Foundations of AI Introduction

LECTURE 1

BRIEF HISTORY OF AI

OVERVIEW OF AGENTS

CS5100

Today

What is AI

What can it do

Timeline

Fears-Ethical Issues

Course Details

In-class survey

So what is Artificial Intelligence?

Historical perspective:

- Handbook of AI: the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behavior

Do we believe that?

So what is Artificial Intelligence?

Historical perspective:

- Handbook of AI: the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics **we associate with intelligence in human behavior**

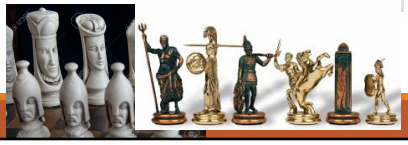
Which is harder for a machine? Why?

What makes chess hard? What makes it easy?

Decide on moves

VS

Recognize pieces and move them



Artificial Intelligence?

Historical perspective:

- Handbook of AI: the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics **we associate with intelligence in human behavior**

What we think requires intelligence is often wrong

- Elephants don't play chess -- Rodney Brooks
- People perform behaviors that on the surface seem simple
- **since they require little conscious thought.**
- Eg. Recognizing a friend in a crowd.

- Why does it seem easy to recognize a friend?

Artificial Intelligence?

Historical perspective:

- Handbook of AI: the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics **we associate with intelligence in human behavior**

It's a moving Target: once we come up with an algorithm or technology to perform a task, we tend to re-assess our beliefs that it requires intelligence or is AI

- Beating the best human chess player was a dream of AI from its birth
- Deep blue eventually beats the best
- "Deep Blue is unintelligent because it is so narrow. It can win a chess game, but it can't recognize, much less pick up, a chess piece. It can't even carry on a conversation about the game it just won. Since the essence of intelligence would seem to be breadth, or the ability to react creatively to a wide variety of situations, it's hard to credit Deep Blue with much intelligence." **Drew McDermott**

Artificial Intelligence?

Historical perspective:

- Handbook of AI: the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics **we associate with intelligence in human behavior**

Moral: What tasks we think are the hallmark of intelligence has been in flux since the dawn of the industrial age.
What constitutes intelligence in humans is widely debated even in psychology.

What is AI?

Russell & Norvig: Views of AI fall into four categories:

Thinking Humanly	Thinking Rationally
Acting Human	Acting Rationally

The textbook organized around "acting rationally"

but lets consider the others as well…

Thinking Like a Human: AI & Cognitive Science

1960 "AI and the Cognitive Revolution": information-processing psychology replaces behaviorism

- Spurred on by a close connection to AI research that provided mechanism for information processing

Cognitive science brings together theories and experimental evidence to model internal activities of the brain

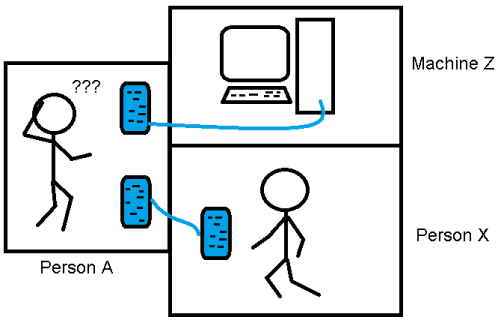
- Explored at different levels of abstraction: Knowledge or Neural Circuits
  - Approaches to validation
    - Predicting and testing behavior of human subjects (top-down)
    - Direct identification from neurological data (bottom-up)
    - Building computer/machine simulated models and reproduce results (simulation)

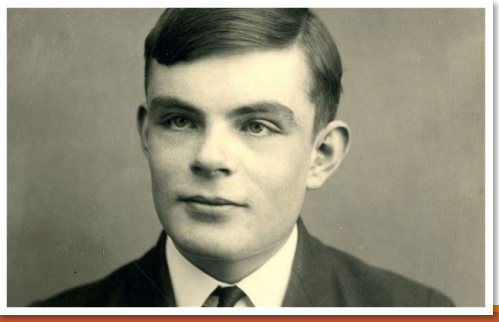
Acting humanly: Turing Test

Turing (1950) "Computing machinery and intelligence":

"Can machines think?"

- → "Can machines behave intelligently?"
- → Can machines' behavior be indistinguishable from human





What would a computer need to pass the Turing test?

Natural language processing: to communicate with examiner.

Knowledge representation: to store and retrieve information provided before or during interrogation.

Automated reasoning: to use the stored information to answer questions and to draw new conclusions.

Machine learning: to adapt to new circumstances and to detect and extrapolate patterns.

Total Turing Test

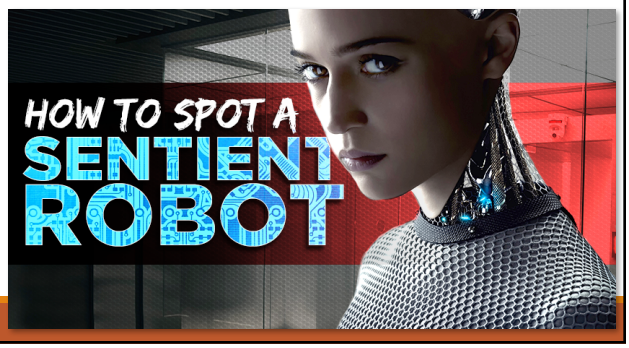
Vision: to recognize the examiner's actions and various objects presented by the examiner.

Motor control: to act upon objects as requested.

Other senses (total test): such as audition, smell, touch, etc.

Acting humanly: Turing Test

- Stimulated the development of:
  - *Natural Language Processing*
  - *Knowledge Representation*
  - *Automated Reasoning*
  - *Machine Learning*

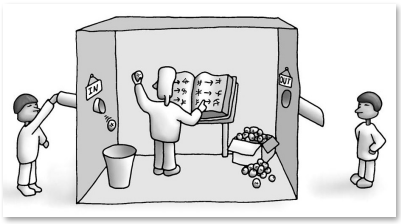


Acting Humanly: The Full Turing Test

In Turing Test, machine is hidden

What about physical interaction?
- Perception and action

True understanding?
- Searle's Chinese Room →→



Social Persuasion?
- "The machine needs to convince the human to do things for it -- to fall in love with it, to serve its own purposes." Tim Tuttle

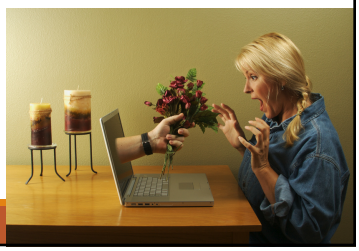
Creativity?

Ex Machina →→

Virtual Human Research

People respond to virtual humans *almost* as if they were real

- **Impression Management**
- e.g. People give socially desirable responses when asked by virtual human (Krämer et al 2003)
- **Mood management**
- e.g. Bad news evokes less anger when delivered by virtual human (Krämer et al)
- **Stereotype bias**
- Whites more threatened by black agents Blascovich et al (McCall et al 2009; Lok et al 2008)
- **Attentiveness**
- e.g., virtual face attracts attention; can distract from other tasks (Takeuchi & Naito, 1995)
- **Persuasion**
- e.g., Messages more persuasive if delivered by virtual human (Bailenson & Yee 2005)
- **Trust**
- e.g., Trust increases using anthropomorphic interface (Rickenberg & Reeves, 2000)



Thinking Rationally: Rational Decision Making

The study of agents that exist in an environment and perceive and act rationally.

Key issue: How do we define rational?

- We will explore this issue from a decision theory perspective

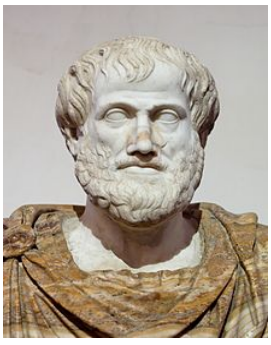
Thinking rationally: "laws of thought"

Aristotle: codify right thinking

- what are correct arguments/thought processes?

Syllogisms provided patterns of correct thinking

- "Socrates is a man, All men are mortal, therefore Socrates is mortal"
- Such laws of thought could/should govern intelligent thinking

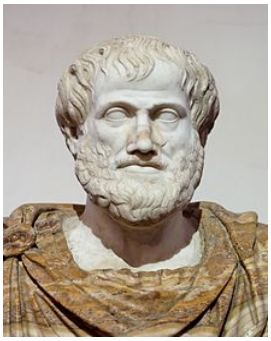


Thinking rationally: "laws of thought"

By the 19th century, logicians had developed the notation for making statements about the world

By 1965, programs existed that could solve any problem expressed in such notation

Logicist tradition within AI builds on such programs to create intelligent systems



Thinking Rationally: Laws of Thought

Problems:

1) **Sometimes there is no correct solution or no time to determine it**
2) **Uncertainty:** Not all facts are certain (e.g., *the flight might be delayed).*
3) **Resource limitations:** Difference between solving a problem in principle and solving it in practice under various resource limitations like time
4) **Representation of knowledge:** Hard to take informal knowledge and express it in a logic

These problems are fundamental to all of AI

- We will see them repeatedly throughout this course

Acting Rationally: rational agent

Agent: entity that perceives and acts autonomously in order to achieve its goals

Rational agent: choose behavior to achieve best possible outcome, given the resource constraints (such as available information or limited time) and uncertainty

AI's Impact Today

What can I help you with?

24 Hours a day AI spawned technologies are interacting with you, assisting you, profiling you through your data, manipulating your money.

Assistants: Siri

Legal Profession
- Discovery phase: NLP, KR, data mining, pattern detection, SNA
- Detect Conspiratorial Behavior

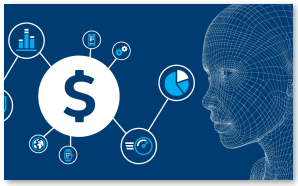
Banking & Investment

Medical Diagnosis: Interpretation of medical images

Game playing: Chess (Deep Blue), video game bots

Intelligent Tutors

Autonomous Vehicles: Satellites, Cars, Drones, Vacuum cleaners



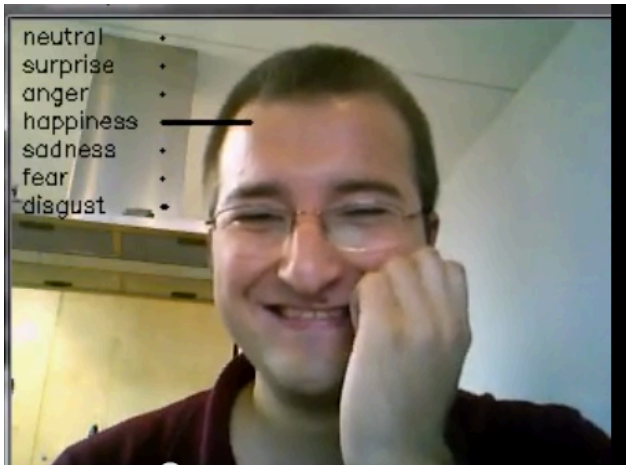
Example: Autonomous Cars

Early On (20th Century, mid 70s on)
- Eureka Prometheus Project (80s and 90s) (1.5 billion)

DARPA Grand Challenge
- 2004: 250 mile route in Mojave desert, best performance 7 miles
- 2005: 5 teams succeed
- Urban Challenge: Had to obey traffic laws, 6 teams succeed

Car manufacturers promise self-driving cars by 2020

Computer Vision



https://www.youtube.com/watch?v=LdQw8PSV2P8

Example: Financial Systems

Wall street
- dominated by high speed computers making buy /sell decisions

Your ATM transactions
- overseen by sophisticated fraud detection algorithms

Example: Our Social Interactions

Your web behavior
- Being tracked by preference elicitation algorithms to determine your likes and wants

Email and phone transactions
- Analyzed by a range of security agencies

Facebook Manipulated User News Feeds To Create Emotional Responses without their knowledge

How did we get here? AI Timeline

AI was a millennia-old dream of making machine in humankind's image, waiting for scientific and more importantly technological progress to catch up.

Here are just a few key points….

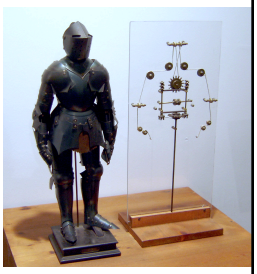
AI: timeline of ideas

Pre-AI Groundwork: a very small sampling

First millennia BC:

- Chinese, Indians and Greeks: formal/mechanistic deduction
- Aristotle works on logical deduction
- Machines to calculate navigation are developed

15th Century: Da Vinci's mechanical knight design



AI: timeline of ideas

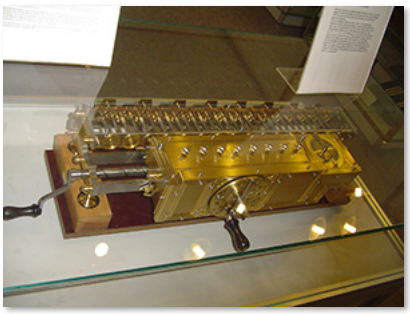
Pre-AI Groundwork: a small sampling

17th Century: Liebniz builds calculating machines

- Also advocates for binary system but used based 10 in his machine

17th Century: Liebniz, Hobbes, Descartes argue rational thought as a form of calculation

- Lays the groundwork for the physical symbol system hypothesis of Newel and Simon



AI: timeline of ideas

18th Century: Birth of economics from philosophy

- Jeremy Bentham, Adam Smith, Bernouli lay out ideas about how utility and probability provide a framework for human decision-making
  - U: Utility as a measure of pleasurable/desirable an outcome is
  - Pr: Probablility – the likelihood of that outcome
  - Pr x U: Expected utility – people seek to maximize their expected utility
- As we will see these ideas have strongly influenced AI's approach to decision-making under uncertainty

18th Century: David Hume argues that emotion is critical to human reasoning

- 'Reason is and ought to be the slave of passion'
- Mirrors early AI work by Simon, Minsky and Sloman and current work on affective computing

"The principle of utility judges any action to be right by the tendency it appears to have to augment or diminish the happiness of the party whose interests are in question... "

AI: timeline of ideas

19th Century: Industrial Revolution reinforces idea of intelligence as mechanical

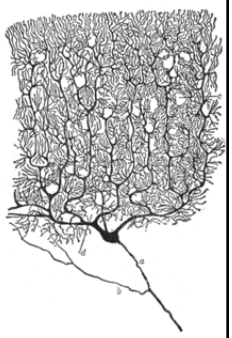
- Babbage designs Analytical Engine, a programmable machine

19th-20th century

- Santiago Ramon y Cajal uses Golgi's staining to meticulously isolate and draw the brain's neuronal structures, proposes the neuronal doctrine.

20th Century: the birth of the computer

- Turing, von Neumann and others lead to development of digital computers
- WWII (1940s) code breaking machines use Turing's and von Neumann's ideas



AI: timeline of ideas

20th Century
- Study of mathematical logic by Boole and Frege
- Church-Turing Thesis: Mathematical reasoning could be mechanized

1956: Birth of AI at Dartmouth Conference
- Minsky McCarthy, Shannon, Selfridge, Samuel, Newell, Simon, Amarel, etc.
- "intelligence can be so precisely described that a machine can be made to simulate it."
- Newel and Simon propose the concept of **Intelligence as search** – which we will discuss

1957: Rosenblatt creates the perceptron algorithm, early neural network design
- 1969 Perceptron book by Minsky and Papert leads to stagnation of research on neural networks
- 1980 sees re-emergence with work on Parallel Distributed Processing by McClelland and Rumelhart
- If time permits we will touch on perceptrons

AI: timeline of ideas

1980: Knowledge Based Approaches
- Expert systems ranging from computer configuration (R1) to medical diagnosis
- Knowledge Acquisition bottleneck

1990s: Statistical techniques in Machine learning
- Lead to dramatic improvements

1990s: Agent based approaches

2000s: the birth of big data
- Machine Learning merges with statistical techniques

AI's cycle of failed expectations

Examples

1958, Simon and Newell: "within ten years a digital computer will be the world's chess champion"

1965, Simon: "machines will be capable, within twenty years, of doing any work a man can do.

1967, Minsky: "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved.

1970, Minsky: "In from three to eight years we will have a machine with the general intelligence of an average human being."

Such optimism lead to AI winters as AI failed to meet expectations

- Reduced attendance at conferences, reduced federal funding

What were the roadblocks?

Limited computer power: There was not enough memory or processing speed to accomplish anything truly useful.

Intractability and the combinatorial explosion. Karp: *many problems can probably only be solved in exponential time (in the size of the inputs).*

Commonsense knowledge and reasoning. Many important artificial intelligence applications like vision or natural language require enormous amounts of information about the world

Moravec's paradox: Proving theorems and solving geometry problems is comparatively easy for computers, but a supposedly simple task like recognizing a face or crossing a room without bumping into anything is extremely difficult.

Failure actually good for AI

Forced AI to Explore new ideas

Statistical techniques revitalized Machine Learning

Old ideas reinvigorated using new approaches and technologies as well as new applications

Neural Networks
- Early 1950s work on neural networks falls out of favor after Minsky and Papert book on Perceptrons identifies representational limits
- Deep Learning: Now back in a wide range of applications involving large data sets that are now available

Work on Emotion
- Initially argued as critical for AI by Simon and Minsky
- Fell out of favor during rational period
- Now a key new area Affective Computing: as man and machine increasingly interact.

Earlier ideas about knowledge representation re-entering ML,
- May transform purely statistical techniques

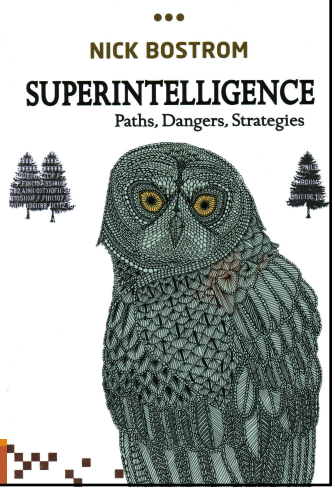
Success brings fears and ethical concerns

Elon Musk
- "Competition for AI superiority at national level most likely cause of WW3"
- "If I were to guess at what our biggest existential threat is … With artificial intelligence, we are summoning the demon"
- AI "potentially more dangerous than nukes."

Stephen Hawkins
- "I think the development of full artificial intelligence could spell the end of the human race"

Vladimir Putin
- "Artificial intelligence is the future not only of Russia but of all of mankind,"
- "Whoever becomes the leader in this sphere will become the ruler of the world."

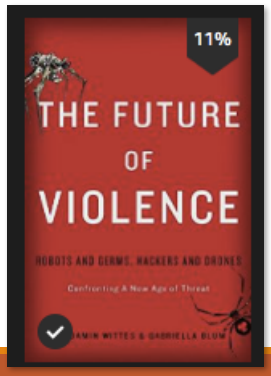
Real immediate threat likely not some super AI making people irrelevant



AI in service of people

People using AI techniques to model and influence human behavior

- "There is a digital mosaic of you stored on public and private computer servers around the world."
- "websites you visit, the toll booths you pass through, the purchases you make online or with credit cards, the prescriptions you fill, the phone numbers you dial, the e-mails you send, the library books you check out, the specific pages you have read on your Kindle, the restaurants at which you make online reservations, the steps you take as measured by your Fitbit, the photos you post on Facebook, and the photos that others post of you"
- This data **is** being used by individual, groups, governments and corporations to analyze you, using AI spawned technologies



Immediate ethical issues from AI *in service of people*

How would you program …

A Battlefield Robot's decision to shoot

Self-Driving Car's decision whether to avoid

- A cat crossing the road
- A child crossing the road

AI in Fiction: Mirroring our expectations and fears

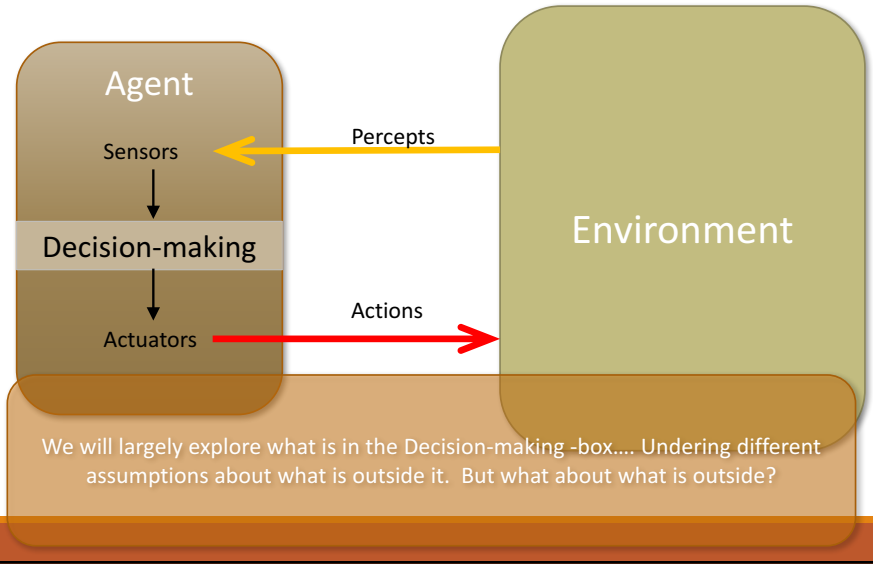
Bubo, Clockwork Owl in ancient Greek mythology **(~3 milennia ago)**
- Sent by the god Athena to help Perseus

Stanislaw Lem (Polish Sci-Fi author) predicted (1964):
- political parties compete to show voters who had best AI to solve country's problems
- swarms of extremely small robots would attack enemies

The Good, Bad and Ugly Robots in Film/TV:


"Intelligent" AGENTS

Agent Based AI



Acting rationally: rational agent

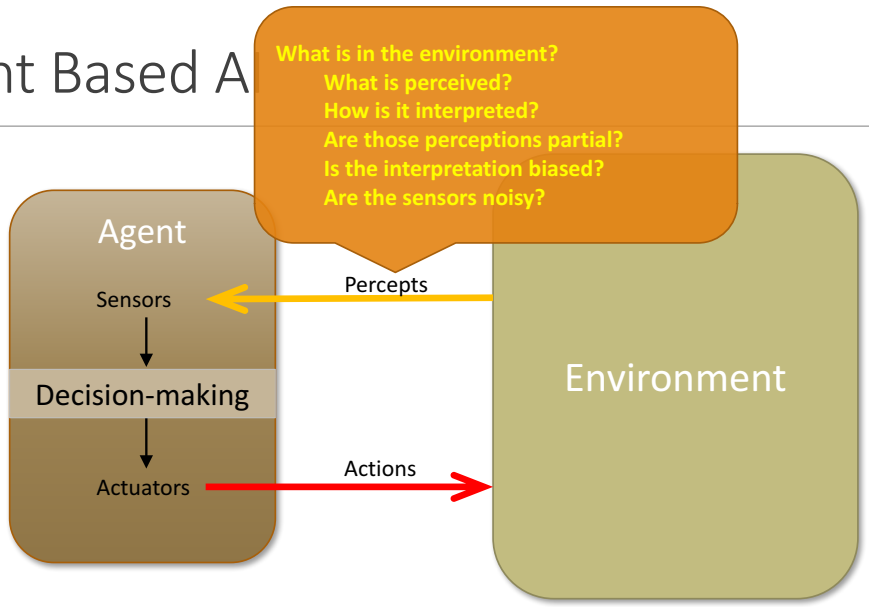
Agent: entity that perceives and acts

Rational Agent: choose behavior that maximize goal achievement, given the available information

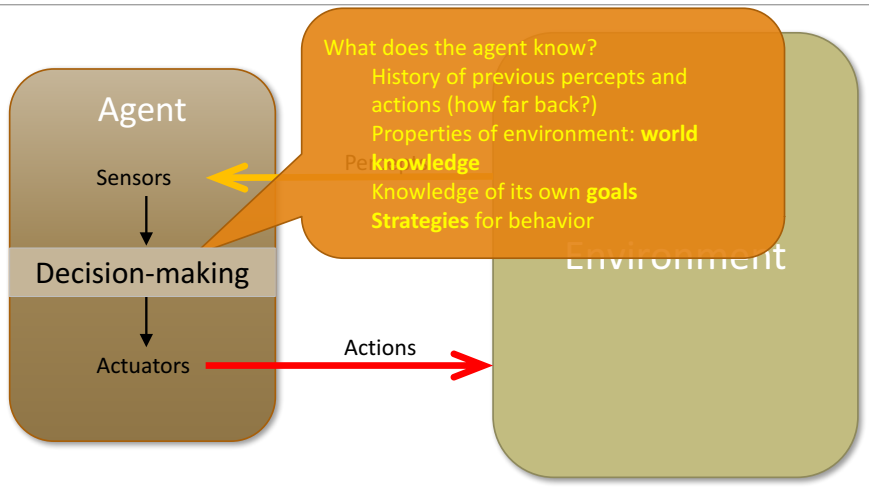
For each possible **sequence of percepts**, a **rational** agent should select an action that is expected to **maximize its performance measure**, given the **evidence provided by the percept sequence** and whatever **built-in knowledge the agent has.**

What makes this hard? What makes it hard for you?

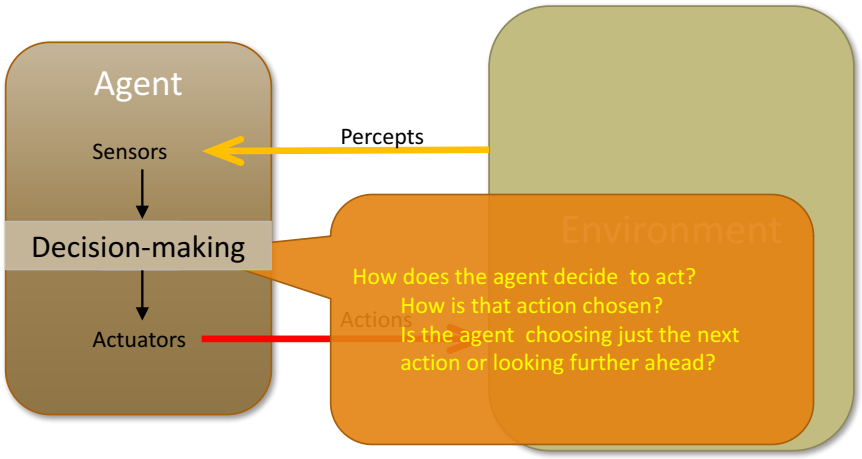
Agent Based A



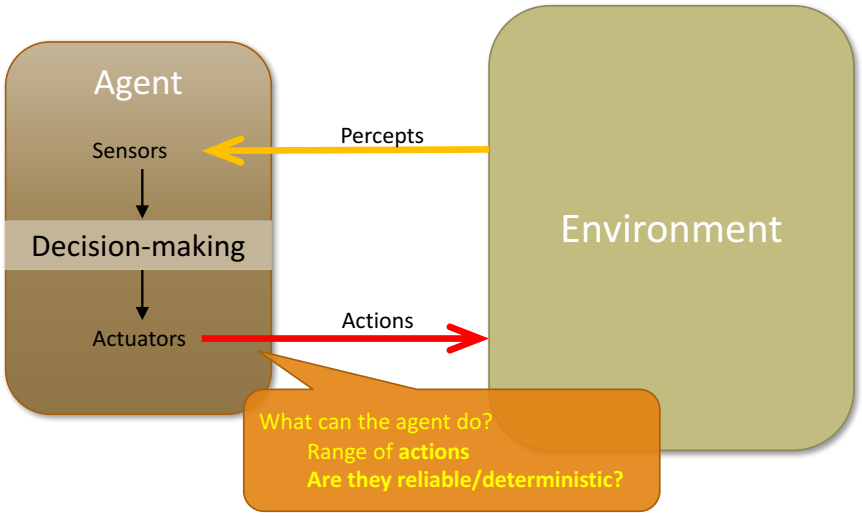
Agent Based AI



Agent Based AI



Agent Based AI



Perception: What makes it difficult?

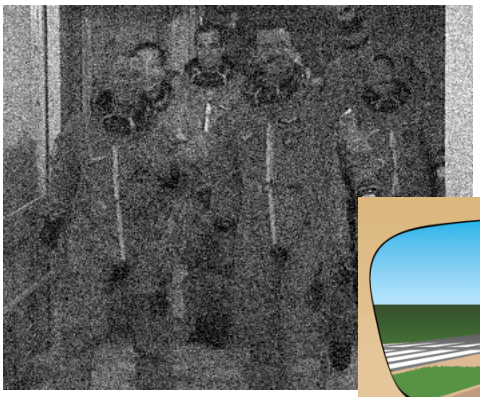
226	226	223	223	226	226
226	226	223	223	226	226
226	226	223	223	226	226
125	125	133	128	120	116
125	125	133	128	120	116
125	125	133	128	120	116
125	125	133	128	120	116
125	125	133	128	120	116
125	125	133	128	120	116
125	125	133	128	120	116
123	123	113	111	120	115
119	119	116	115	125	112
121	121	106	114	120	116
106	106	112	110	118	127
104	104	109	117	102	109
106	106	103	118	107	102
110	110	104	116	103	108
115	115	108	117	103	97
106	106	111	110	111	116
108	108	112	111	94	115
115	115	107	109	102	112
110	110	112	107	118	104
113	113	111	107	103	111
113	113	115	104	115	118
104	104	106	107	111	115

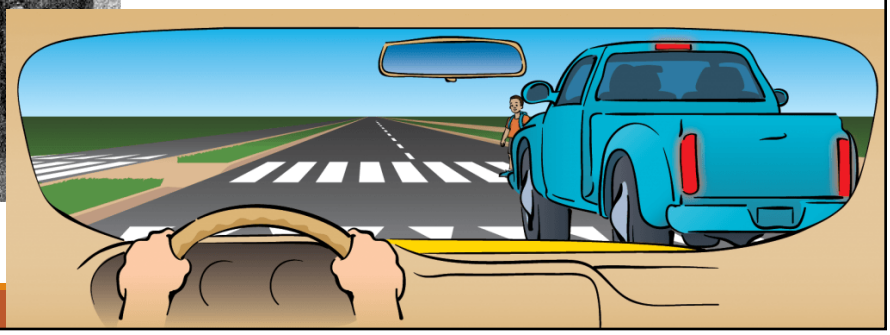
What objects are in image?

- Is it a person
- Is it a car coming at you on the road?

Key:
A major portion of our brains are involved in visual processing
We are "blind" to that processing or its limitations
We tend to underestimate the difficulty of the task

Perception: noise and partial information





Perception: What to attend to depends on decision task

Decision-making

- Imagine you want to take a winter holiday road trip
  - Eg you are writing winter trip planning software for NU students
- You need to formulate a representation that allows your software to effectively search for good trips
- **What are the challenges?**
  - **What is your performance criteria**
    - **Define the goals – what is a good trip**
    - **Define effectively/efficiently search**
  - **What factors to be used in your representation**
    - **Graph that identifies roads and distances**
      - **Speed on roads,**
    - **Interesting road side attractions for stops**
  - **How do you deal with uncertainty**
    - **Holiday traffic, Weather conditions**

Key AI Themes:

Intelligence as Search

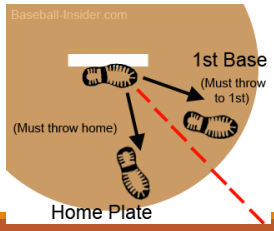
Knowledge Representation
Formulate good representations
(in which to search "smartly")

Reasoning under uncertainty

Decision-making: Agent's knowledge of how the environment works can be a source of uncertainty

Known vs. Unknown: An agent may not know the laws that govern the environment

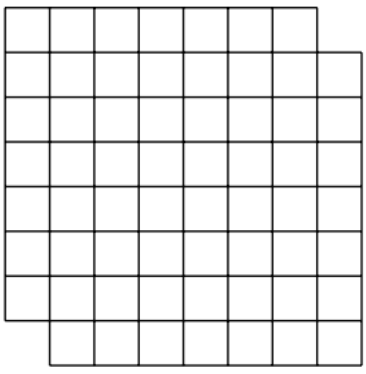
- Often incredibly hard problem.
- Imagine watching a baseball game for the first time
- Balks, infield fly rule, 3rd strike steal, fouling being or not being a strike – all these exceptions

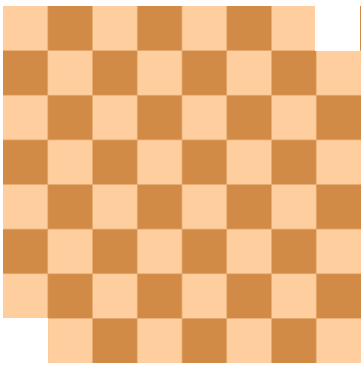


Decision-making: The space in which you search matters

The right Knowledge Representation / Formulation can make problems easier to solve

Is it possible to exactly cover the array with 2x1 tiles?





8×8 Mutilated array.

Decision-making: Knowledge Representation / Formulation can make it easy or hard

Multiplication in Hindu-Arabic numerals

177
_24

Multiplication in Roman numerals

CLXXVII
XXIV

Problem Reformulation is one of the areas of AI from its birth

Agent's decision-making depends on how it represents/models the world and its own actions

Minimally, an agent needs to represent

- What is the world like now?
- What action(s) can it I take?
- How do those actions effect the world.

Action Execution

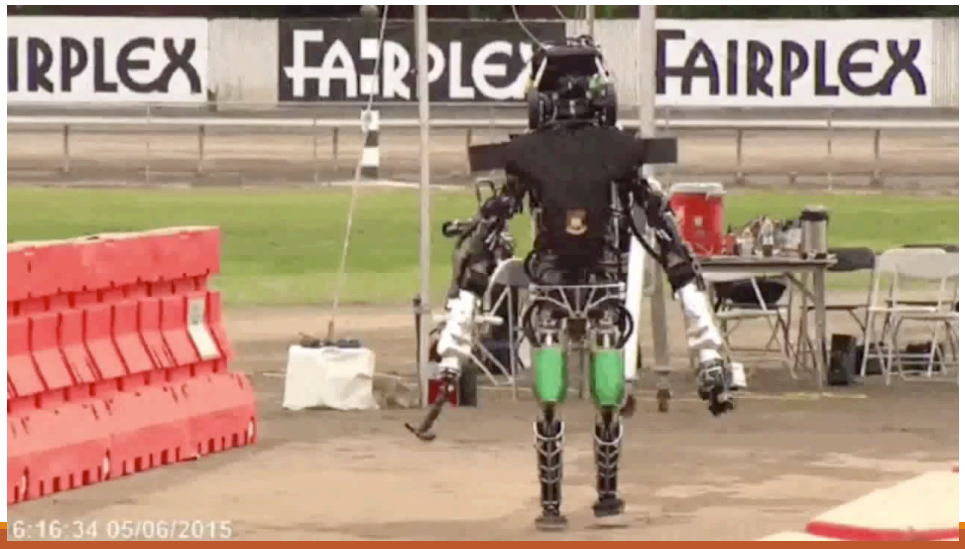
What makes the execution of actions difficult?

Changing state & partial knowledge of the world

The plan may become invalid
- Weather, road conditions, traffic
- Don't know state in which action is executed

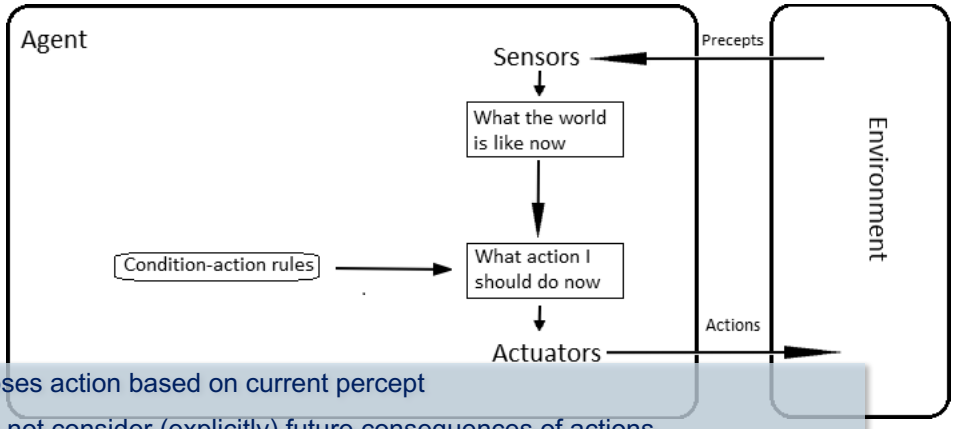
The effects of the actions may be uncertainty
- Icy conditions → brakes no longer stop the car
- Other agents
  - Drivers changing lanes
  - Pedestrians

Action Execution (Darpa Humanoid Challenge)



Types of Agents

Simple Reflex (Reactive) Agent

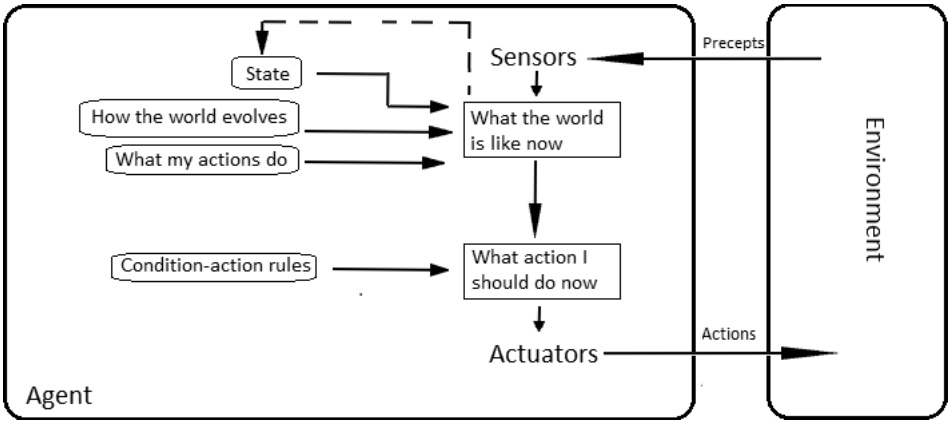


- Chooses action based on current percept
- Does not consider (explicitly) future consequences of actions
  - *Not how the world will be*
- When does this work well/poorly?
- Would you make a Self-Driving Car like this?

John Hopkins Beast

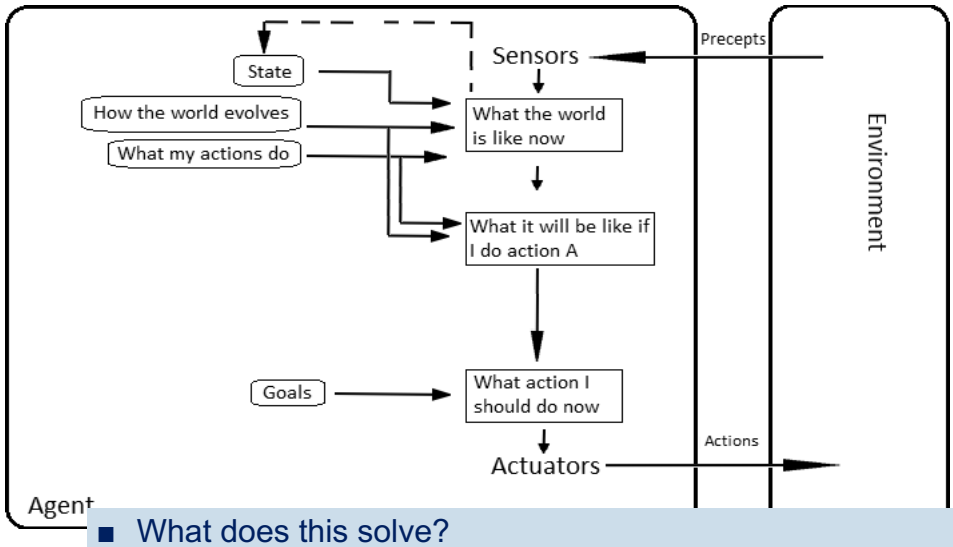
Videos from JHU

Model-Based Reflex Agents



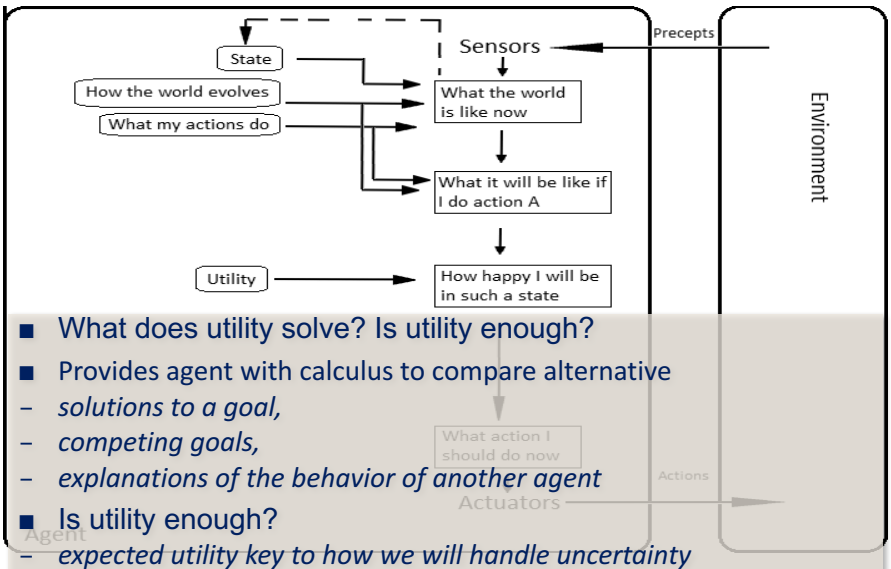
- How is this more capable/robust then Simple Reflex Agent?
- What are its limitations?

Goal Based Agent



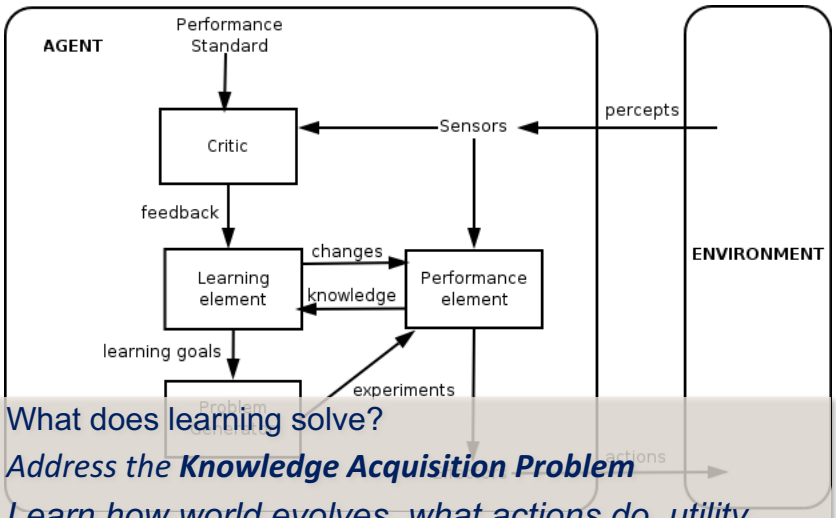
- What does this solve?
- What may be lacking?

Utility Based Agent



- What does utility solve? Is utility enough?
- Provides agent with calculus to compare alternative
  - *solutions to a goal,*
  - *competing goals,*
  - *explanations of the behavior of another agent*
- Is utility enough?
  - *expected utility key to how we will handle uncertainty*

Learning Agent



- What does learning solve?
  - *Address the **Knowledge Acquisition Problem***
  - *Learn how world evolves, what actions do, utility value of actions, optimal plans or action sequences*

Recap: What makes problems hard for an agent?

Uncertainty in various forms

- About the way world is now
  - Perceptions are often ncomplete/partial, noisy
- What the effects of action and plans of actions are
  - E.g., Driving on ice
  - In a world with other drivers/pedestrians
- What the future foretells
  - about the way the world will be, including actions of other agents

Tractable Decision-making

- Finding good solutions in large search spaces
- Finding good representations
- Determining one's goals in an uncertain world

This Course

This course is largely about problem solving in increasingly uncertain environments and agents with more complex tasks/goals in those environments

And the more sophisticated approaches to knowledge representation and agent design that are needed to be effective in those domains

Piazza

Please register on

- http://piazza.com/northeastern/fall2017/cs5100

Course overview

Announcements

Homework will be submitted using Blackboard

In class assignments will be on paper