Penn CIS 5210 - Chapter 1

Jonathon Delemos - Dr. Chris Callison Burch

June 3, 2025

This course investigates algorithms to implement resource-limited knowledge-based agents which sense and act in the world. Topics include, search, machine learning, probabilistic reasoning, natural language processing, knowledge representation and logic. After a brief introduction to the language, programming assignments will be in Python. — Description of CIS 4210/5210 in course catalog

1 Section One: Introduction

1.1 What is AI?

```
The field of artificial intelligence (AI) is concerned
with not just understanding intelligence, but also
building intelligent entities|machines that compute
how to act effectively and safely in a variety of novel
situations.  We can define intelligence as the better
part of rationality-loosely speaking, the ability to
do the right thing at the right time.
```

1.2 Acting Humanly: The Turing Test

```
Natural Language Processing is the computers ability
to communicate successfully in a humnan language.
Knowledge Representation is the ability to store what
it knows or hears.
Automated Reasoning to answer questions and to draw
new conclusions.
Machine Learning to adapt to new circumstances and
to detect and extrapolate patterns.
Computer Vision to perceive objects and scenes.
Robotics to move and manipulate objects.
```

1.3 Thinking Humanly: Cognitive Modeling

Introspection is the process of examining one's own thoughts and feelings. **Psychological Experiments** is observing a person in action.
Brain Imaging is observing the brain in action
Cognitive Science brings together the field Psychology and computer models to construct precise and testable theories of the human mind.

1.4 Thinking Rationally: The Laws of Thought Approach

Logic is a pattern for argument that is irrefutable. If we don't have something that is absolutely certain, what we can do is fill that gap with probability analysis.

1.5 Acting Rationally: The Rational Agent Approach

Agent is something that acts. Of course all computer programs do something, but agents are expected to operate autonmously, perceive their environment, etc.
Rational Agent is a model that acts to achieve the best outcome, or best expected outcome. Methods are based on probability-AI has focused on the study and construction of agents that do the right thing.
Using rational agents to accomplish these best possible outcomes is referred to as the **Standard Model.**
Limited Rationality is the act of acting appropriately when there isn't enough time to calculate all the possible computations.

1.6 Beneficial Machines

The goal of this section is to define the multiple parameters that need to be weighed when making an agent. For example, in a self-driving car, one might the objective is simply to make it to your destination. However, you might be concerned with injury to passengers, annoying other drivers, following the rules of the road, etc. This is where **Value Alignment Problems** come in. The values of the human must be aligned with the values of the agent. The result of not properly placing limits on value assignments is a computer doing seemingly irrational things to reach it's sole goal.

1.7   The Foundations of Artifical Intelligence

Dualism is the counter argument against that if the mind is governed by physical laws, we have no more free will than a rock falling down. This claims our mind is exempt from physical laws. There are loads of other philosophies here. Most of which are very interesting. Refer to pages 24-25. Moving forward.
General Problem Solver is another term for greedy algorithm.
Example Hello, World! This is an example for Sai.

1.8   Mathematics

Tractability is the word to describe problems at approach incomputability. This section deals with basic definitions of NP-complete, probability, and incompleteness.
Markov Decision Processses is the term used to describe sequential decision problems. This is when the first decision doesn't produce the best result.