
Assignment-3 : AI and its Foundations

Harshvardhan Patidar
Department of Artificial Intelligence
Indian Institute of Technology Hyderabad
ai24btech11015@iith.ac.in

The section 1.1 gives us a basic idea of what AI actually is and how it is planned to work. The focus of section 1.2 is on explaining how the foundation for AI was laid by the different fields and the advancements in them. [1]

1 Section 1.1

Many people prefer AI to think and act like humans, while others want it to be rational. There are also diverging views on how intelligence should be defined, some focus on how machines should think, that is their internal thought process, while others give importance to how it should behave.

From these two different sets of views, i.e. humanly vs rationally and thought vs behavior, we get four different ways to define AI, all of which have had some research programs. For example, if you want AI to act humanly, you need to study about human behavior and psychology, but for rational behavior, you need to use math, statistics and control theory.

1. **Acting Humanly** : The objective of this approach is to make AI capable of thinking like humans. Alan Turing had proposed the turing test which mainly tested if the machine could think. A computer could only pass this test if it is able to generate responses which a human checker cant distinguish if it is written by a machine or a human. This focuses on behavior rather than internal thinking of the machine.
2. **Thinking Humanly** : This focuses on making AI think like humans. For this they tried to model how human brain works. Cognitive science was used to achieve this. The goal is to replicate human thinking processes by machines.
3. **Thinking Rationally** : This approach talks about developing AI to think rationally, that is thinking by using logic and making use of the concept of probability in cases where there is some uncertainty. The goal is to make machines think in a perfect, logical way.
4. **Acting Rationally** : The idea of this is to not make machines think like humans, but should act in a way such that it achieves the best outcome or the best expected outcome. Machines should achieve their goals in the most efficient manner, no matter if they think like humans or not.

The section also talks about if a machine which is designed to achieve a specific object is actually valuable or not. There might be cases where just achieving the goal in real life is not practically possible or ethically viable. Instead we need to set the objective of the machine in such a way that it both achieves the goal, and also doesn't act in an unacceptable manner.

2 Section 1.2

This section gives us a brief history of all those fields which contributed in some way or other in the development of AI.

1. **Philosophy** : People had been thinking about reasoning and logic for long, even before computer were invented. The philosophers had been trying to make an artificial machine that could actually work on rationality and logic.
2. **Mathematics** : This discipline is actually very fundamental for AI as logic and probability are very important for AI. The incompleteness theorem shows that there are statements that have no proof.
3. **Economics** : The models built to solve economical problems, like how to maximize profits or minimise losses have also contributed to development of AI. It sometimes require to take some decision making a trade-off, which it can optimize by knowledge of such models.
4. **Neuroscience** : We now also have some data on mapping between areas of brain and parts of body they control or receive input from. This mapping data helps to develop AI models capable of optogenetics, i.e. to control the activity of neurons by using light. The development of brain-machine interfaces promises to restore function to disable individuals.
5. **Psychology** : In this field people focus to get information about what someone's thought process would be like while performing some task. This data can later be used to make AI think in a similar manner to achieve human-like results.
6. **Computer Engineering** : This is a very obvious foundation of AI. All the operating systems, programming languages, and tools needed to develop AI programs were because of the advancements in computer science and engineering. The LLM's nowadays require a lot of compute power which is supplied by this field.
7. **Control Theory and Cybernetics** : This field tries to find the answer to the question, "How can artifacts operate under their own control?". This field helps AI in adapting and modifying itself according to the place its working. It focuses to make AI homeostatic, which has feedback loops to achieve a stable behavior.
8. **Linguistics** : This field focuses on language learning. To understand the language, understanding the structure of the sentence is not enough. An understanding of the subject is also necessary to understand the language.

References

- [1] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, Hoboken, NJ, 4th edition, 2020.