AI

and its implications

Vorwissenschaftliche Arbeit
vorgelegt von Adriel Ondas

BetreuerIn: Manuel Menzinger

Schuljahr 2020-2022

Contents

1 Introduction

1.1 Why is this topic important?

In the last few years, AI has been the topic of many discussions, often without a very realistic view on AI.

While many see AI as these super-strong real-life robots, many do not get just how broad one can define AI in a sense of intelligent agents, and just how close it is to all of our lives, even through we are currently not using as versatile agents, as Hollywood would like us to believe.

The most obvious examples of use of AI are found around computers, more closely around social media, and almost everywhere those mysterious "Algorithms" are used, for example to help advertisers advertise [1] or to sort content [2].

1.2 Questions this thesis seeks to answer:

The Author wants to bring this topic in broader view of the public, and wants to answer the questions of:

- What AI is (in a very abstract sense).
- How AI affects society.
- What previously impossible to solve tasks AI can help to solve.
- What new problems arise.
- How we can build aligned and save AI, to help mitigate those problems.

1.3 How was this thesis made?

Using mainly content found in the book "Artificial Intelligence, an Introduction" [3], the author would like to answer these questions, using other works to back it up in very specific topics not discussed as extensively in this book.

2 What is AI?

2.1 Definition

First it is necessary defining what is meant by AI in this thesis, as there is a very broad range of definitions.

To get an overview it is very helpful to view an AI as an entity, a sort of black box. This construct can be called an ”Agent”. This thesis will shed some light into these black boxes.

2.1.1 Agents

This agent has to somehow get a sort of input, like sound from a microphone, click position on a website, ... Then, it has to process this input, and lastly it has to have some sort of output, like driving somewhere, or showing different things on a screen, ...

The input can be called ”Percept”, and the output ”Action”. Image

These Agents vary widely in their implementation and function, and can range from bots on the web, to the microchip in your smart stove or Roomba. Agents are this diverse because they have to act in very different environments, for example a computer, a room where a robot is driving around, a map, a game, ...

This thesis however, will mostly focus on the internal structure of the agents, which can be similar, even in very different environments. The goal of this thesis is an overview of these different internal structures.

As the concept of an agent is so openly defined, it helps to narrow it down further. For example a random number generator has all the prerequisites for an agent, but this thesis will focus only on intelligent agents.

2.1.2 Intelligence

For this, we must define intelligence first.

Commonly, intelligence is viewed as ”the ability to learn, understand and think logically about things; the ability to do this well”

Important to stress here are two separate concepts: Learning, and thinking logically. In AI, these concepts can be separated, as to say there is a learning, and an applied phase, where the logical thinking comes into play. This is called supervised learning. There are also models combining these two phases, where learning takes place while also applying the previously learned. This is then called unsupervised learning.

2.2 Conclusion

This restriction of only intelligent agents is still not very narrow, as intelligent agents can still be a lot of things: A calculator has input, processes it intelligently, and shows the result, as does a player of a game, ...

2.3 Usage

There are very different models for AI, as will be explained in more detail in the next chapters, but of course they all have their own little problems but also advantages. So in practical use, different models are often chained together, to leverage the advantages of each sort of agent, to create an agent better than the individual agents.

3 Categorization

Simple map Description

3.1 symbolic vs subsymbolic

A first big distinction is to make between symbolic and subsymbolic AI. The difference is in how knowledge about the environment is stored and used. In symbolic AI, knowledge is coded into an agent, while in subsymbolic AI knowledge is generally learned by an agent through observations (=data). These observations can either be pre-recorded or acquired by trial and error. (The structure is still hard coded through)

Which sort of AI is better to use really depends on the circumstances: Symbolic AI for a complex subject normally requires more time to program in comparison, and needs a programmer well versed in the subject. Subsymbolic AI on the other hand depends on Data being available. Of course, it still is complex to program such a system, but the programmer does not have to know the subject at hand. Another aspect to keep in mind is the accuracy: Symbolic AI can have a 100 percent accuracy rate when programmed correctly, as it is just logical statements and numbers chained together by some operators, while with a subsymbolic AI it is normally very hard to get a 100 percent accurate model, and most models are capped at some (high) percentage of accuracy. This is in part because of incorrect/incomplete data, but also because of considerations like time and resource constrains while training. Subsymbolic AI can still be more accurate in certain domains: When there is no one correct answer but good data for example. As it is often the case, there is a cost-accuracy payoff, and whether subsymbolic or symbolic AI is better suited really depends on the use case. Often they are combined to achieve better results. This work will focus on subsymbolic AI.

3.1.1 Data

The requirement for data in subsymbolic AI is also the reason why data collection is getting increasingly important in our time.

4 Examples

4.1 Decision trees

4.1.1 Definition

A decision tree is a tree consisting of nodes, in which each node answers a yes/no question.

A decision tree can also be made symbolically, but to create such a tree subsymbolically, data is needed, which is used to shape the tree accordingly. The hard part is figuring out what question the next node will use for its split.

One method of measuring the usefulness of a question is the Gini method. It is calculated as follows: Gini impurity $= 1 - (yes/n)^2 - (no/n)^2$ The smaller this impurity, the less mixed the output is. Which is what we want, as less mixed means the tree is surer of its output. So to choose the next node to add to a tree, we measure the GINI impurity of all possible parameters, and choose the one with the smallest impurity. We repeat this process until we only have few samples left in each leaf.

Once created, to find an answer, one must walk down the tree, following the path the nodes lead you, until an end (a leaf) is met.

4.1.2 Usage

There are two advantages of decision trees in comparison to other forms of AI. One is the ease of understanding it, as we can retrace decisions with relative ease, which is very important in applications where trust in an algorithm is an issue. So for example in an environment where job applications or prison sentences are guided by algorithms, it might be required to be able to check the algorithm for, for example, discriminatory behavior.

Another benefit decision trees offer is the low computational complexity, in both training and application. Of course, they can grow arbitrarily large, but with the right techniques they can be brought back to reasonable sizes, which just capture the essentials.

Of course, they are not perfect: For one they require a lot of data to be conclusive, and as they are one of the simpler forms of AI, they need to get quite big to capture complex concepts.

4.2 Neural nets

4.2.1 Definition

List of Figures

Bibliography

[1] facebook. Werbung auf instagram | instagram business.

[2] google. So funktioniert die google-suche | suchalgorithmen.

[3] Stuart Russell and Peter Norvig. *Artificial Intelligence A Modern Approach*. Pearson Education, Inc., 3 edition, 2010.