Advanced Artificial Intelligence

In The Terminator, released in 1984, James Cameron was of the opinion that by the year 2029, technology would get to the point where machines could hold dominion over mankind. In 1984, he probably thought that a 45-year time gap was a reasonable enough period for the audience to believe that technology could catch up with the premise of the film. Well, we are seven years away from a global nuclear war. Similarly, in Avengers: Age of Ultron, released in 2015, an artificially sentient being decides within minutes of gaining consciousness that the only way to save Earth is the extinction of humankind. This begs the question of whether we should allow the field of artificial intelligence to gain traction and embed it into the very core of our daily processes.

Right now, in this program and elsewhere, the conversation is heavily focused on bias in AI. Now, let's assume that we manage to fix the issue of developer bias in AI and all other programming-related issues in developing sentient AI. And we time travel 100 years into the future, where artificial intelligence is present in every little aspect of life. Can we be sure that AI won't develop a bias of its own?

If our goal is to reproduce human intelligence in an artificial being, we should be ready for the artificial intelligence to make poor decisions the way humans do. And if we build parameters into the artificial intelligence model that it cannot function outside of, then is it really sentient? I believe that sentient artificial intelligence is a paradox. We cannot expect artificial intelligence to inculcate the decent part of humans and leave out the absurdity without downgrading what it means to be truly artificially intelligent.

Artificial consciousness. This is the term I am using to define the next iteration of artificial intelligence. This is the stage where artificial intelligence systems are no longer performing automated functions as they have been instructed; instead, they are changing the parameters of the models as they see fit, to the point where they can discard the models altogether if they deem them to be irrelevant. Humanity does not need to fear artificial intelligence; it needs to fear artificial consciousness.

Can different ideologies of artificial consciousness collide with each other? What happens when two sets of artificially intelligent systems are supposed to complete one task, but the systems cannot mutually agree on a solution? Does that mean we need to have a "superior" AI in place that can act like a manager? Is this how far we wish this field would expand, or do we want to stop it from developing any further after reaching a certain level of maturity?

History has taught us about lost civilizations, the downfall of powerful empires, and the way society has transformed over millennia. Ironically, the one thing that has remained constant throughout history is that nothing ever stays the same. So, my question is: what will replace AI? Are we being silly by investing all these resources into studying artificial intelligence when it might not even matter in some time? I understand that AI has the power to transform humanity, but what will humanity look like when the next big thing takes over?

The far-reaching consequences of artificial intelligence bleed into industries such as health care, research, construction, automobiles, films, love, and countless others. We can write entire theses about how AI can transform each one of these activities in the long term. But for now, we are constrained by the limitations of a linear timeline that stops us from learning how the world will change over the next thousand millennia.