CSE 3401: Intro to Artificial Intelligence & Logic Programming Introduction

- Required Readings: Russell & Norvig (R & N) Chapters 1 & 2.
- Optional readings: Poole & Mackworth (P & M) Chapters 1 & 2.
- Lecture slides adapted from those of Fahiem Bacchus.

What is Artificial Intelligence?

- What is **AI**?
- What is **intelligence**?
- What features/abilities do humans (animals? animate objects?) have that you think are indicative or characteristic of intelligence?

Webster says: a. the capacity to acquire and apply knowledge. b.the faculty of thought and reason. …

Alternate Definitions (Russell + Norvig)

	Like humans	Not necessarily like humans
Think	Systems that think like humans	Systems that think rationally
Act	Systems that act like humans	Systems that act rationally

Human intelligence

- Is imitating humans the goal?
- Pros?
- Cons?

Human intelligence

- The Turing Test:
  - A human interrogator. Communicates with a hidden subject that is either a computer system or a human. If the human interrogator cannot reliably decide whether or not the subject is a computer, the computer is said to have passed the Turing test.

Human intelligence

- Turing provided some very persuasive arguments that a system passing the Turing test is intelligent.
- But too much emphasis on deception.
- Moreover, the test does not provide much traction on the question of how to actually build an intelligent system.

Human intelligence

- In general there are various reasons why trying to mimic humans might not be the best approach to AI.
  - Computers and Humans have a very different architecture with quite different abilities.
    - Numerical computations
    - Visual and sensory processing
    - Massive and slow parallel vs. fast serial

Human intelligence

- But more importantly, we know little about how the human brain performs its higher level processes. Hence, this point of view provides very little information from which a scientific understanding of these processes can be built.
- However, Neuroscience has been very influential in some areas of AI. For example, in robotic sensing, computer vision, etc.

Rationality

- The alternative approach relies on the notion of **rationality**.
- Typically this is a precise mathematical notion of what it means to *do the right thing* in any particular circumstance. Provides
  - A precise mechanism for analyzing and understanding the properties of this ideal behavior we are trying to achieve.
  - A precise benchmark against which we can measure the behavior the systems we build.

Rationality

- Mathematical characterizations of rationality have come from diverse areas like logic (laws of thought) and economics (utility theory, how best to act under uncertainty, game theory, how self-interested agents interact).
- There is no universal agreement about which notion of rationality is best, but since these notions are precise we can study them and give exact characterizations of their properties, good and bad.
- We'll focus on acting rationally
  - this has implications for thinking/reasoning

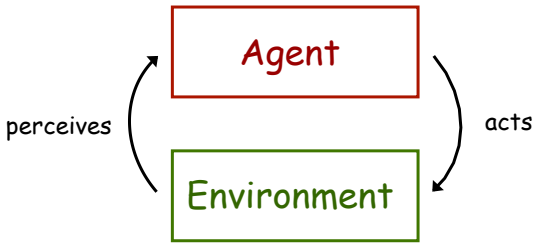
Computational Intelligence

- *AI tries to understand and model intelligence as a computational process.*
- Thus we try to construct systems whose computation achieves or approximates the desired notion of rationality.
- Hence AI is part of Computer Science.
  - There are other areas interested in the study of intelligence, e.g., cognitive science, which focuses on human intelligence. Such areas are very related, but their central focus tends to be different.

Agency

- It is also useful to think of intelligent systems as being **agents**, either:
  - with their own goals
  - or that act on behalf of someone (a "user")
- An *agent* is an entity that exists in an *environment* and that *acts* on that environment based on its *perceptions* of the environment
- An *intelligent agent* acts to further its own interests (or those of a user).
- An *autonomous agent* can make decisions without the user's intervention, possibly based on what it has learned

Agent Schematic (I)

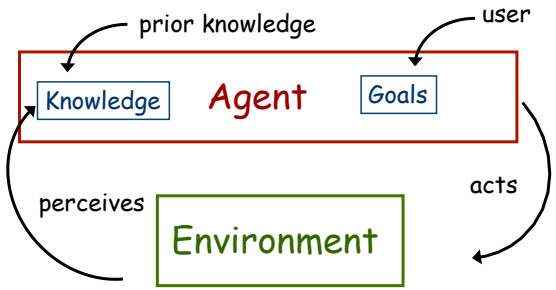


- This diagram ignores the internal structure of the agent.

Types of Agents

- **Simple reflex agents**: apply simple condition-action rules to decide next action based on current percepts
- Model-based reflex agents: maintain a model of the world, apply rules to decide next action based on current world model
- **Goal-based agents**: decide next action based on current model of world state and current **goal(s)**; may do planning; more flexible!

Agent Schematic (II)



- Supports more flexible interaction with the environment, ability to modify one's goals, knowledge that is applied flexibly to different situations.

Types of Agents (cont.)

- **Utility-based agents**: choose actions to maximize their expected **utility** in uncertain worlds

For all types of agents, can add a learning mechanism: explore space of possible rules/actions/models, evaluate performance, and modify agent to improve

Types of Environments

- Fully Observable vs. Partially Observable
- Deterministic vs. Stochastic
- Episodic vs. Sequential
- Static vs. Dynamic
- Discrete vs. Continuous
- Single-Agent or Multi-Agent
- Known Dynamics vs. Unknown Dynamics

See R & N Sec. 2.3 for details

Agent Architectures

- Agents may have a more complex architecture than what we have seen so far
- In particular, embodied agents such as robots typically have complex hierarchical control architectures with multiple layers, with low-level layers handling local motion and collision avoidance, mid-level layers handling path planning and following, and high-level layers handling task planning

See P & M Sec 2.3 for details.

Degrees of Intelligence

- Building an intelligent system as capable as humans remains an elusive goal.
- However, systems have been built which exhibit various specialized forms/degrees of intelligence.
- Formalisms and algorithmic ideas have been identified as being useful in the construction of these "intelligent" systems.
- Together these formalisms and algorithms form the foundation of our attempt to understand intelligence as a computational process.
- *In this course we will study some of these formalisms and see how they can be used to achieve various degrees of intelligence.*

AI Successes

- In 1997 IBM's Deep Blue beat chess world champion
- In 1999, NASA Remote Agent used AI planning to control a spacecraft
- In 2005 Stanford team won DARPA Grand Challenge 132mi race in desert
- In 2011, IBM's Watson beat the top Jeopardy winners
- In 2016, Google DeepMind AlphaGo beat decade's top player
- Many successes in speech recognition, machine translation, robotics, scheduling, spam fighting

Reasons for Recent Progress

- Better hardware
- AI techniques are improving, especially:
  - Search methods and heuristics
  - Improved representations
  - Machine learning, large corpuses

Subareas of AI

- Perception: vision, speech understanding, etc.
- Robotics
- Natural language understanding
- Machine learning
- Reasoning and decision making (our focus)
  - **Knowledge representation**
  - **Reasoning** *(logical, probabilistic)*
  - **Decision making** *(search, planning, decision theory)*

Prospects for AI

- Recent progress has been rapid
- Concerns about the risks of developing AI
- Are current learning-based AI systems really intelligent?
  - Winograd Schema Challenge, e.g.

    The city councilmen refused the demonstrators a permit because they **feared** violence. Who feared violence?

    vs

    The city councilmen refused the demonstrators a permit because they **advocated** violence. Who advocated violence?

Some Interesting & Entertaining Videos

- James May's Big Idea Man-Machine episode where he meets Honda's Asimo robot programmed so it can learn to recognize objects http://www.youtube.com/watch?v=QfPkHU_36Cs
- Google's self-driving car https://www.youtube.com/watch?v=TsaES--OTzM
- Google self-driving care Waymo (recent) https://www.youtube.com/watch?v=uHbMt6WDhQ8
- Google Deep Mind AlphaGo win http://www.theguardian.com/technology/video/2016/mar/09/alphago-computer-beats-go-champion-video