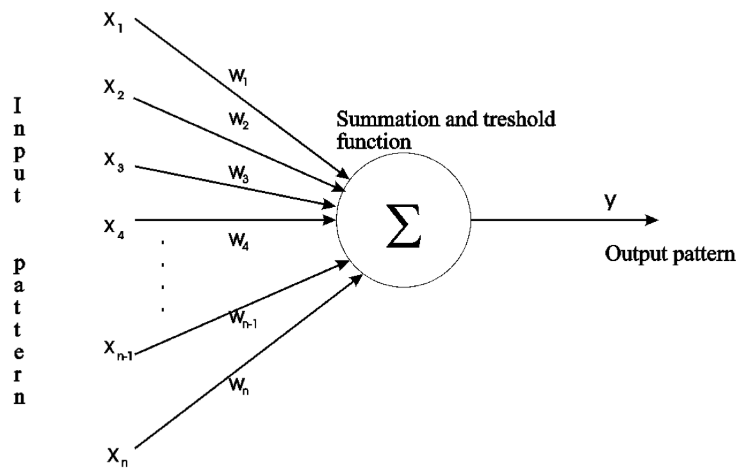


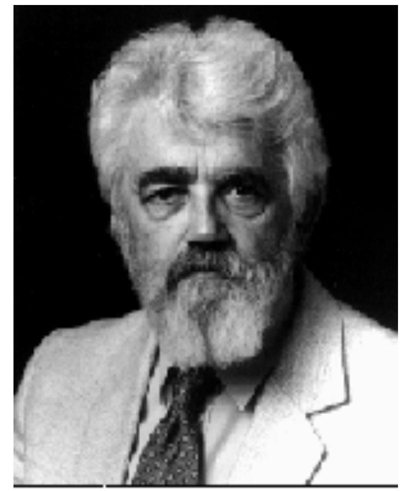




- Wolfgang von Kempelen invented *The Turk* in 1770
- Fashioned from wood, powered by clockwork, dressed in a Turkish costume—and capable of playing chess.
- Connected to Napoleon, Franklin, Poe
- Exposed in 1834, Destroyed by fire in 1854

AI: the early years (1943-1956)





John McCarthy

- McCulloch and Pitts (1943): artificial neuron
- 1956 Dartmouth conference (10 people, 2 months) organized by John McCarthy
  – Coined the name “Artificial Intelligence”

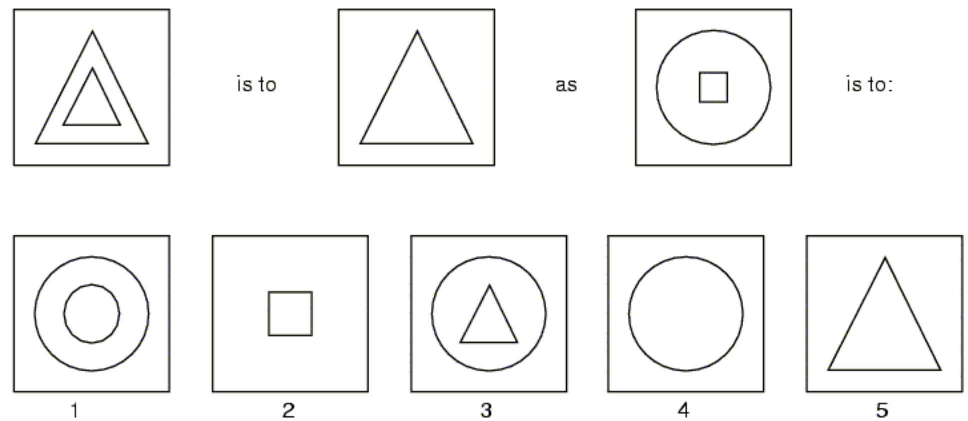
Era of Great Expectations (1952-1969)

SAINT: Symbolic Automatic Integrator

$$\int \frac{x^4}{(1-x^2)^{\frac{5}{2}}} dx$$

Try $y = \arcsin x$, yielding:

$$\int \frac{\sin^4 y}{\cos^4 y} dy$$



- Newell and Simon's GPS
- Samuel's checkers program (1952)
- Slagel's SAINT program (1963)
- Evans's ANALOGY program (1968)

Early Predictions of Success

- “It is not my aim to surprise or shock you ... . But the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until -- in a visible future -- the range of problems they can handle will be coextensive with the range to which the human mind has been applied.”
- “Machines will be capable, within twenty years, of doing any work that a man can do.”
- “…within a generation the problem of creating 'artificial intelligence' will be substantially solved.”

A Veneer of Intelligence

- ELIZA: A Rogerian therapist created by Joseph Weizenbaum in 1966

- Weizenbaum’s observations:
  - Emotional attachment
  - Believed to be real
  - Worried practicing psychotherapists
  - Worried AI researchers -- a serious solution to producing natural language?

Eliza: Hello. I am ELIZA. How can I help you?

Me: I am very tired today.

Eliza: How long have you been very tired today?

Me: It seems like I've been this way all week.

Eliza: Please go on.

Me: I just can't seem to sleep.

Eliza: Can you elaborate on that?

Me: I close my eyes, but I keep worrying.

Eliza: Why do you say your eyes but you keep worrying?

A Dose of Reality (1966-1969)

- AI failed to live up to expectations
- Intelligence was symbolic, logical computations without any context
  - Chess Playing
  - Solving Logic Problems
  - Solving Calculus Problems
- Input/Output tasks are seen as secondary
  - 1966: Summer Vision Project at MIT

Knowledge-Based Systems as the road to intelligence? (1969-1979)

- Stand-alone, single-purpose systems with specific knowledge of a domain (Expert Systems)
  - DENDRAL: infer molecular structure based on mass spectrometer readings
  - Edward Feigenbaum (AI), Joshua Lederberg (genetics), Carl Djerassi (chemistry)

If there are two peaks at *x1* and *x2* such that

(a) x1 + x2 = molecular mass + 28;

(b) x1 – 28 is a high peak;

(c) x2 – 28 is a high peak;

(d) At least one of x1 and x2 is high

Then there is a ketone (C=O) group

AI goes Industrial (1980-1988)

- Expert systems saved money!
  - R1 at DEC (John McDermott, 1982)
    - Helped configure orders for new computer systems
    - Saved an estimated $25-40 million per year
- Concentration on real-world, practical tasks
  - First commercial machine vision systems for parts inspection
  - Assembly-line robotics

Fad Computing (1986-present)

- Belief Nets
- The return of neural nets
- PDP : parallel distributed processing
- Hidden Markov models (HMMs)
- POMDPs

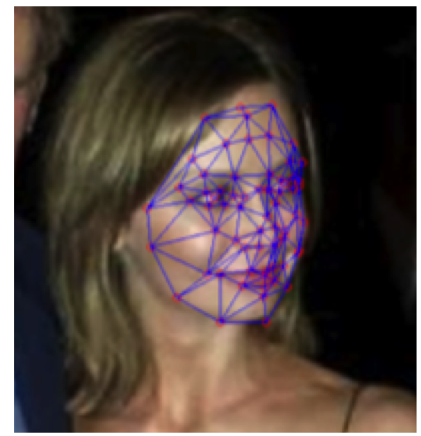
Agents, Big Data, and AI for the Masses (2001-present)

- Refocus on the complete package
- Integration becomes an issue
- Many put their hope in reliance on amassing huge data sets
- AI projects become mainstream

Why don't we have intelligent machines?

+

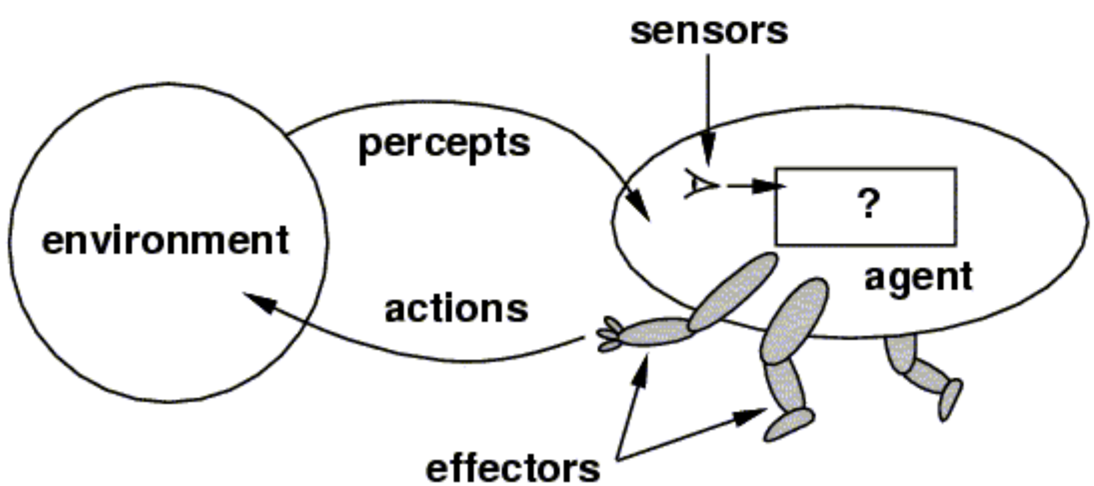
+



= ?

To answer this, we also need a bit of formalism…

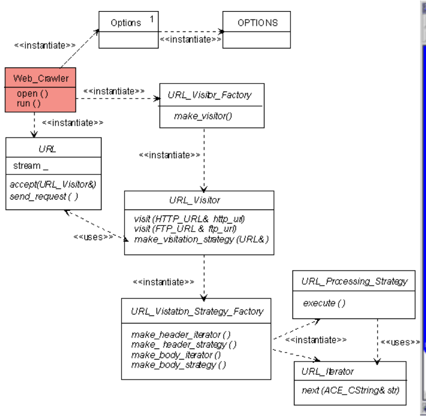
A Unifying Framework: The Intelligent Agent



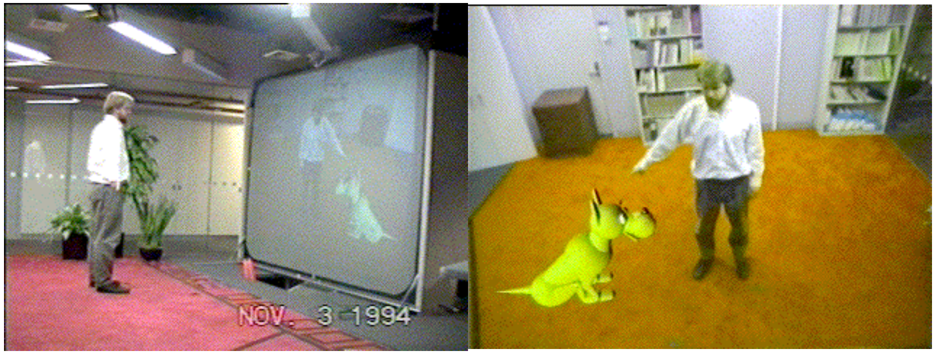
- Rational Agent
- Performance Measure

Agent Examples

Web Crawler

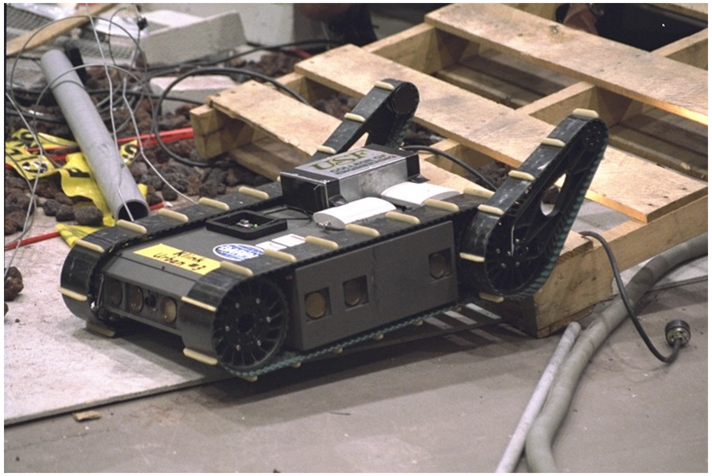


ALIVE (Maes, 1994)

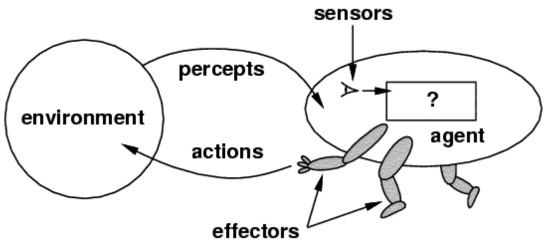


Simulated Soccer

Search and Rescue



Agent Examples



Agent Type	Performance	Enviornment	Actuators	Sensors
Web crawler	# of archived pages, minimal bandwidth	Hyperlinks, HTML and PDF documents	Web navigation, text processing	Network traffic
Simulated Soccer	Scoring, compute cycles	Simulated (standardized) soccer arena	Interface to physical simulator	Interface to simulated vision sensors
ALIVE: Silas	User satisfaction	Fixed room, contents of that room	Display screen, speaker system	Cameras in room, microphones
Search and Rescue Robot	Survey location, find survivors	Collapsed building	Travel, signal base, emit sounds	Camera images, temperature, touch sensors, etc.

Characterizing Environments

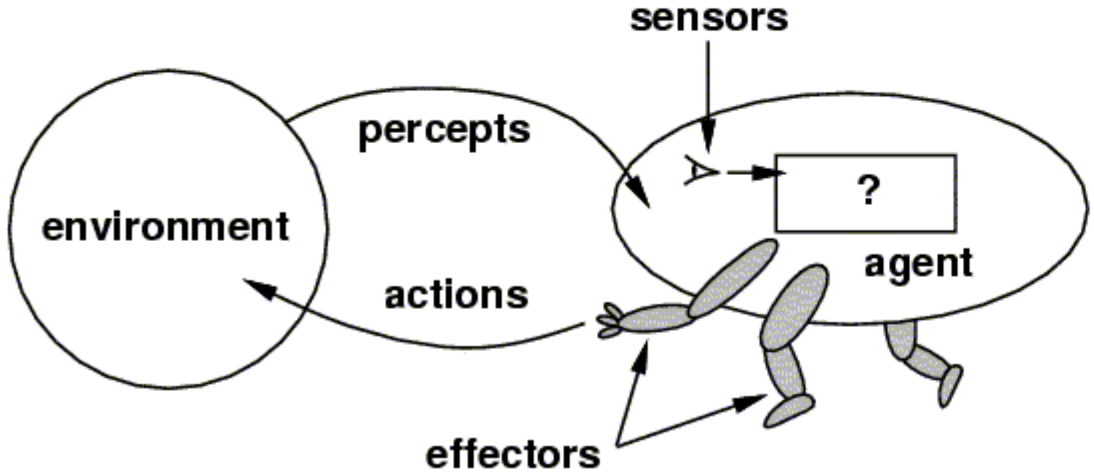
- Fully Observable vs. Partially Observable
  - Do sensors give complete world state?
- Deterministic vs. Stochastic
  - Can next state be derived from current state and action?
- Episodic vs. Sequential
  - Does the quality of an action depend only upon the current sensory state?
  - (Are sensing/action pairings atomic?)
- Static vs. Dynamic
  - Does the environment stay the same while the agent decides to act?
- Discrete vs. Continuous
  - Are there a limited number of distinct percepts and actions?

Characterizing Sample Environments

Environment	Observable	Deterministic	Episodic	Static	Discrete
	Do sensors give complete world state?	Can next state be determined by current state and action?	Does quality of an action depend only on current state?	Does the env. stay the same while the agent thinks?	Are the number of percepts and actions limited?
Chess (no clock)	Fully	Yes	No	Yes	Yes
Poker	Partially	No	No	Yes	Yes
Taxi driving	Partially	No	No	No	No
Image analysis	Fully	Yes	Yes	Semi	No
Part-picking robot	Partially	No	Yes	No	No

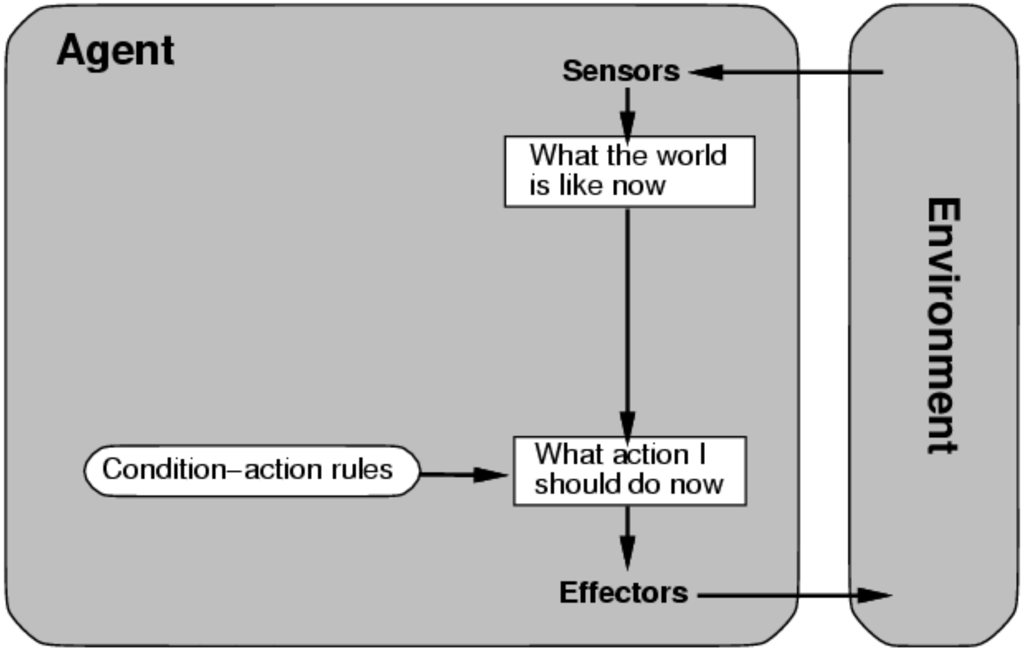



Characterizing Agents



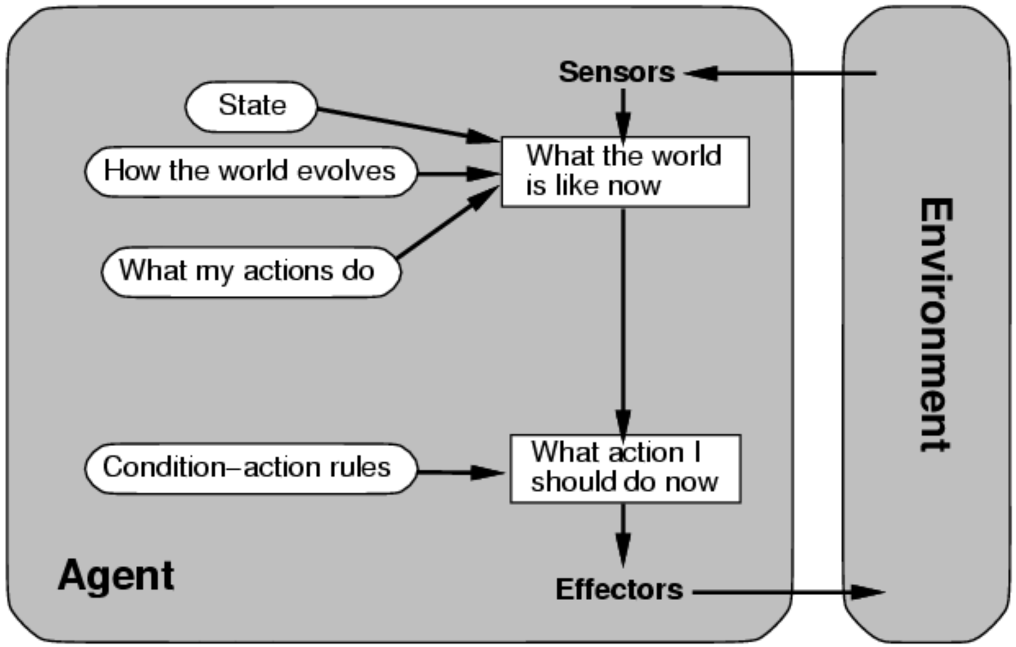
- We can characterize the environment
- Are there important differences in the agent itself?

Reflexive Agent



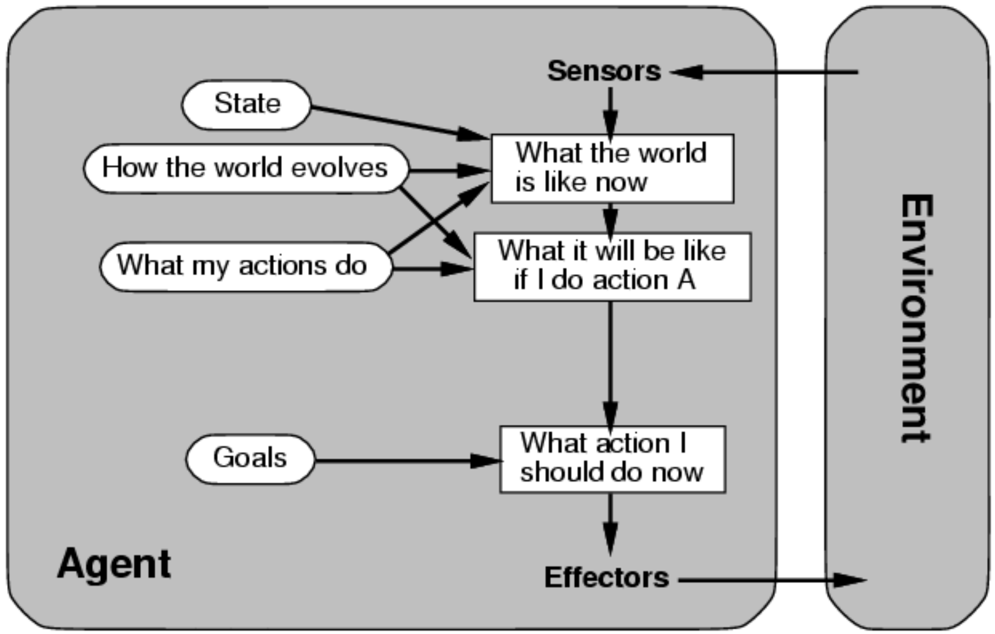
- Fixed set of condition-action rules
  - If car-in-front-is-breaking then apply-breaks
- Efficient implementations, limited applications

Reflex Agent with Internal State



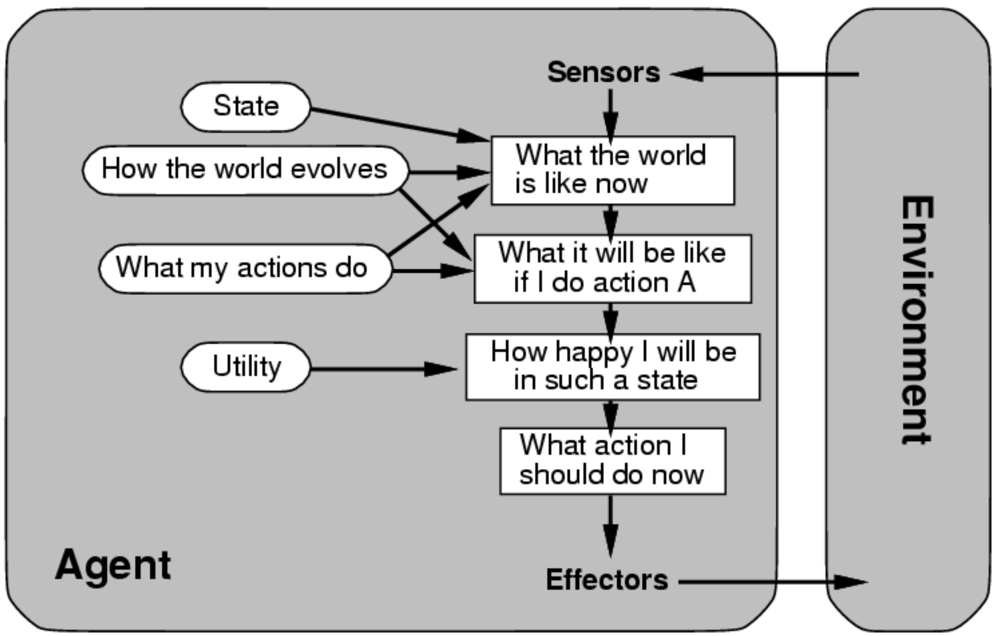
- Keeps track of the world or its own actions
- Allows for variation in behavior
- Must have a model of how this internal state evolves

Goal-Based Agent



- Consideration of future actions
- Allows for actions to be directed
- Varied implementations, varied usefulness

A utility-based agent



- Utility function maps all states onto reals
- Act to maximize utility
- Varied implementations, varied usefulness

What does this formalism get us?

- Describe variations among
  - Goals, percepts, actions
  - Environments
  - Agent architectures
- How easy is it to cross these boundaries?

OK, so why don't we have intelligent machines

- Lessons from history
  - We are not really good judges of intelligence
  - We want to believe the hype
  - Easy to make glorious predictions
- Why is it so hard?
  - Range of environments and goals makes it difficult to leverage other work
  - Our understandings and expectations are changing

Administrivia

- Sign up on canvas
- Schedule and reading assignments
- Remember:
  - If you have questions
    - Piazza
    - Email scaz@cs with ”CS470” in the subject

Up Next…

- Today: Problem Set 0 out, due next Wednesday
- Friday: Introduction to Python with Meiying
- Monday: No class – MLK day
- Wed+Fri: Solving problems with search