21st Century Knowledge Worker: the Centaur

Daniel Kiss

Introduction

The centaur is a well-known mythological creature, half-human half-horse. The most famous of centaurs was Chiron, the teacher of Asclepius, Theseus and Jason among others, a being known for his wisdom and his skill in medicine, music, hunting and prophecy.

Nowadays centaurs spend their time playing chess. In centaur chess a player is a team made up of a human player and a computer. Centaur chess, or advanced chess was popularized by Garri Kasparov, the world chess champion defeated by Deep Blue, the IBM chess computer. Centaurs tend to be stronger than pure human or pure computer players, sometimes a team of an average human and a computer achieving higher ratings than top grandmasters.

Centaur chess is just one example of human-artificial intelligence cooperation. Hybrid systems, where people work together with AI, are used in fraud detection and on the stock market. We can go further, and say that anybody with an internet connection is a centaur. We cooperate with AIs, when looking for a restaurant on Google, or when we are buying presents on Amazon.

So how can knowledge workers utilize these systems and the new ones introduced every day, to their full potential? We should stop looking at AI as a tool, and we should treat them as fellow knowledge workers. In Managing Oneself Drucker provides a framework to analyze a knowledge worker, so we do just that.

1. What are the strengths of AI?

Drucker writes, that to understand ones strengths and weaknesses one must perform a feedback analysis. We decide on a goal, look for historical evidence whether that goal has been reached, and evaluate how and why was that goal reached or not reached.

Full AI

The original goal of artificial intelligence research was to create a machine that can perform as well or better than a general human. The field of AI started from this

motivation in the 1950s, was an inspiration for HAL 9000 and other fictional computers, and there is still significant ongoing research towards this goal.

The lack of results are obvious: if there were anything coming even close to human intelligence, we would have already heard about it. Chatter bots, AIs that have been designed to hold humanlike conversations, while improved a lot since the 1950s, have not even come close to beating the Turing test, in other words were unable to fool a suspicious human into thinking that they are talking to an actual person.

However this researched showed the challenges of AI research. The three most important are commonsense, intractability and Moravec's paradox.

Every human being, or even animals holds a truly vast amount of organized and connected information about the environment it lives in, we call this information base commonsense knowledge. We can not only remember this information, but make deductions based on it, we call this skill commonsense reasoning. Translating commonsense knowledge to a formal language understood by computers is a difficult problem, and it is inevitable that a lot of information would be lost in translation. As the original database is lacking, the possible deductions using the database will likely be useless.

It is commonsense, that if the number of variables is increased in a problem the necessary time to calculate the answer also increases. In many problems the necessary computation time grows exponentially. In computer science this is called intractability, and it imposes the limitation that only the simplest versions of problems are solvable in a reasonable time.

Moravec's paradox states that tasks that need high-level reasoning, such as calculating a product, require little computation, but tasks that are low-level for us, such as walking, require an incredible amount of computation. An explanation for this by Moravec was:

"Encoded in the large, highly evolved sensory and motor portions of the human brain is a billion years of experience about the nature of the world and how to survive in it. The deliberate process we call reasoning is, I believe, the thinnest veneer of human thought, effective only because it is supported by this much older and much more powerful, though usually unconscious, sensorimotor knowledge. We are all prodigious olympians in perceptual and motor areas, so good that we make the difficult look easy. Abstract thought, though, is a new trick, perhaps less than 100 thousand years old. We have not yet mastered it. It is not all that intrinsically difficult; it just seems so when we do it."

Narrow AI

Facing these very difficult problems, we lower our expectations, instead of creating a general-purpose full AI, we propose an AI that can only do one or two things, a so called "narrow" AI. As the scope of the possible inputs and problems is greatly reduced, we can skirt around the commonsense problem, as the environment the AI interacts with becomes smaller and simpler. The intractability problem becomes less significant as well. Furthermore these AIs can learn. Every time the AI produces an output, the user evaluates it. New inputs add new information to the database the AI can draw from in the future, and the evaluations can help optimize the processes of the AI, decreasing the computation time and lessening the effects of intractability.

Google's search engine, Amazon's recommendation system or some stock trading algorithms are examples of a narrow AI. As narrow AIs are the ones commonly used today, from here on when saying AI we will mean narrow AI.

The strength of these AIs are their ability to handle enormous amount of information, and based on this information giving the answer, with the highest probability of correctness. However, as the inputs become less and less abstract, and closer to human communication, their responses become less and less useful. Furthermore, the more peculiar the inputs are, the less likely you will get a reasonable response. Just try looking up "SPICY PORN circuit" on Google. You will not be shown a link to the electrical circuit simulator software you were looking for.

2. How does an AI do the work and perform?

Are AIs readers or listeners?

AIs are readers, as they need the input to be well formatted, whether it's new information to expand on its database, or a request. Anybody who ever dabbled in data-mining or machine learning knows that usually 50-60 percent of the work is acquiring, formatting and cleaning the data for the AI to process. As natural language processing techniques are developed, AIs might become competent listeners, but they will never understand all the subtleties of human communication, and it might become a source of errors. As the joke goes, the robots will rise against us not because of some design flaw, but because of misunderstood sarcasm.

How does an AI learn?

There are many different learning algorithms in use today, but their basic structure is the same. The AI produces an output, responding to an output, then the quality of output is evaluated according to some value function. We then try to change the AI in some way so the value of the output increases. We repeat this process, until we cannot increase the value of outputs anymore.
The important thing to keep in mind is, that no matter how good the learning algorithm is, an AI can only be as smart as its training data allows it. If we want to create an AI that separates fruit, and we use only apples and oranges to train it, it will not be able to handle mangoes very well.

In what relationship does an AI work well?

AIs work best as subordinates. As these are narrow AIs, they are unaware of the big-picture, while their advice should be considered, when making a decision, an AI should not be a critical decision-maker.
In situations when a huge number of routine decisions have to be made, and the occasional error is acceptable, such as sorting post, an AI might be economical and useful, but even then some supervision is required.
To illustrate this point, just take a look at stock trading AIs, sometimes called 'algos'. It is said that 70 percent of trades today are performed by algos, as the microsecond-scale of high-frequency trading, cannot be achieved by humans. There have been instances when malfunctioning AIs shut down the London Stock Exchange and the New York Stock Exchange.

What are the values of an AI?

When it comes to AIs we have to talk about two kinds of values. The first kind are the values explicitly coded into the value function of an AI. For a stock trading algorithm this could be the whether the algorithms values short-term returns or long-term returns when choosing an investment.
The second kind are implicit, and are derived from the training data, sort of rules of thumb used by the algorithm. For example, we want to create an AI that gives a recommendation of sentence to a judge in criminal cases. To do this we input all kinds of data about previous cases, demographical data about victims and perpetrators, the

crimes etc. and of course the sentencing. We set the value function so the outputs closest to actual sentences are rated highest.

We reason that inputting the sentences of many judges we can recommend a fair sentence, since the sentences of the "outlier" judges will be drowned out by the sentences of the "average" judges. However when we start testing the AI, we found that if we input the same crime, for a white and a black perpetrator, it recommends a longer sentence for the black one. When we analyze the training data we find, that in American court cases black men were given a longer sentence, and the AI learnt this tendency, as it tried to match its outputs to actual sentences. Again, an AI is only as good as its training data.

3. What should be the contribution of AI?

An AI can aggregate huge amounts of data and find patterns in a completely mathematical, inhuman way. Remember, we are just guests in the world of numerical data, while AI was born there. Just as a poet has different perspectives from a plumber, an AI looks at the world in a different way from humans.

In other words an AI used by a knowledge worker should bring tools fundamentally different from ours to the table, and as Drucker wrote: *"Effective work is actually done in and by teams of diverse knowledge and skills."* Drucker is not the only one with this opinion, Scott E. Page puts it into three words: '*"Diversity trumps ability"*. The reason centaur chess players can beat even top human and computer players, is not because of their individual skills, but because of their diverse skills and knowledge.

4. Relationship Responsibility

The important thing when working with an AI is that every part of the communication process depends on you. You have to make sure the input data (that can also become training data) is correct and well formatted, to ensure useful outputs.

When evaluating results one has to take them with a grain of salt, however even if they seem impossible or just plain stupid one should consider them. Remember AIs do not have commonsense which is one of its shortcomings, but it is also one of its strengths.

While commonsense simplifies our environment to a level that it becomes livable, it is not necessarily correct all the time. On the contrary, most great discoveries seemed to defy the commonsense of their time.

5. The Second Half of the Life of AI

AIs are capable of more and more every day, be it speech recognition, machine translation or driving a car. As technologies like this mature, it will inevitably displace many human workers, call-center operators or drivers, as an AI does not breaks or salaries for that matter.

This direction of research tries to work its way down from the high-level, abstract tasks that are easy for a computer to the lower levels, which are hard for AI, opposing Moravec's paradox through sheer computational power and effort. However sooner or later, the paradox will win, as the limitations imposed by intractability and the commonsense problem become too hard to overcome. For example when somebody says something to us, we do not only use the meaning of the words and grammar to understand what he or she is staying, but we also consider who the other person is,.or the situation we are in.

Drucker writes: *"One cannot build performance on weaknesses, let alone on something one cannot do at all."* Instead of trying to copy human capabilities and intelligence, we can progress AI to another direction, creating a fundamentally different, inhuman artificial intelligence. Why would we want to create another human intelligence, we already have seven billion of them. This inhuman intelligence would have new perspectives, perspectives that no human can have. This would be an easy and good way to increase cognitive diversity. Instead trying to slightly improve the hammer, we should create new tools altogether.

The 21st century knowledge worker has to learn how to cooperate with AIs, how to become a centaur, wise and skillful like Chiron. If he or she does not, it will be just like a man trying to overpower a horse, not very successful.