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 helpful defining what AI is, so this chapter will start by defining AI. The field of AI is very broad, and defining this field is quite challenging, but also useful for this thesis.

To do this, there are a few prerequisites: First, the AI has to somehow get a sort of input, like sound from a microphone, click position on a website, ... Secondly, the AI 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, ...

In cite(modern approach), this whole construct is called an ”Agent”, the input is called ”Percept”, and the output ”Action”.

These agents can act in very different environments.
This definition is, perhaps not coincidentally, quite similar to how functions might be defined:
In mathematics, a function is a binary relation between two sets that associates each element of the first set to exactly one element of the second set.
In the case of an AI, the first set are the inputs, and the second set are the different outputs.
Important to note here is, that the input might also be what happened until now, not only what happens at this instance in time.

These Agents vary widely in their implementation and function, and can range from bots in the web, to your stove. In this thesis however, we will mostly focus on the internal structure of the agents, which can be similar, even in very different environments. In this thesis the author wants to create 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 I will focus only on intelligent agents.

Generally, intelligence is viewed as ”the ability to learn, understand and think in a logical way about things; the ability to do this well” Important to stress here are two seperate concepts: Learning, and thinking in a logical way. In AI these concepts are sometimes seperated, as to say there is a learning, and a applied phase, where the logical thinking comes into play.

But there are also models combining these two, where learning takes place while applying the previously learned.

Intelligent agents can still be a lot of things: A calculator has input, processes it intelligently

Data=Learning/Knowledge=programming symbolisch/subsymbolisch

Mixing

3 Examples

Q-table

decision tree

neural net back propagation

evolution

3.1 Definition

A decision tree is a form of AI, which is, in comparison, rather easy, and not that complex. It exports a tree consisting of nodes, in which each node answers a yes/no question.

To create such a tree, 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: 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.

3.2 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.

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.