AI and the secular vision to redefine life itself

Lucien Tuinstra

Author Max Tegmark (born Shapiro) is a Swedish-American physicist and cosmologist, a professor at MIT,[1] and co-founder of the Future of Life Institute (FLI), a 'beneficial AI movement' (more on this below). In this, his second, book he discusses a wide range of scientific topics, including information storage systems (memory), learning, intelligence, and consciousness. As a believer in evolution and big bang cosmogony, who is actually best known for his cosmological theories, some of his statements will not resonate with all readers. Nonetheless, his historical views don't really affect what is going on at present and what may (or may not) happen in the future with Artificial Intelligence (AI). Actually, arguments from design are rife throughout the book.

Books on AI-related topics (e.g. transhumanism) abound but Life 3.0 gives an overview of what the future may hold, to laymen and experts alike. Tegmark's previous book is titled *Our Mathematical Universe: My quest for the ultimate nature of reality.*[2] Perhaps the conclusion(s) of that work inspired the title of this book's opening chapter: "Welcome to the most important conversation of our time" (p. 22). It's a great one-liner to keep the reader engaged or lure in prospective buyers looking at the contents of this best seller, but not one with which Christians will agree!
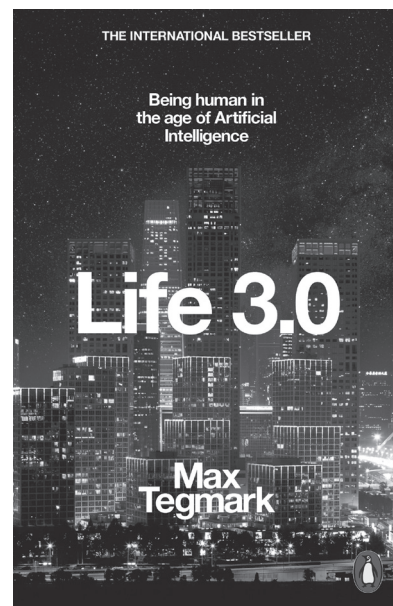
A momentous topic?

Tegmark envisions three distinct stages in life's evolution (figure 1) and qualifies this important conversation about Life 3.0 "in terms of both urgency and impact" (p. 37). The urgency stems from a timescale shorter than the threat of "climate change" (p. 37), with the impact deemed to be worldwide. Life 1.0 is biological and replicates. What sets humans (Life 2.0) apart is that we have an added cultural aspect. Life 3.0 is the next stage, where technology will increasingly be implemented. Tegmark points out that humans really are at Life 2.1, as we already use implants, prostheses, and other such technology.

Assuming technology continues to increase, how do we feel about AI outperforming human beings on cognitive tasks and when do we think this will happen (figure 2)?

Undoubtedly, current technology outperforms us on particular tasks (e.g. pocket calculators, the best computer chess programs such as Stockfish). These examples of computational power are rather bespoke devices. Neither is good at anything else but its designed task (arithmetic and chess respectively) and cannot 'think' outside the box.

A breakthrough happened when AlphaGo (which plays the ancient Japanese board game Go) made an unexpected, counter-intuitive (creative?) move, going against thousands of years of human intuition, perhaps better explained here as a great confidence in long-term strategy. The surprise play proved key to the victory, which became clear approximately 50 moves later. This may not seem very impressive at first sight, but the theoretical number of possible resulting games after this point of the game, move #37, is as follows:[3]

$$\binom{324}{50} \; which \; is \; > 2 \times 10^{59}$$

AlphaGo's achievement is impressive, but could AI improve on man-made strategies in cases such as the military, investment, and politics? Diligence is required given the great loss caused by some preventable (with hindsight) past situations. One company lost "$440 million in forty-five minutes after deploying unverified trading software" (p. 96). Tesla's self-driving car was involved in a deadly crash when the bright side of a lorry was interpreted as part of the bright sky. However, these examples should not turn us into Luddites.[4] Humans, unlike AI, trespass the law with speeding and jumping the orange (or red?) light when they should stop. The root cause in the above examples is found to be, (a) sloppy programming (sometimes due to incorrect assumptions) or (b) incorrect

interpretation of sensing indicators. Thus, further steps are vital:

a. Verification—"Did I build the system right?"
b. Validation—"Did I build the right system?" (pp. 96–97)

We can agree that verification and validation relate to design, programming and assembly, software, and hardware. The question is: can these be done by non-human intelligence?

Did intelligence evolve from matter?

The easiest of these concepts to come to terms with is assembly. This already happens, for instance in car manufacturing. Software drives the hardware of those machines which assemble the various parts into a functional car. Life 1.0 (figure 1) is depicted by what some call a simple cell. Actually, surviving and replicating cells are anything but simple. They are bestowed by their creator with very sophisticated software that actively performs computations and stores them in memory. Memory is defined as an information storage system (e.g. a genome) and computations transform information, i.e. changing its memory state. Assembly cannot occur without software (whether computer-based or cell-based), which in turn demands an intelligent source.[5]

What about design and programming? Tegmark explains that a goal can be pursued by human intelligence (Life 2.0) but also by an adequate AI. He also believes evolution is goal-directed. The big question is whether AI can become creative and even conscious (more on consciousness below). People, made in God's image, are creative—think of the arts—and make discoveries by "thinking God's thoughts after Him".[6] *Some* might say that AlphaGo was also creative. However, possessing lots of knowledge and calculating power does not make one intelligent even if it gives the appearance of that very thing. A computer could calculate numerous routes to goal achievement. This is not intelligence, and in some cases even wastes time. What about the human experience—a gut feeling or intuition—that something might not work, so an alternative is chosen and fleshed out? Educated guesses are common for humans and they start early in life.

A key aspect of AI is learning (adding knowledge/gaining experience). Take face recognition, for example. That this is not yet flawless can be seen at airport security gates fitted with this feature, despite the strict rules imposed on passport photos (no smiles/hats/glasses etc.). If the person standing in front of the camera looks identical to the photo stored on file, this would be straightforward. Nevertheless, anyone who has travelled on a long-haul flight knows that one's appearance can change dramatically over the course of hours—think of a grumpy, unshaven, baggy-eyed face after the person has been jammed in a seat for the better part of half a day, not to mention the use of make-up, which can change appearances even quicker.

What about wisdom? Can AI become evil? No, not intrinsically. Only morally culpable beings can knowingly commit evil. However, an action may be perceived as evil by human beings, but for AI to know it as evil, it would need to be conscious. A hypothetical *evil* computer would *know* the correct answer to 1+1, but return something other than 2 on purpose, with the *intent* to deceive. In reality, most glitches are software related; rarely is there a mechanical failure. The perception of evil is likely due to misalignment of goals, and this is a sticky point, because we humans cannot get our goals aligned. Even when we think we have achieved it, there can be person-to-person differences in interpretation (sometimes underpinned by unspoken assumptions), leading to all sorts of consternation! Tegmark aptly highlights this with the story of the genie in the lamp. More often than not, the person's third wish ends up undoing the previous two wishes.

What might the future hold?

Computing memory size per dollar drops every two decades by a factor of 1,000. Computational power per dollar doubles every couple of years. It is now accepted that AI already exists on some level and specialists are increasing its sophistication continually.[7] Is there any doubt that AI will multiply in years to come? AI outperforming human beings on specific tasks will certainly increase. This is not a cause for concern. Among billions of global citizens, chances are there is somebody better than you at any given task. Even if you are the most skilful in something, there are countless other tasks in which you are not the best. However, humans have a fabulous ability to do many things well, and yes, excel in a few. It is quite possible we could program a goal into an AI and it will master it

	Life 1.0 (simple biological)	Life 2.0 (cultural)	Life 3.0 (technological)
Can it survive & replicate?	✓	✓ Hello	✓ Hello
Can it design its software?	✗	✓ Hola	✓ Hola
Can it design its hardware?	✗	✗	✓ Bye!

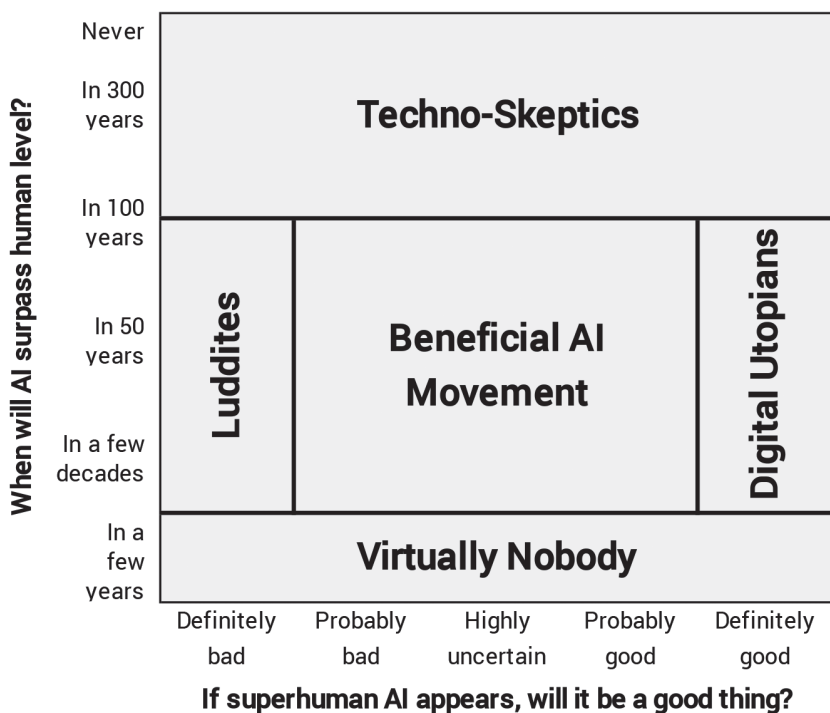**Figure 1.** The three stages of life (after Tegmark, figure 1.1, p. 26)

Figure 2. Distinct schools of thought regarding AI (after Tegmark, figure 1.2, p. 31)

(shown already with the Atari game 'Breakout'[8]). In the future, Tegmark believes it could even build a robot (hardware) to beat humans at tennis, for example. But to master all its goals, this AI would need a vast number of combinations of software and hardware.

Tegmark describes a list of aftermaths following such a future intelligence explosion (rapid acceleration of AI). One of them is the 'Protector god' scenario: "Essentially omniscient and omnipotent AI maximizes human happiness" (p. 162). The AI operates in the background. This in contrast to the 'Benevolent dictator' AI, where everybody knows who is running the show.

Another scenario is the malevolent 'Conqueror', where AI "decides that humans are a threat/nuisance/waste of resources" (p. 162). This of course is lucrative fiction for Hollywood blockbusters, but, like the 'Protector god' and other scenarios, is based on an idea that intelligence can result from inanimate matter, which can serve as information storage, but has no ability to freely manipulate that information in order for intelligent properties to emerge. Information comes from intelligent beings, not from goalless processes involving random chance. AI operates at the level of syntax and cosyntics (code+syntax), but lacks semantics, pragmatics, and apobetics.[9]

Human objectives regarding AI

Tegmark says that "most matter on Earth that exhibits goal-oriented properties may soon be designed rather than evolved" (p. 258) by AI. Indeed, the inanimate matter does not generate the goals, but the intelligent designers do. Critically, AI can only be useful if it has its goals aligned with those of humans. The author recognizes that humans don't all have the same goals, which poses a greater problem underlying the next stages for AI to achieve (p. 260):

1. Learn our goals
2. Adopt our goals
3. Retain our goals

It all hinges on the words 'our goals'. Whose goals? Analogically, if a law is legislation of morals, the question is: whose morals? On multiple occasions, Tegmark invites people to join the discussion, but more often than not, it is the rich or the powerful minority (usually both) that dictate how things turn out. Soberly but astutely, atheist Yuval Noah Harari commented: "If the future of humanity is decided in your absence, because you are too busy feeding and clothing your kids—you and they will not be exempt from the consequences."[10] Given the political (and spiritual) state of affairs, it is unlikely that a global consensus will be found. Despite Tegmark's noble endeavours to make this a global approach, it will probably become a fragmented effort, with different authorities working on their own implementation of AI with different goals.

Tegmark shows his optimism of AI's abilities with calculations demonstrating that future gains in technology are only limited by the laws of physics. Yet, he goes further, suggesting essentially that alchemy will make a comeback; that is, AI will be able to turn any material made up of quarks into any other material. It is not quite *creatio ex nihilo*, but not far from it. The underlying thought is that hydrogen was (presumably) converted into the other elements in the past so, going forward, this will be a process controlled by AI!

The fiction continues with the artificial assembly of humans, starting with "two gigabytes of information needed to specify a person's DNA and then incubating a baby to be raised by the AI, or the AI could nano-assemble quarks and electrons into full-grown people who would have all the memories scanned from their originals back on earth" (p. 225)! On the other hand, without a body, "the prospect of infinite computation could translate into subjective immortality

for simulated life forms" (p. 232). Tegmark's future vision of "many humans that take solace in a belief that their minds will live on after their physical bodies die" (p. 237) may very well appeal to people with his worldview, but does not apply to the Christian who has confident hope of also being granted a new body (1 Corinthians 15:44; Philippians 3:21). Needless to say, stripping out our five senses—among many other things—is a massive downgrade from, indeed a destruction of, real human life.

Superhuman intelligence and beyond

Tegmark considers that consciousness is by far the most remarkable "trait" of human intelligence (p. 184), but fails to recognize (or accept) that it is God's image borne by humans that makes them special. His worldview is rather plastic. On the one hand, he says life evolved the way it did due to its goals (the fallacy of reification), but then acknowledges that (according to some) life on our planet evolved because of "a wild stroke of luck" (p. 243). He probably means that the fittest survived and survival, after all, was its goal (fallacy of begging the question). In the context of his broad definition that "consciousness = subjective experience" (p. 283), he quotes Yuval Noah Harari—who believes superhuman intelligent AI is a threat—from his book *Homo Deus*:[11] "If any scientist wants to argue that subjective experiences are irrelevant, their challenge is to explain why torture or rape are wrong without reference to any subjective experience" (p. 283). Unfortunately, he fails to discuss the moral implication of the word 'wrong'. It is clear that conscious, morally culpable beings (i.e. humans) have judicial laws that make these two examples a crime. Bible believers point to *objective* standards, established by

a wholly good Lawgiver. Therefore, torture and rape are wrong, irrespective of the experience of the subject. This does not mean they deny the reality of the victim's horrendous experience.

The 'redundancy' of the physical senses of post-mortem humans (whose intelligence was then 'stored' computationally), as hypothesized by AI proponents, would mean any experience that remained would be non-physical. There are many people that have one or more senses not working, yet they are conscious. Even people whose memories last a minute or less can be perfectly conscious. The Bible is clear that some living things (e.g. plants) were good for food from the beginning (Genesis 1:29–30). Other living things had an additional life principle (*nephesh chayyāh*). Members of the former category are not conscious, whereas the latter category contains (at least some) conscious members. Humans are a category set apart. They bear the image of God, and are also moral beings. From a Christian perspective, some might assume AI could be conscious but it would still not be alive in the biblical sense, let alone morally aware.

Tegmark ends on the following note: "It's not our Universe giving meaning to conscious beings, but conscious beings giving meaning to our Universe" (p. 313). The first clause of this statement does not follow from the secular belief that conscious beings have ultimately emerged from inanimate matter, which he believes. And the Christian must also take issue with the idea that we are the ultimate attributors of meaning, a prerogative that belongs to God alone.

Conclusions

Tegmark's *Life 3.0: Being human in the age of Artificial Intelligence* is a blend of science (verifiable prescriptions) and fiction (future

scenario descriptions). This makes the book readable, despite the somewhat abstract topics.

Life 1.0 is biological life and can survive and replicate. Life 2.0 is more than that; it is cultural, and can design its own software—e.g. learn additional languages. Life 3.0, according to Tegmark, will be able to do the same, plus design its own hardware.

He has a positive outlook on where Life 3.0 might be heading. Holding to a worldview embracing big bang cosmogony and evolution, he believes Life 1.0 and Life 2.0 came about by random chance processes (rejecting divine design), albeit with goals of some sort. However, he is adamant that Life 3.0 will be designed, yet does not see the glaring irony of his belief. A member of the beneficial AI movement, Tegmark believes that working together it *will* be possible in the next century or so to have superhuman intelligence, consciousness not necessarily excluded. Will anything then be impossible (cf. Genesis 11:6)? If AI with superhuman intelligence should come about, people would no longer be the 'wisest' entities (as designated by our species name *sapiens*). Therefore, he suggests rebranding human beings, in the age of Artificial Intelligence, from *Homo sapiens* to *Homo sentiens* (laying the emphasis on our five senses rather than our then-inferior intelligence).

Haven't we been here before? 'Uploading' people would certainly stop them from being dispersed over the face of the whole earth (reminiscent of Babel). Bringing this full circle, Tegmark considers the discussion around Life 3.0 to be the most important of our day, whereas Christians instead focus on Life 4.0 (eternal life, only for those born twice). Therefore, the most important life focus is fulfilling the Great Commission (Mark 16:15).

References

1. Massachusetts Institute of Technology, Cambridge, MA, USA.

2. Alfred A. Knopf, New York, 2014.

3. There are 19 x 19 squares on a Go board and 37 positions have already been taken, leaving 324 available locations. My example here assumes that a further 50 moves have occurred.

4. A person opposed to new technology or ways of working.

5. Smith, C., Lost in translation: the genetic information code points to an intelligent source, creation.com/genetic-code-intelligence, 6 May 2010.

6. Usually attributed to Johannes Kepler (1571–1630).

7. Megget, K., AI has arrived. *PharmaTimes*, October 2018, pp. 30–33.

8. Breakout is a game where you position a platform to bounce back a ball that hits tiles. The goal of the game is to score as many points, by removing as many tiles as possible. This means not losing the ball off the bottom of the screen.

9. Gitt, W., *Without Excuse*, Creation Book Publishers, Atlanta, GA, 2011.

10. Harari, Y., *21 Lessons for the 21st century*, Jonathan Cape, London, 2018; as cited in Russell, J., A chilling warning, *Sunday Times*, 19 August 2018.

11. Harari, Y., *Homo Deus: A brief history of tomorrow*, Vintage, New York, 2017.