Thoughts on „AI Debates & Watson on Jeopardy“

Intro to Philosophy
Yavor Stoychev

As a Computer Science student, I have taken a course in Artificial Intelligence, and I am familiar with both the „Good Old Fashioned Artificial Intelligence“ approach and the novel method of using statistics to build a model of an existing environment. I know about the giant leaps that computers have made in the last 50 years in both hardware and programming. But even though they can now accomplish tasks that many thought were too complex for the „big calculators“, computers are still far away from achieving real intelligence. And I doubt that they will ever get there.

First of all, I would like to explain what The „Good Old Fashioned Artificial Intelligence“ approach actually means. Imagine a parallel universe which is nothing more than a 9x9 Sudoku puzzle. In this universe, there exists only one being – a computer, which is supposed to solve the puzzle. The computer doesn't know that Sudoku is. It doesn't understand the meaning of winning, or losing. It doesn't grasp the concept of success, or failure. It is governed by a rule-book, which tells it what to do, and what not to do in certain cases. It is just a machine following a predefined set of rules. It is as intelligent as an elevator, that takes you to the third floor, if you press the right button.

This approach cannot lead to real intelligence, because is makes two very basic assumptions, which do not generally hold true in the real world. The first assumption is that the world is static. This is very much true in the case of the classic Sudoku, where the number of squares is fixed, but, if more squares start appearing from out of nowhere, the rule-book will be rendered obsolete, and then the computer in this Sudoku world will not have a clue what to do. The second assumption is, that the world is simple. Simple enough, that a set of unambiguous rules, addressing every possible situation, can be created. This assumption does not hold in the real world, since although creating such rules might be theoretically possible, it is very much unfeasible in practice.

The new statistical approach to Artificial Intelligence tries to address these issues. It is a huge step forward, as it allows the computer to actually build its own rule-book by observing the real world, following it to a certain extent, and changing it when it deems it necessary. This method is quite successful at simulating intelligence. In fact, I believe that, given the right programming and an appropriate amount of training, a computer using this approach might be able to pass the Turing test. Turing would argue that this is all there is to it, and the questions whether machines can think is irrelevant. He believes that we should use the “imitation game” to help us better understand “the intellectual capacities of man” and help us build better machines. (Turing 1950).

A computer might be “smart”. It might perform admirably at the “imitation game”. However, that does not mean it can think. If we look under the hood, we will not see an intelligent being. We will rather see a very expensive and complex statistics calculator at work. We will see a machine that can process syntax, but has no notion of semantics. We will see a machine, that can extract knowledge, but not meaning. It it still bound by the same limitations, that John Searle showed in “The Chinese Room“ thought experiment (Searle 1980). It does not care, it does not experience wonder, it does not have fundamental beliefs, reasonable or otherwise. And, most of all, it is bound by rules.

In his paper „What did the Watson Computer Do“, Stanley Fish argues that, humans do not follow rules blindly, but rather take a moment to evaluate how the rules might relate to a specific context. He also states that computers do not do that (Fish 2011). While this is certainly true for the „Good Old Fashioned Artificial Intelligence“ approach, it is something that the novel method of Artificial Intelligence allows for. However, as the Watson computer has shown, this approach is far from perfect either, as it made Watson give the wrong answer to a rather straightforward inquiry. The statistical approach has its merits, but it also has its limitations. It assumes that, given enough training, a computer can adapt its rule-book to every situation, and act accordingly. But how much training is “enough”? It obviously depends in the complexity of the real world, and our world is fairly complex. So, it is entirely possible, that using this approach, a computer will have to go through tens of thousands of years of training examples, before it can start delivering the right answers.

Why? Why is it so difficult for a computer to even play the “imitation game”? I think that the answer lies in the fact that a computer works very differently from the human brain. The human thought is something that is not constrained by any rules or guidelines. We, humans, do not treat rules as absolute. A computer, however, does. It can be programmed to derive its own changing rules from the surrounding environment, but, if you look deep enough into its programming, there are ALWAYS predefined rules that it must follow. It is bound by logic, which is something that the human mind can easily transcend. And I believe, that is the primary reason why we have intelligence, and a computer does not, and might never do.

Works Cited:

1. Searle, John. *"Minds, Brains, and Programs"*. Behavioral and Brain Sciences. 1980.
2. Fish, Stanley. *“What Did Watson the Computer Do?”*. The New York Times. 2011.
3. Turing, Alan *“Computing Machinery and Intelligence”*. 1950.