## Introduction to Artificial Intelligence

- AI ‘food for thought’
- Agent framework
- Example Systems



## What is AI?

- What do you think?
  - Make programs that behave like the brain behaves?
  - Make programs that behave like the brain SHOULD behave?

## Which of these exhibit intelligence?

- You beat someone at chess.
- You prove a mathematical theorem using a set of known axioms.
- You need to buy some supplies, meet three different colleagues, return books to the library, and exercise. You plan your day in such a way that everything is achieved in an efficient manner.
- You are a lawyer who is asked to defend someone. You recall three similar cases in which the defendant was guilty, and you turn down the potential client.
- Someone taps a rhythm, and you are able to beat along with it and to continue it even after it stops.
- You are some sort of primitive invertebrate. You know nothing about how to move about in your world, only that you need to find food and keep from bumping into walls. After lots of reinforcement and punishment, you get around just fine.
- A stranger passing you on the street notices your watch and asks, “Can you tell me the time?” You say, “It is 3:00.”
- You are told to find a large Phillips screwdriver in a cluttered workroom. You enter the room (you have never been there before), search without falling over objects, and eventually find the screwdriver.
- You are a six-month-old infant. You can produce sounds with your vocal organs, and you can hear speech sounds around you, but you do not know how to make the sounds you are hearing. In the next year, you figure out what the sounds of your parents' language are and how to make them.
- You are a one-year-old child learning Arabic. You hear strings of sounds and figure out that they are associated with particular meanings in the world. Within two years, you learn how to segment the strings into meaningful parts and produce your own words and sentences.

3

## Why study AI?

- AI makes computers more useful
- Computers with intelligence would have a huge impact on civilization
- AI is cited as “the field I would most like to be in” by scientists in other fields
- The computer is a good metaphor for talking and thinking about intelligence
- Turning theories of intelligence into working programs forces us to work out the details
- AI yields good results for computer science, including time sharing, interactive interpreters, automatic storage management, etc.
- Computers make good experimental subjects

4

## What is the definition of AI?

- Russell and Norvig characterize AI definitions into four categories:

| | |
|---|---|
| Systems that think like humans. | Systems that think rationally. |
| Systems that act like humans. | Systems that act rationally. |



## What is the definition of AI?

- **Bellman, 1978**
  "[The automation of] activities that we associate with human thinking, activities such as decision making, problem solving, learning" (Systems that think/act like humans.)

- **Charniak and McDermott, 1985**
  "The study of mental faculties through the use of computational models" (Systems that think rationally.)

- **Haugeland, 1985**
  "The exciting new effort to make computers think *machines with minds*, in the full and literal sense" (Systems that think like humans.)

- **Kurzweil, 1990**
  "The art of creating machines that perform functions that require intelligence when performed by people" (Systems that act like humans.)

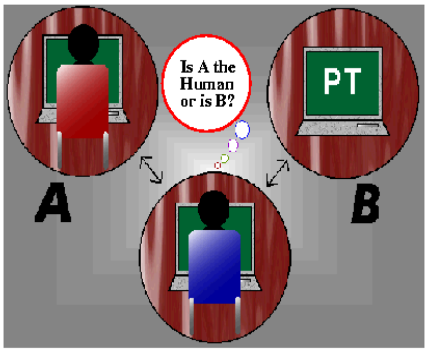
# What is the definition of AI?

- **Schalkoff, 1990**
  "A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Systems that act rationally.)
- **Rich and Knight, 1991**
  "The study of how to make computers do things at which, at the moment, people are better" (Systems that act like humans.)
- **Winston, 1992**
  "The study of the computations that make it possible to perceive, reason, and act" (Systems that think rationally.)
- **Luger and Stubblefield, 1993**
  "The branch of computer science that is concerned with the automation of intelligent behavior" (Systems that act rationally.)
- **Nilsson, 1998**
  "Many human mental activities such as writing computer programs, doing mathematics, engaging in common sense reasoning, understanding language, and even driving an automobile, are said to demand intelligence. We might say that [these systems] exhibit artificial intelligence" (Systems that act like humans.)
- **Nilsson, 2005**
  "Machines exhibiting true human-level intelligence should be able to do many of the things humans are able to do." (Systems that act like humans)

7

# Approach 1: Acting Humanly

- The Imitation Game (i.e. The Turing Test) (Turing, 1950)
  - "Three people:  a man (A), a woman (B), and an interrogator (C)"
  - "The interrogator stays in a room apart front the other two"
  - "Object of the game for the interrogator is to determine which of the other two is the man and which is the woman."
  - "In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten…"




8

# Approach 1: Acting Humanly

- The Imitation Game (i.e. The Turing Test) (Turing, 1950)
  - “We now ask the question, **‘What will happen when a machine takes the part of A in this game?’** Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‘Can machines think?’”

9

# How effective is this test?

- Agent must:
  - have command of language
  - have wide range of knowledge
  - be able to demonstrate human traits (humor, emotion)
  - be able to reason
  - be able to learn
- Total Turing Test: Computer Vision, Robotics
- Simulate understanding vs. true understand (John Searle)
- Loebner prize competition is modern version of Turing Test
  http://www.aisb.org.uk/events/loebner-prize
- AlexaPrize: https://developer.amazon.com/alexaprize
- Microsoft’s XiaoIce: https://arxiv.org/abs/1812.08989

10

# Chinese Room Argument

- Imagine you are sitting in a room with a library of rule books, a bunch of blank exercise books, and a lot of writing utensils. Your only contact with the external world is through two slots in the wall labeled “input” and “output”. Occasionally, pieces of paper with Chinese characters come into your room through the “input” slot. Each time a piece of paper comes in through the input slot your task is to find the section in the rule books that matches the pattern of Chinese characters on the piece of paper. The rule book will tell you which pattern of characters to inscribe the appropriate pattern on a blank piece of paper. Once you have inscribed the appropriate pattern according to the rule book your task is simply to push it out the output slot.
- By the way, you don't understand Chinese, nor are you aware that the symbols that you are manipulating are Chinese symbols.
- In fact, the Chinese characters which you have been receiving as input have been questions about a story and the output you have been producing has been the appropriate, perhaps even "insightful," responses to the questions asked. Indeed, to the outside questioners your output has been so good that they are convinced that whoever (or whatever) has been producing the responses to their queries must be a native speaker of, or at least extremely fluent in, Chinese.

11

# The Chinese Room Argument

- The question that Searle asks is - Do you understand Chinese?
  - Searle says NO
  - What do you think?
  - Is this a refutation of the possibility of Artificial Intelligence?
  - The Systems Reply
    - The individual is just part of the overall system, which does understand Chinese
  - The Robot Reply
    - Put the same capabilities in a robot along with perceiving, walking, and any other relevant activities. Such a computer would seem to have genuine understanding and mental states.

12

## Approach 1: Acting Humanly

- Problem:
  - “The quest for ‘artificial flight’ succeeded when the Wright brothers and others stopped imitating birds and started using wind tunnels and learning about aerodynamics.  Aeronautical engineering texts do not define the goal of their field as making ‘machines that fly so exactly like pigeons that they can fool even other pigeons” (R&N, 2010).



## Approach 2: Thinking Humanly

- OK…so how do humans think?
- Ways to capture all that:
  - Introspection
  - Psychological experiments
  - Physical brain imaging and measurement

## Approach 2: Thinking Humanly

- Cognitive science (F&M, 2008)
  - Brings together psychology, neuroscience, and computer science
  - Creates “computational theory of mind”
    - Maps cognitive processes to operational models
  - Initial emphasis = Clear segregation of mental processes
  - Over time = Mental processes influenced by (and influence) behaviors and environment
    - Enactive Perception



## Approach 2: Thinking Humanly

- Problems:
  - Requires knowledge of brain function
  - What level of abstraction? Knowledge, circuitry, chemical?
  - How can we validate this?

## Approach 3: Thinking Rationally

- Logic
  - Aristotle attempted to define irrefutable reasoning processes
  - What are correct arguments or thought processes?
  - Provided foundation for much of AI
- Problems:
  - It is not easy to formalize informal knowledge
  - Difference between solving a problem “formally” and in practice
  - Not all intelligent behavior controlled by logical

17

## Approach 4: Acting Rationally

- Act to achieve goals, given a set of beliefs
- Rational behavior is doing the “right thing”, or that thing expected to maximize goal achievement
- More amenable to scientific approaches
  - Thinking rationally is only a part of what makes up rationality
  - Humanness is hard to quantify
- Focus of most AI research

18

## Approach 4: Acting Rationally

- Problem:
  - Is achieving perfect rationality of action realistically achievable?
  - Is this something humans would be comfortable interacting with?



## History of AI

- **Philosophy**
  - 450 B.C., Socrates asked for an *algorithm* to distinguish pious from non-pious individuals.
  - Aristotle developed laws for reasoning (to deduce conclusions from premises).
- **Mathematics**
  - 1847, Boole introduced formal language for making logical inference.
- **Psychology**
  - 1879, Helmholtz studied human vision.
  - Cognitive psychology, study of human cognition.
- **Linguistics**
  - 1957, Skinner studied behaviorist approach to language learning.

# History of AI: Computer Science

- Computer-science-based AI is commonly agreed to have started with “the Dartmouth Conference” in 1956.
- Some of the attendees:
- John McCarthy
  - LISP, application of logic to reasoning
- Marvin Minsky
  - Popularized neural nets
  - Slots and frames
  - The Society of the Mind
- Claude Shannon
  - Computer checkers
  - Information theory
  - Open-loop 5-ball juggling
- Allen Newell and Herb Simon
  - General Problem Solver

21

# AI Questions

- Can we make something that is as intelligent as a human?
- Can we make something that is as intelligent as a bee?
- Can we make something that is evolutionary and self improving and autonomous and flexible?
- Can we save the AF $20 million a year with pattern recognition?
- Can we save a bank $50 million a year by automatic fraud detection?
- Can we start a new industry of handwriting recognition agents?

22

## State of the Art

- Which of the following can currently be done?
  - Play a decent game of table tennis
  - Drive along a curving mountain road? Or, how about downtown Cairo?
  - Play a decent game of bridge
  - Discover and prove a new mathematical theorem
  - Write an intentionally funny story
  - Play an instrument in a jazz band
  - Give competent legal advice in a specialized area of law

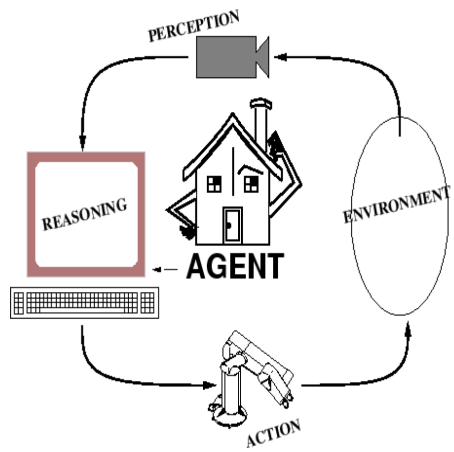


## Our AI System

- Weak vs. Strong AI
- The framework of an “Agent”
- Weak AI as:
  - Search
  - Knowledge Representation

## Components of an AI System

- An agent perceives its *environment* through *sensors* and acts on the environment through *actuators*.
- Human: sensors are eyes, ears, skin, nose, tongue; actuators (actuators) are hands, legs, mouth.
- Robot: sensors are cameras, sonar sensors, lasers, ladar, bump sensors, actuators are grippers, manipulators, motors.
- This course will focus largely on the **Reasoning** component.




25

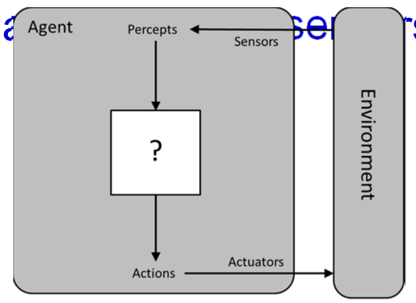
## Rational Agents

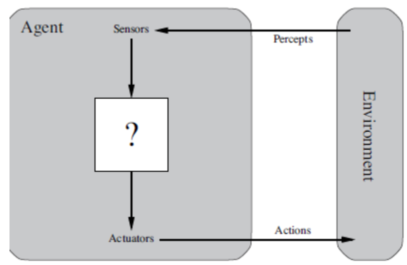
- A *rational agent* chooses the *action* which maximizes the *expected* value of the *performance measure* given the *percept* sequence to date
- Types of agents (increasing in generality)
  - Simple reflex agents
  - Reflex agents with state
  - Goal-based agents
  - Utility-based agents
  - Learning agents

26

# Agent Types

- How good is an agent?
- We need a *performance* measure
- In designing a rational agent, must know PEAS (performance measure, environment, a[...] se[...]rs)



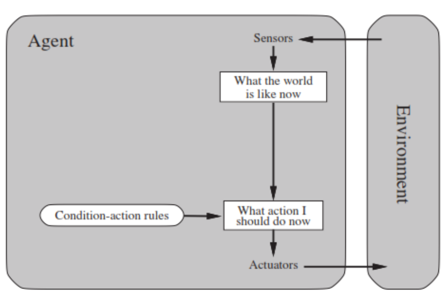






# Simple Reflex Agents

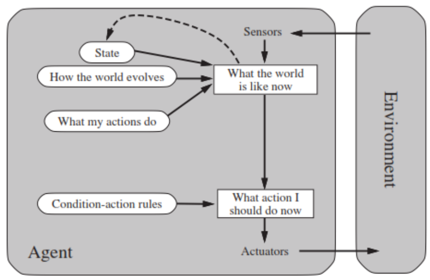
- Use simple "if then" rules
- Can be short sighted

```
SimpleReflexAgent(percept)
   state  = InterpretInput(percept)
   rule   = RuleMatch(state, rules)
   action = RuleAction(rule)
   return action
```



## Reflex Agent with State

- Provides memory
- Extends time horizon of agent
- Stored information may be inaccurate

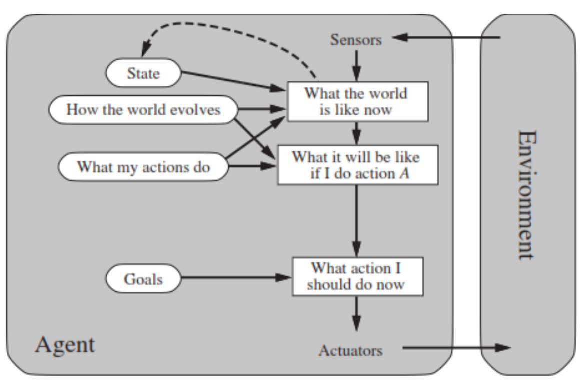



```
ReflexAgentWithState(percept)
   state  = UpdateState(state, percept)
   rule   = RuleMatch(state, rules)
   action = RuleAction(rule)
   state  = UpdateState(state, action)
   return action
```

29

## Goal Based Agent

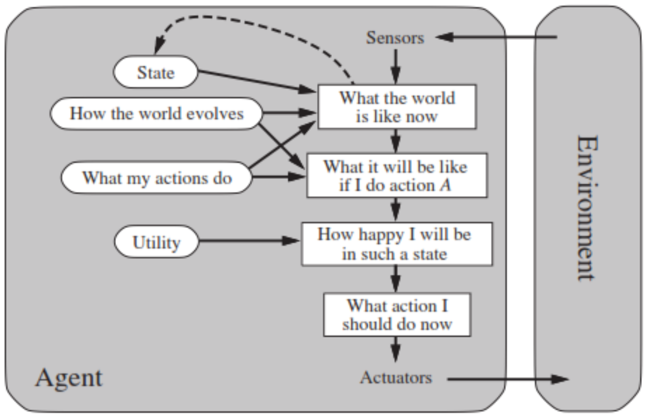
- Goal reflects desires of agent
- May project actions to see if consistent with goals
- Requires time, world may change during reasoning




30

# Utility Based Agent

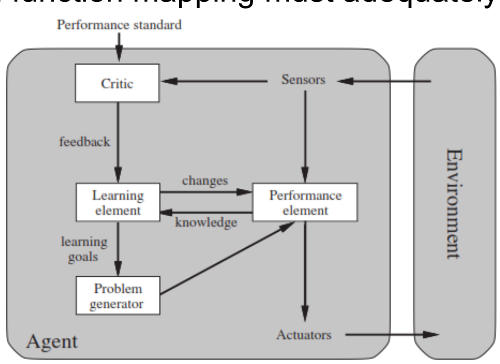
- Evaluation function to measure utility f(state) → value
- Useful for evaluating competing goals




31

# Learning Agent

- Learns a function mapping input to output
- Agent can operate in initially unknown environments
- Must have some mechanism for recognizing 'good' behavior, and function mapping must adequately express the domain.




32

# Environment Properties

| | | |
|---|---|---|
| Accessible | | Inaccessible |
| Deterministic | | Non-deterministic |
| Episodic | **VS** | Non-episodic |
| Static | | Dynamic |
| Discrete | | Continuous |

33

# Environment Examples

| Environment | Observable | Agents | Deterministic | Episodic | Static | Discrete |
|---|---|---|---|---|---|---|
| Chess with a clock | | | | | | |
| Poker | | | | | | |
| Backgammon | | | | | | |
| Taxi Driving | | | | | | |
| Medical Diagnosis | | | | | | |
| Image Analysis | | | | | | |
| Robotic part-picking | | | | | | |
| Refinery control | | | | | | |
| Interactive English tutor | | | | | | |

34

## Example AI System

- **Robotic Vehicles (DARPA Grand Challenge ⇒ Google Car)**
- Description:
  - Perception: cameras, lidar, radar, GPS, dead-reckoning
  - Reasoning: Markov model induction, A* searching, Bayes classification, behavior based systems
  - Action: wheeled Akerman actuation

http://http://code.google.com/p/cornell-urban-challenge/

35

## Example AI System

- Pathfinder Medical Diagnosis System
- Description:
  - Perception: Input symptoms and test results
  - Reasoning: Bayesian networks, Monte-Carlo simulations
  - Action: Output diagnoses and further test suggestions

http://research.microsoft.com/~horvitz

36

# Example AI System

- IBM Watson
- Description:
  - Perception: 200 million documents
    Natural Language Processing
  - Reasoning: Multiple Pattern matching and search algorithm
  - Action: Report highest ranked result



# Example AI System

- TDGammon World Champion Backgammon Player
- Description:
  - Perception: Keyboard input
  - Reasoning: Reinforcement learning, neural networks, Minimax Search
  - Action: Graphical output showing dice roll and piece movement

http://forum.swarthmore.edu/~jay/learn-game/systems/td-gammon.html

## Example AI System

- Alpha Go
- Description:
  - Perception: Keyboard input
  - Reasoning: reinforcement learning of two very large neural networks and Monte-Carlo Tree Search
  - Action: Graphical output showing stone placement



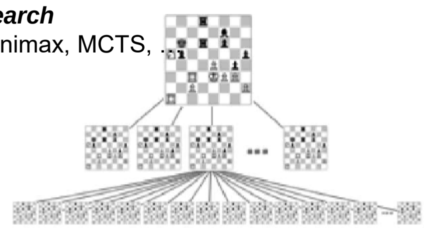
## What do weak AI Systems look like in practice?

Percept

***Knowledge Representation***
A state is a set of board features:
piece count, piece location, threat count

***Search***
Minimax, MCTS, ...

Action



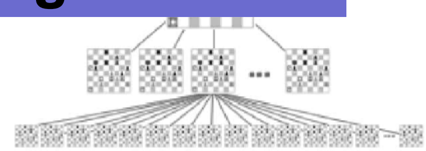
# But AlphaGo…

Percept

***Knowledge Representation***
Board + features into a deep CNN

Note: Learning the CNNs is it's own agent. KR is a sum of weights
hat are then
hill climbing

**Theme: Agents Inside of Agents**



41

# Recent Applied AI

- Billiards robot
- Throwing robot
- Soccer-playing robots
- Radiotherapy treatment planning
- Juggling robot
- Credit card fraud detection
- Lymphatic system diagnoses
- Mars rover
- Sky survey galaxy data analysis
- Train maintenance and route scheduling

42

# Subcomponents of Reasoning

- Search and Knowledge Representation
  - General problem solving
  - Planning
  - Machine learning
  - Uncertainty reasoning
  - Computer vision
  - Robotics

43

# Review

- Definitions of AI
- Definitions of Agents
- Example Systems

- For the remainder of the semester:
  - Search
  - Knowledge Representation
  - Introduction to algorithms in which search and knowledge representation combine to generate good solutions for many real world problems.

- Next Class:
  - Concepts and Algorithms for Search

44