32002: Artificial intelligence **Spring 2014**

Lecture 2: January 02, 2014

Lecturer: K.R. Chowdhary *: Professor of CS (GF)*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

2.1 Goals and Roots of AI

2.1.1 Goals of AI

The *scientific goal* is to determine theories about knowledge representation, learning, rule-based systems, and search that explain various sorts of intelligence.

The *engineering goal* is to solve real world problems using AI techniques such as knowledge representation, learning, rule systems, search, and so on.

2.1.2 Roots of AI

Artificial Intelligence has identifiable roots in a number of older disciplines, particularly:

- Philosophy
- Logic/Mathematics
- Computation
- Psychology/Cognitive Science
- Biology/Neuroscience
- Evolution

Philosophy

350 BC: *Aristotle* formulated different styles of deductive reasoning, which could mechanically generate conclusions from initial premises, e.g. *Modus Ponens:*

$$If \text{ A } \rightarrow B \text{ and } A$$
$$then \text{ B}$$

1596 1650: Rene Descartes gave a concept of mind-body dualism, as per that part of the mind is exempt from physical laws. Otherwise, how do we have free will?

1646 1716: *Wilhelm Leibnitz* was one of the first to take the materialist position, which holds that the mind operates by ordinary physical processes this has the implication that mental processes can potentially be carried out by machines.

Logic and Mathematics

1815 - 1864: *George Boole* introduced his formal language for making logical inference in 1847 - *Boolean Algebra*.

1848 - 1925: *Gottlob Frege* introduced a logic that is essentially the first-order logic that today forms the most common knowledge representation system.

1906 - 1978: *Kurt Godel* showed in 1931 that there are limits to what logic can do. His *Incompleteness Theorem* showed that in any formal logic powerful enough to describe the properties of natural numbers, there are true statements whose truth cannot be established by any algorithm.

Computation

1912 - 1954: *Alan M. Turing* tried to characterize exactly which functions are capable of being computed. The Church-Turing thesis states that a Turing machine is capable of computing any computable function, is generally accepted as providing a sufficient definition. Turing also showed that there were some functions which no Turing machine can compute (e.g. Halting Problem).

1903 - 1957: *John von Neumann* proposed the von Neuman architecture which allows a description of computation that is independent of the particular realization of the computer.

Psychology and Cognitive Science

Modern Psychology and Cognitive Science are concerned to how mind functions, how we behave, and how our brains process the information.

Language is an important part of human intelligence. Much of the early work on knowledge representation was tied to language, and was produced by research into linguistics.

It makes sense to explore the properties of artificial systems (computer models/simulations) to test our hypotheses concerning human systems.

Biology and Neuroscience

Our brains (which give rise to our intelligence) are made up of tens of billions of neurons, each connected to hundreds or thousands of other neurons. Each neuron is a simple processing device (e.g. just firing or not

firing depending on the total amount of activity feeding into it).

The field of Neural Networks attempts to build artificial systems based on simplified networks of simplified artificial neurons. The aim is to build powerful AI systems, as well as models of various human abilities.

The major difference between neurons and human reasoning is that neural networks work at a *sub-symbolic level*, whereas much of conscious human reasoning appears to operates at a symbolic level.

Evolution

Advantage humans have over current machines/computers is that they have a long evolutionary history. As per the evolution theory, the fitter individuals will naturally tend to live longer and produce more children, and hence after many generations a population will automatically emerge with good innate properties.

This has resulted in brains that have much structure, or even knowledge, built in at birth. Computers are finally becoming powerful enough that we can simulate evolution and *evolve* good AI systems. We can now even evolve systems (e.g. neural networks) so that they are good at learning.

A related field called *genetic programming* has had some success in evolving programs, rather than programming them by hand.

2.2 Artificial Consciousness

A discussion of artificial consciousness gives rise to several philosophical issues:

- Can computers think or do they just calculate?
- Is consciousness a human prerogative?
- Does consciousness depend on the material that comprises the human brain, or
- can computer hardware replicate consciousness?

2.3 Common Techniques used in AI

Representation: Knowledge needs to be represented somehow - perhaps as a series of *if-then rules*, as a *frame based* system, as a *semantic network*, or in the connection weights of an *artificial neural network*.

Learning: Automatically building up knowledge from the environment such as acquiring the rules for a rule based expert system, or determining the appropriate connection weights in an artificial neural network.

Rules: These could be explicitly built into an expert system by a knowledge engineer, or implicit in the connection weights learn by a neural network.

Search: This can take many forms perhaps searching for a sequence of states that leads quickly to a problem solution, or searching for a good set of connection weights for a neural network by minimizing a fitness function.

2.4 Sub-fields of AI

AI now consists many sub-fields, using a variety of techniques, such as:

- *Neural Networks* brain modeling, time series prediction, classification
- *Evolutionary Computation* genetic algorithms, genetic programming
- *Vision* object recognition, image understanding
- *Robotics* intelligent control, autonomous exploration
- *Expert Systems* decision support systems, teaching systems
- *Speech Processing* speech recognition and production
- *Natural Language Processing* machine translation
- *Planning* scheduling, game playing
- *Machine Learning* decision tree learning, version space learning

References

[GFL09] D. George F. Ludger, "Artificial Intelligence - Structures and Strategies for Complex Problem Solving," *5th Edition, Pearson Education, India*, 2009, Chapter 1.