Thoughts on Artificial Intelligence

Some years ago i learned of the concept of the Singularity, the idea that once (or if) an artificial intelligence arises that is at human intelligence level (but much faster due to the use of electronics), it will then apply itself to the task of designing improved versions of itself. Those new versions would in turn continue to design better versions, resulting in a rapid rise in intelligence, far exceeding that of humans in a relatively short time. Estimates place the start of the singularity as soon as 2050, and probably no later than 2100 (if it occurs at all). Probably the best books on the subject are those by Ray Kurzweil.

Believers in the inevitability of the singularity fall into two broad camps. One camp posits (without any evidence, IMO) that the greater intelligence and speed of AI will result in a benevolent super race of machines that will take care of all human needs, leaving us the leisure to do what we like, without the necessity of earning a living. The other camp believes (based on analogy with the generally poor destiny of less technologically developed humans after they come into contact with those with greater technology) that humans will suffer greatly, and maybe become extinct, when machines are in control.

Progress in AI has been remarkably rapid in the last few years, with the results that machines can do pretty amazing things in the areas of image recognition, language "understanding", medical diagnosis, as well as more traditional computational tasks such as modeling (climate, aircraft and other designing, protein behavior, anti-viral testing, etc.). This progress has been made chiefly by so-called "deep learning", which uses massive computing power and huge amounts of example data to build networks that are generally very good in predicting system behavior.

A recent, much publicized example of deep learning is ChatGPT, an application that people can interact with in a fashion similar to humans. You can (e.g.) ask it questions or request it write essays or poetry about various subjects. In some cases its responses simulate that of human writers pretty well, though you can't always trust the "facts" that it returns. But it is sufficiently good that some teachers are beginning to worry about its use by students to complete writing assignments. I recently asked it to write a parody of a Dylan song; you can see the result on my website at

https://rlgreene.net/song-writing/parodylans/i-shall-be-relieved/. The raw result was pretty poor by human standards, and probably not singable. However, it had some interesting ideas and rhyme suggestions; i ended up using (with modification) perhaps 25% of the output.

Although this deep learning approach will continue to be used to build better and better tools that will exceed human capacity in many fields (and indeed have already exceeded that of humans in a lot of areas), in my view it is unlikely to lead to sentient AI of the sort envisioned by the singularity advocates. I think that for machines to become sentient, less data-intensive approaches will be required – i.e., machines that are capable of rapid learning from relatively few examples, like humans do. That probably requires hardware with complexity comparable to the human brain, which is likely is still a decade or so away. I personally believe that it can happen, though probably not in my lifetime.

Will it happen? I actually doubt that it will, because i am of the opinion that humankind is going to fall back into a dark age, or perhaps even become extinct, due to problems caused by our overpopulation (pandemics, climate destruction) and/or our propensity for violence (nuclear or biological war). I don't think our leaders have the political gumption or that our species is sufficiently far-sighted enough to prevent such a disaster.