Introduction

Metaphysics of Physics is the much needed and crucial voice of reason in the philosophy of science, rarely found anywhere else in the world today. We are equipped with the fundamental principles of a rational philosophy that gives us the edge, may make us misfits in the mainstream sciences but also attracts rational minds to our community.

With this show, we are fighting for a more rational world, mostly by looking through the lens of the philosophy of science. We raise awareness of issues within the philosophy of science and present alternative and rational approaches.

You can find all the episodes, transcripts, subscription options and more on the website at metaphysicsofphysics.com.

Hi everyone! This is episode fourteen of the Metaphysics of Physics podcast. I am Ashna, your host and guide through the hallowed halls of the philosophy of science. Thanks for tuning in!

Today we are going to start our review of "Life 3.0" by Max Tegmark. This will be the first part of a series where we go many of the central ideas presented in this terrible book.

Today we will cover the prelude and the first chapter. Later parts will cover further chapters at about two or three chapters per part. Meaning that the entire series will be about three or four parts long.

But, without further ado, let us start with a quick introduction to the book itself.

The book is called "Life 3.0" and it is subtitled: "Being human in the age of artificial intelligence". Which, to be fair, does give you a fair idea of what you should expect.

Here is the end of the blurb provided on the inside jacket of the copy I have before me:

"What sort of future do you want? This book empowers you to join what may be the most important conversation of our time. It doesn't shy away from the full range of viewpoints or from the most controversial issues - from super-intelligence to meaning, consciousness and the ultimate physical limits on life in the cosmos".

Basically, it argues that artificial intelligence in the form of greater than human level intelligence is all but inevitable. And that we should start thinking about what this implies for us. Now, rather than in the future when Max Tegmark believes it will be too late.

The book starts by making the case that the issue of how to handle the possible rise of artificial intelligence is the most important issue of our time.

It then goes on to show the possible benefits and dangers of AI and how it might drastically alter our lives and civilizations. And what we should do to make sure AI does not prove to be dangerous enough to wipe us out.

Before, we go any further, when I say "AI", it should be assumed that I mean "strong AI" or "human-level intelligence" unless otherwise stated. Alright, now we have that noted, let us continue.

What do we think of all of this? Well, the main issue we have is that it makes a huge, huge leap: That AI is possible in the first place. We have argued that in fact, it is not.

You can see our argument for this presented way back in episode four:

https://metaphysicsofphysics.com/episode-four-the-possible-and-ai/

If we were to assume that such AI is indeed possible, then we would probably leave the book alone. Since if AI was indeed possible, then some of it would certainly follow.

We disagree with this premise, so we are not going to leave his book alone. Instead, we are going to deal with his arguments for AI and whatever other philosophically dubious ideas we encounter.

This being Max Tegmark, we should not have a lot of trouble finding quite a few philosophically dubious ideas.

Now that you have some idea of what the book is all about, let us start our criticism of the book. Starting with the prelude and working our way through chapter by chapter. We will only deal with the philosophically interesting parts of the book and leave the others alone.

But why pick on Tegmark at all? Well, because a lot of other AI researchers agree with many of the ideas presented in this book.

This episode will deal with the prelude and will then cover chapter one and chapter two.

Prelude

The prelude is a fantastical scenario where a bunch of programmers called the Omega Team covertly creates some pretty creates some artificial intelligence, which they call Prometheus. They then proceed to secretly unleash it upon the world.

Prometheus eventually proves to be up to any task given to it. It can make money on the stock market, make movies better than most human studios, it can even design and create amazing new technology. You name it, Prometheus can probably do it better than any human could hope to.

It massively disrupts the economy and starts dominating virtually every industry in existence. It and the army of AI's it creates put countless people out of work and people all over the world start depending on Prometheus for virtually everything.

This allows it to expand massively, virtually without limit. Until it is entrusted with almost everything. Including political power. At this point, it essentially runs the world while humanity sits back and reaps the benefits.

While this makes somewhat interesting science fiction, I hardly consider it very realistic. But it seems that this chapter represents one of the possible courses of development Tegmark considers AI could take. And one of the better ones, at that.

However, this is apparently the kind of future Tegmark believers might happen. He is far from alone in this and he is joined by many, such as Larry Page from Google and others.

I am not so sure I agree with this. Nor do I agree that humans would let themselves become so unproductive simply because AI exists which can make so many of their decisions for them.

Even if the kind of super-intelligence he portrays could exist, I think it would be used to supplement human decision-making. It would do a lot of the work humans do not excel at or which is very dangerous or otherwise undesirable.

There would still be countless productive things humans could do. Including the creative arts, academic pursuits or anything else which the AI did was not quite as good as people. Which is actually quite a lot of things.

Which is why I do not greatly worry about things like robots taking jobs or the like. There will always be plenty of things for humans to do.

Then there is the fact that the human population is in many areas of the world, expected to seriously decline. Despite widespread paranoia about overpopulation. In this case, then there may very well not be nearly enough people to fill the jobs that need to be done to maintain an advanced civilization and we should welcome robots filling in for the shortage of people.

There is not really a great deal else to say about the prelude. It is largely intended as Tegmark presenting speculative fiction for the purpose of setting the scene for later chapters.

Chapter One

In this chapter, Tegmark defines life as a process which can retain its complexly and replicate. This is, of course, a drastically over-simplified definition of life that is not limited to biological life or any anything of the sort.

In fact, B to this definition, a purely mechanical system which manages to main at a constant (or improving) level of complexity and which is able to replicate itself is alive! So, that would mean that an assembly line able to create a copy of itself would be alive!

This is clearly a problematic and ridiculously oversimplified definition of life. But, Tegmark seems to have defined it so that it is not limited to biological life and can easily include anything that can replicate itself and maintain its complexity. This presumably makes it easier to include machines which meet little or none of the criteria for life used by biologists.

He then proceeds to divide life into three developmental stages. These three stages are:

Life 1.0: A biological stage where its hardware and software are evolved.

By hardware, he means the physical aspects of a life-form. In other words, its physical body. Its software is the information that life-form possesses.

This information would include memories, associations and the like. Lifeforms at this stage have no control over their software. Or in other words, they have no control over the contents of their brains and cannot "reprogram" it in any way.

Life 2.0: A cultural stage. At this stage, lifeforms have control over their software. That is, they can design their software by learning.

Life 3.0: A technological stage where a being can not only design its own software, but also its hardware.

This kind of being can design or redesign its physical form. It also does not need to evolve and can be created independently of any process of evolution.

Now we see why Tegmark defined life as he did. He wants to make it easy to classify machines as potentially being a form of life.

This raises the question: What is learning?

He does not tackle the question of what learning is until chapter two! Even there, he does not really pin it down in any definite way.

He seems to consider learning to be, in my own words: "the ability to perform some task".

To call that learning is rather ridiculous. Yes, a computer can acquire the abilities that they did not originally have. For instance, a computer can acquire the ability to recognize human faces in photos or to identify fraudulent financial transactions. But does that mean that it has learned anything?

Learning is in fact, the acquisition of *knowledge*. Which is a mental grasp of reality acquired by perceptual observation or a process of reasoning based on perceptual observation.

If a computer learns to identify human faces in a photo has it learned anything? Has it any kind of mental grasp of reality? No. It has no mental grasp of anything, it has no mental abilities at all! In fact, it has no mind!

Does have any capacity for perceptual observation? No. It has no perceptual faculties of any kind. It is merely a machine and has no consciousness of any kind.

So, how could it perceive anything? The ability to process input from a camera or such is not the same thing as consciousness.

It most certainly has no ability to reason! Reasoning is the ability of conscious beings with volition to direct their mental processes for the purpose of attempting to understand some aspect of reality.

Computers do not have volition, nor any kind of mental faculties with which to reason. So, if they are not aware of anything nor have any mental faculties, how can they be said to learn?!

This is why Tegmark avoids defining what learning is and tries to suggest that learning is simply acquiring the ability to do something. That way, at least for now, he does not have to deal with the issues of consciousness or whether or not something has any kind of mind.

Which brings us to his definition of intelligence...

Intelligence: The ability to acquire complex goals.

What does he define goals as? He doesn't define what he means by "goals".

We can turn to his definition of "having a goal" for some clue:

Having a goal: Exhibiting goal-orientated behaviour.

Alright, what is "goal-orientated behaviour"?

Goal-orientated behaviour: Behaviour more easily explained via its effect rather than its cause.

What this eventually leads to in a later chapter, is treating a goal as some kind of "desired outcome". As though certain things act in order to do something.

You see, Tegmark likes to use teleology as an explanation for things. Many physicists are guilty of this, but Tegmark is very explicit.

For instance, physicists sometimes claim that gases behave the way they do so because entropy must always increase.

That is not why gases act that way. That is trying to explain the behaviour of gases as though they somehow know how they are meant to behave and that this explains why they behave that way. As though things act the way they do in order to achieve some end state.

But this explains nothing. Things act the way they do because of their nature. There is something about their nature which results in them acting that way. They do not act that way in order to reach a certain state.

Tegmark is doing something similar here. He is pretending as though machines have a goal simply because they tend to operate in a way that results in certain outcomes.

It is then easier to assume that the reason this happens is because machines have the goal to achieve that outcome! This kind of thinking leads him to conclude that "even missiles have goals"!

Which is absurd. No, missiles do not have goals. Missiles do not strike their target because of some desired outcome. They do so because things have been set up in such a way that it is what must happen as it is in the missiles nature to do so in that context.

Why does he use goal in this way? Well, because anything with the ability to acquire complex goals is apparently intelligent! So, as long as it is capable of achieving a complex desired outcome it is intelligent!

The thing does not, apparently, need to be conscious or have any kind of mind to be intelligent. It simply has to achieve the desired outcomes.

Desired by whom? Does not matter. Apparently, it can be desired by the thing itself, or its creator or just someone using it! Presumably, as long as it is able to achieve those outcomes, it is intelligent.

He gives no real criteria for what qualifies as a "complex" goal, making it very hard to know exactly what would qualify as intelligent.

According to this definition, a pocket calculator is certainly intelligent. It can be used to achieve the complex goal of performing complex mathematical operations such as finding square roots. Or, is that not complex enough?

It is clear that he would consider a computer intelligent if it was able to pick up human faces in a photo. Or to recognize possibly suspect

financial transactions. Those tend to be fairly complex tasks, even if the former can be easily performed by most people.

Of course, this is ridiculous. Yes, a computer can perform very complex tasks, but that does not make it intelligent. Even if it does so according to the goals of its user.

Clearly, machines can have no goals of their own, as they have no volition or awareness of any kind.

So, here we see his terrible definition of intelligence, built on very shaky grounds.

The rest of the chapter goes on to ramble about to debate about different people debating whether or not AI is possible and when we should start worrying about it. But. that is not all that interesting and we shall not discuss it, at least not yet.

Outro

That brings us to the end of this episode. I hope you enjoyed our first part of the "Life 3.0: A Slow Death" series. The second part will not be next week but will come out in the form of a blog post at some point within the next month or so.

Next episode will be another Quora episode. There we will answer Quora questions on mathematics.

In April we will be launching our subscription content. This will be content which can be accessed for the very small monthly fee of $2.

Thanks for listening!

Remember to check out the website and subscribe if you like our podcast, sign up to our mailing list or follow us on Facebook or Twitter to get the updates!

You can also look at our Metaphysics of Physics merchandise if you wish. All profits from these go back into the show. Click here to take a look.

As always, you are welcome to send in questions about any of the things talked about in this episode or about irrational stuff in physics or the

philosophy of science in general. Send them to questions@metaphysicsofphysics.com.

Please tune in for the next episode and start thinking of some questions! Until then, stay rational!