STEPHEN N WILLIAMS

Walton lecture: Thinking Machines? A Christian Perspective.[1]

The Thinking Scene

If we propose to offer a Christian perspective on thinking machines, it would seem that we are attempting to bridge not only two worlds, but also two worlds doubly apart. Firstly, it is a truism that Christianity is culturally marginal in much of the West. Where it is not, it standardly pops up socially wearing the shabby cloak of Fundamentalism, which does not seem exactly decked out for intellectual engagement with Artificial Intelligence (AI). Secondly, even if we bridge the worlds of Christianity in general and thinking machines in general, what can Christians sensibly say on the subject of thinking in the case of thinking machines? If they want to say something about the social and cultural significance of AI, so be it – the public square in a self-respecting democracy is open to a plurality of voices, including from the margins. But if we aspire to comment on the possibility of machines being able to think, how can there be a Christian perspective on the science and technology in question, more than there could be on quadratic equations, the chemical behaviour of calcium carbonate, or the metal conduction of electricity?

Those acquainted with Alan Turing's endlessly-referred to paper on 'Computing Machinery and Intelligence' (1950) – the one that introduced to the public the Turing Test – will suggest an answer.[2] Here, Turing described and responded to nine objections to the claim that machines might be able to think, and the first of these was theological. The objection goes like this: the organ of thought is the soul, to which the brain is instrumental; machines lack a soul; therefore, machines lack the capacity for thought. Turing made clear that he did not think much of theology in general, but while the objection is on the table, he would point out that it fails. Why? Because theologians believe that God is omnipotent. Thus, God could bestow a soul on an elephant, if he so wished, concomitantly conferring on the elephant, in the course of a mutation, an 'appropriately improved brain to minister to the needs of the soul'.[3] Thus, Turing claims that there should be no difficulty in principle in a divine exten-

1 This article is a lightly revised version of the Walton Lecture for Christians in Science, Ireland, given in October, 2021.

2 The essay is included in B. Jack Copeland (ed.), *The Essential Turing: Seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life plus the secrets of Enigma* (Oxford: Clarendon Press, 2004) 441-63.

3 *ibid*., p. 449-450.

sion of this capacity to machines, imparting to them a soul, and so enabling their mechanisms to think.

Turing's argument is not my direct concern; I merely note that he acknowledged a Christian (or theistic) perspective on the possibility of thinking machines. But how relevant is this over seventy years on? Has belief in the soul not been discredited not just outside, but within Christianity, with theologians standardly, though certainly not universally, insisting that the kind of separation between soul and body which underlies the position Turing tackled is theologically mistaken? If the mistake is admitted, we are back where we started, wondering what Christianity has to say to the possibility of thinking machines. It is best to start at the other end, and ask what leading figures in the world of AI themselves have to say on the subject.

The answer is a bewildering variety of things.[4] Let us pick out an influential account found in what is far and away the leading text-book in the field of AI, recently published in its fourth edition: Stuart Russell and Peter Norvig's *Artificial Intelligence: A Modern Approach*', which runs to over a thousand pages.[5] The question: 'Can Machines Really Think?' surfaces only at its 'Conclusion', constituting a sub-section of less than two pages to a discussion on 'Philosophy, Ethics, and Safety of AI'. While the authors by no means speak for everyone in the field, it is instructive to note how they answer their question, because their answer commands both sympathy and influence within their world.

'The question of whether *Machines Can Think*…is about as relevant as the question of whether *Submarines Can Swim*.' Thus the computer scientist, Edgar Dijkstra, whom Russell and Norvig quote. The relevance in question is relevance to design. Aircraft design is unaffected by the question of whether aircraft strictly fly, if adjudication on this matter is in the hands (or claws) of our feathered friends. The fact is that it makes obvious sense to speak of aircraft as flying. Our authors uphold Turing's conviction that incessant arguing over the question of whether machines think in some fully adumbrated philosophical sense of 'thinking', is rather futile. In this, Russell and Norvig are quite representative; those working in the field of AI will typically proceed along operational, rather than theoretical, lines in their practical discourse about and pursuit of intelligence. The familiar mantra that an intelligent machine is one that is able to do things which, if done by people, would be judged to require intelligence is applicable to machine thinking too, in that the pattern of attribution is the same. Russell and Norvig hold that if a machine behaves

4 For an indication of the range, see John Brockman (ed.), *What to Think About Machines That Think: Today' Leading Thinkers on the Age of Machine Intelligence* (New York, NY/London etc: Harper Perennial, 2015).

5 Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. (Harlow: Pearson, 2021).

intelligently, we should ascribe to it the capacity to think, as an extension of the 'polite convention' which we adopt in relation to other human beings who behave intelligently. We have no direct access to others' mental states, but the entirely reasonable thing to do is to affirm that they, like us, can and do think. So it is with machines. Intelligence surely entails awareness, and awareness is an aspect of consciousness.[6] Consciousness is a scientific mystery for us at present, but it is a reasonable presumption that this nexus of features – intelligence, thinking, awareness, consciousness – hang together in some form in AI, as they do in some form in human beings. Russell and Norvig do not want to ascribe to machines baldly, straightforwardly, studiously or simplistically the capacity to think, but for all practical purposes, we must regard them as possessing that capacity.

Our two authors make some huge conceptual leaps in order to achieve their conclusions, and there are plenty in the field of AI who would agree. As ever, those who come to it from outside that field need to reckon with an use of language which differs from the one which may obtain in other domains. It is not that a uniform concept of 'thinking' stalks the world of AI; on the contrary, this whole article could be devoted to different concepts of 'thinking' found in, let alone outside, that world. But in connection with terminology, we need to notice the way 'agency' is used. While there may be no standard use of that term in philosophical or theological discussion, it is frequently attributed, if not just to human beings, at least just to sentient beings. But Russell and Norvig describe 'an **agent'** as 'something that acts…'[7] Wolfgang Ertel (who wrote an introduction to AI partly because Russell and Norvig's was too long and expensive for many students), remarks that *'*[*a*]*gent denotes rather generally a system that processes information and produces an output from an input.'*[8] Max Tegmark, professor of Physics at MIT, notes that computer scientists will refer to 'entities that collect information about their environment from sensors and then process this information to decide how to act back on their environment' as 'intelligent agents'.[9] Those who work in AI will be surprised that I even bother to quote here; this is standard stuff. ' 'Agent' can just be a learning program.'[10]

I pick out 'agency' not as a random example of words to watch in inter-disciplinary discussions of AI, but because the language of agency exemplifies the way object and person are bracketed together in AI literature, and this alerts us to the context in which talk of machines that think goes on. What Russell and

6 'Self-awareness' and 'attention' are the other two aspects the authors invoke.

7 *op.cit.* (5)

8 Wolfgang Ertel, *Introduction to Artificial Intelligence* (London: Springer, 2011), 9.

9 Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (London: Penguin, 2017), 25.

10 Melanie Mitchell, *Artificial Intelligence: A Guide for Thinking Humans* (London: Pelican, 2020), 161.

Norvig say about rational and intelligent agency, in a discussion succeeding the introductory chapter, steers what follows in their volume. They deploy what they call a 'standard model' of rational agency: 'A **rational agent** is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome'; '*AI has focused on the study and construction of agents that **do the right thing***' and what counts as the right thing 'is defined by the objective that we provide to the agent'.[11] Accordingly, in the following chapter on 'Intelligent Agents', humans are treated as a species of the genus: 'rational agents'. The most prominent examples of rational agents powdering this chapter are enhanced vacuum cleaners and autonomous taxis. For moral philosophers and theologians, as for lay persons, the phrase 'doing the right thing' usually refers to human moral action. Russell and Norvig easily accommodate this: humans and vacuum cleaners both share rational agency, and moral agency is a form of rational agency applicable to humans, but not to vacuum cleaners.

There I leave Russell and Norvig, and AI literature in general. To repeat: our authors do not speak for the whole AI community; but neither are they idiosyncratic or marginal. This is what students at a number of educational institutions – schools, institutes of higher education, colleges, universities - are reading. In sum, their position is roughly: 'AI is rational *in our terms* (i.e. terms of computer science, of which AI is a sub-field) and intelligent *in our terms*, and thinks *on our terms*, and *our terms* are perfectly sensible, to say the least, and impeccably rational.'

In Christian Perspective

In response, what may be offered in the way of a Christian perspective on the possibility of thinking machines? The question of the soul in Christian theology is too big to handle here. Generally, and perhaps even universally, the biblical word, 'soul', whether in Hebrew or in Greek, does not pick out a disembodied entity capable of disembodied existence. If we take our lead from the Hebrew vocabulary in the OT, we might describe the preponderant meaning of 'soul' as the vital, animating principle of body – its life-principle. This is not necessarily in strict distinction to the solid, corporeal, material body which we all have; rather, the soul makes our bodies to be living bodies and not stones.[12] Animals are ensouled – living beings. But how useful is it to consult biblical terminology at all when addressing the contemporary question of thinking machines? Is it not an exercise in archaic futility? Whatever it has to say about the soul, does not the Bible speak of the *heart* as the organ of thought? The brain is not found

11 *op.cit.* (5), 22. 'Modern AI is based on the theory of "rational agents", arising from work on microeconomics in the 1940s by John von Neumann and others': Brockman, *What to Think*, 12.

12 When the person is dead, the soul is dead according to Numbers 6:6, where the Hebrew word translated 'body' in the case of 'dead body' is actually the word for 'soul'.

there. On the other hand, heart, along with kidneys, liver, bowels and head, for example, all feature in biblical anatomy. If we treat the vocabulary of biblical anthropology in terms of primitive science, the problem with bringing Scripture into dialogue with modern science is not that biblical language is scientifically imprecise. On the contrary; the problem is that it is scientifically precise. And it is precisely wrong. Witness talk of heart in the absence of brain.[13]

However, if reading the Bible in dialogue with modern science means treating it as a primitive science, we err. When we read it on its own terms, scientific precision is exactly what we do *not* find in its anthropology. Different biblical translations give us a clue here. At the milder end of things, alternative translations of one and the same Hebrew word give us 'heart' or 'mind'.[14] That is not just an English translation choice; the Septuagint (the Greek translation of the Old Testament, standardly used as Scripture in the early church), can do the exactly the same thing with the Hebrew. Again, when I pick up my New International Version (NIV) to read Psalm 73:21, I read that 'my heart was grieved and my spirit embittered', but the word translated here as 'spirit' is 'kidneys' in the Hebrew; the usual Hebrew word for 'spirit' is a different one. If you graduate from the NIV to the English Standard Version (ESV), you will read that it was your 'soul', rather than your 'spirit', which was embittered. Indeed, if we track the 'soul' over a wider terrain, we find that it can be translated not only as 'self', but even as 'throat' in the Old Testament.

We might conclude that it is just as well that biblical writers - or, at least, biblical translators - were not called on to perform surgical operations on heart, kidneys, throat or anything else, perhaps with the noble exception of the 'beloved physician', Luke. Such a conclusion would be more than a little unjust. In point of fact, intellectual and emotional functions of the human being are not rigidly anatomically assigned in biblical literature, nor do they need to be in order for its authors to convey what they want to convey. Gordon McConville, from whom I have lifted the examples above, points out that they illustrate 'that the Hebrew writers were working with a kind of unitary understanding of the human that modern readers may not immediately recognize'[15]. While it is an exaggeration to say that terms in the original language (not just in translations) indulge in a rather merry riot, at least they do not keep to an appointed place when it comes to scientific anthropology. Further, anatomical terms in the Bible are not simply used of individuals: the *people* have a heart, just as the people have a memory-bank; *Israel* is to love God with *its* heart, soul and strength.

13 It is true that modern science is discovering that the heart partakes in a form of things we have normally attributed to the brain, but, interesting as this is in light of biblical anatomy, this does not affect the point being made here.

14 For what follows, see especially J. Gordon McConville, *Being Human in God's World* (Grand Rapids, MI: 2016), chapter 3.

15 *Ibid.*, p49.

Old Testament writers plunder their lexical stock in order to convey something other than scientific ideas, whatever scientific ideas go into the mix.

When we come to the New Testament authors, we do not find them about to tighten up anthropological vocabulary in a scientific direction. Paul is the most prolific of them (though Luke wrote more words). In 1 Thessalonians 5:23, he prays that his readers' 'spirit, soul and body be kept blameless', a formulation unparalleled not only in the rest of his literature, but also in the rest of the NT. Commentators on it usually point out that it is virtually, if not absolutely, impossible to distinguish between 'soul' and 'spirit' in Pauline usage elsewhere, and that Paul more generally deploys anthropological terminology in an *ad hoc* fashion. He can use popular language or liturgical language or language adapted to his particular audiences. He can contrast body with spirit, or mind with flesh, with terminological freedom. Briefly to move away from Paul, no early biblical reader would have become scientifically curious when Jesus added the word 'mind' to his quotation from Deuteronomy: 'You shall love the Lord your God with all your heart and soul and strength' (Mark 12:30; Luke 10:27; see Deuteronomy 6:5).

The point of this exercise is to remove a preliminary objection to the attempt to apply biblical anthropology to the question of 'thinking' in a contemporary scientific context. Biblical writers do not rivet their theological anthropology to scientific anatomy. If I were to say today that the heart pumps blood, and tomorrow that the brain pumps blood, you would gently suggest to me that a little consistency and accuracy might do no harm to my scientific street cred. Suppose I responded: 'Who cares – as long as the blood circulates?' Well, you would advise me to grace with my presence another field of study than the study of science and religion. Granted the perils which attend a thought-experiment such as the following, let me venture it anyway. Imagine I sidled up to a biblical writer and said: 'My religious friend, you need to be better informed about what comes out of the heart and what comes out of something called the brain, of which you seem to be lamentably ignorant'. Then I proceed to offer anatomical instruction. He (assuming that all the biblical writers were male) would surely respond: 'Most interesting, my scientific friend. However, it makes not a whit of difference to anything I want to say. My thoughts and desires come from *somewhere* in me in my embodied self – that's how I experience it. Now let's talk about what I am talking about: good and bad desires; right and wrong thoughts. What matters to me is what goes on in that brain of yours when it whirs around thinking, if that's what happens, or whirs around in some sort of connection or correlation with your thinking, if that's what you are saying. You know all about the mechanics of thought; let me tell you a little about its direction, about the disposition of heart or mind, about openness and a truthful relationship to God, and how you need to let your brain whirr or chatter away or whatever it does, to that end.' The language of biblical anthropology can operate in a similar way to the language of biblical cosmology. Talk

of the sun rising is observational or experiential, and not scientific, language, just as it is with us. So it may be with biblical anthropological language; it is characteristically observational, again as it is with us when we talk about our hearts leaping or in our throat.

It is not that primitive science is entirely absent from biblical anthropological language more than from its cosmological discourse. Rather, in order to understand it, we need to attend to what the biblical writers want to put across. It is of permanent theological importance, whatever their notions were about bodily processes involved in thinking. They regard humans as essentially religious beings, whose rational faculties and moral agency are designed by the Creator within *and only within* the bounds of their religious constitution. That is just an abstract way of parsing the affirmation that humankind is made in the image of God. It is made in the image of God, not just as some generically religious kind of being, but as specifically related to the one true and living God in a way that comes to clearest light in Jesus Christ. Created, humankind is to exercise its faculties in proper dominion over the earth. Fallen, it is to exercise its spirit in repentance and faith. Created and fallen, it is to exercise its energies in the love of God and neighbour. Humankind is constituted as an entity for this end; the brain has been so devised by its creator that, when it functions properly, it is directed to those ends, whatever it is up to in its proximate daily operations. The brain is constituted as a religious organ in the sense that it is centrally located within a being made in the image of God. If science instructs us that desiring, loving and thinking have anything at all to do with the brain, then theology instructs us that desire must be healed and re-directed, love renewed and empowered, and thinking adjusted to those ends, so that brain activity, in its daily and living dynamic, is also suitably conformed to the love of God and neighbour.

What bearing does this have on the question of whether machines can think? After all, I have drawn attention rather to the grand overarching purpose of human thought, and the instrumentality of the human brain in that context and to that end, than to the mechanics of its operation. I have been talking in unscientific, teleological, not in causal, scientific terms. If science uses teleological language, it is usually with reference to the kind of immanent material processes which science routinely investigates and describes. Conversely, I have brought into the equation something different - talk of a transcendent purpose, which belongs to a different and entirely distinct domain from the scientific. So how does talk of a transcendent purpose for humankind - whether it is right or wrong on its own terms - tell us anything about the physical operations of the brain and enable us to contribute something to the question of whether machines can think?

It does so because Christianity is not Deism.[16] According to Christianity, the human brain is constituted by its Creator not only so that it is generically religious, in the sense I have described, but also so that there can be direct and distinct communication between God and each particular person. God addresses and is addressed by distinct individuals, not mere tokens of generic humankind, when they are thinking, praying or doing anything else. Nor is humankind an aggregate of such individuals. Individual humans are brought to their fulfilment in the corporate body of the Church, itself addressed by and addressing God in thought, prayer and worship in historically particular, unrepeatable spaces and times. Pivotal in the human being whose destiny is communion with God through Christ and the church, the brain is crafted for this end and wired to attain it. The evolved human brain is ecclesially wired. This is not the scientific scandal which it may immediately seem to be. With scientific rigour, Dan Siegel has led the way in the field of inter-personal neurobiology, showing how the brain is wired not just to inter-personal connection but to inter-personal concord.[17] From a neurobiological point of view, talk of ecclesial wiring is a religious adumbration of the nature of brain inter-connection that is scientifically demonstrable.

All this has implications for the way we think about AI in general and thinking machines in particular. Nothing is or can be greater than God. Nothing is or can be greater for humankind or for any created entity than knowing God in the dynamic of personal relationship. Any created intelligence that does not *pivot* on this reality – that is, that is not constituted on its basis - is both different in kind from human intelligence and, if we have some reason to use the comparative language of greater or superior, is vastly inferior to it. These two things require comment: (a) different intelligence and (b) inferior intelligence.

We take difference first. If human thought, in its broadest sense, can in practice proceed atheistically, not reckoning positively or explicitly with God at any point at all, it possesses that power and possibility only because the human brain is positively and explicitly organised by its creator for an opposite end – religious power and possibility. To say this is just to apply to the brain a Christian perspective on evil and the created good. Evil is not created by God; it is the perversion of the good, the exercise of created powers to an end other than a created end. I can use my hand to strike and wound only because it has the power to comfort and help. Obviously, the mechanism and power of the hand is all the same, whether it harms or comforts. But if it is intended by God for comfort and help, the power which enables it to achieve that end also enables it

16 I use 'Deism' in its loose and popular sense (imprecise by the standards of historical Deism) - belief in a God who created the world but subsequently absented himself from its mundane operations.

17 Daniel J. Siegel, *The Developing Mind: How Relationships and the Brain Interact to Shape Who We Are*, 3rd edition (New York, NY/London: Guilford, 2020).

also to achieve its opposite, a harmful, end. So it is with the brain. The brain can go about its business in all the departments of life with no positive religious reference whatsoever to God, only because it is empowered to operate as the brain of a human being destined to glorify God, made in his image.

Of course, there is talk of explicitly religious AI. Assuming, just for the sake of argument the possibility of robots thinking at all, then if a robotic priest thinks, he or she thinks to a religious end. However, the religious possibility here is contingent. That is, AI is not designed as constitutionally religious. It is not designed so that its thumping electronic heart (to speak poetically) or its chattering electronic brain are related to the living God. Nor could it be so devised. How could any programmer know how to do that? Humankind is religiously programmed; its basic constitution is religious – in the image of God. If some entity has only the constitutional *possibility* of taking up religion or entering an existential religious relationship, it is simply not human or remotely close to being human. An AI whose intelligence is not constitutionally religious is thus an entirely different kind of intelligence from human intelligence. If it thinks at all, it is not wired to think in the way humankind is wired to think, the mechanism of the human brain being designed so that it can receive and respond to a word from God. Of course, many in the world of AI will be entirely unperturbed by this conclusion They will say that they are not trying to achieve an intelligence which replicates or simulates human intelligence, or to produce machines that think as humans think. Indeed so, and it is good to be reminded of this in light of much popular misrepresentation of what everyone in AI is up to. Then all I have done above is to radicalise the difference between human and machine thought, invoking theological, rather than scientific (or philosophical) reasons for doing so.

Talk of radicalising the difference takes us to the question of comparison – superior and inferior, or whichever comparative terms we deploy. It may be claimed that an overall human-machine comparison in such terms is both inappropriate and irrelevant. Perhaps it is appropriate in specified areas for humans and machines; perhaps it is appropriate overall where humans and some non-human animals are concerned. But it is not appropriate overall when humans and machines are compared. So be it; there is no need to insist *a priori* on competitive comparison. However, questions of overall superiority and inferiority do come up in some contexts in connection with AI, for example when discussion of it veers in the direction of transhumanism. The putative post-human being is radically superior to *homo sapiens sapiens* overall, and not just narrowly in particularised intelligence, ability and thought. It is in a different and higher class. But when Christians talk about the highest, they talk about being made in the image of God; when they indicate the exemplar of humanity, they indicate a man who suffered under Pontius Pilate; when they aspire to live the life to which humans are called, they speak of self-denial, taking up the cross and washing others' feet. They may not opt to use the language

of 'superiority' more than they do the language of 'programming', but, if they are impelled to use it, they will say that when body or soul or spirit or mind or brain or heart is programmed for discipleship to Jesus Christ through the love of God and neighbour, it is programmed for the highest which human or non-human or post-human being can achieve.[18]

But Can Machines Think?

Hitherto, I have skated over the question of whether machines can think at all. Is that possibility on the theological table? To what should we be theologically committed? Part of the difficulty with addressing this question has been touched on earlier: there is a range of ways in which the notion of thinking may be understood in and beyond the field of AI. For purposes of discussion, let us take 'thinking' in what may be termed a strong sense, supposing an analogy with or similarity to human thinking, entailing consciousness. As far as I can see, Christians have no *prima facie* theological commitments on this AI possibility one way or another. Christianity directly commits me to some beliefs, but belief in the possibility or impossibility of thinking machines is surely not one of them. Of course, Christian scientists and Christian philosophers may have something definite to say about the possibility. There is a parallel here to other scientific questions. Christianity as such does not commit us to or against the view that the earth moves round the sun or that humans are an evolved species, but our science may commit us. On scientific or on philosophical grounds, the possibility of thinking machines might be convincingly ruled out or in by scientists and philosophers. Christian scientists and philosophers who are of their number will try to join up the lines with their Christianity so as to illuminate aspects of Christianity for those who are neither scientists or philosophers. This contributes to shaping Christian thought, just as in the cases of heliocentricity and evolution. But it is from the scientific (or philosophical) direction, and not from a theological one, that light will come, if at all, on the possibility of thinking machines. I doubt if Christianity can generate a principle out of its inner theological resources which guides us one way or the other.

With respect specifically to machine consciousness, which I connected above with the question of machine thinking, my theological agnosticism echoes that of a neuroscientist and influential Christian thinker of an earlier generation, Donald Mackay, one-time Research Professor of Communication and Neuroscience at the University of Keele.[19] The agnosticism is not comprehensive. The processes of thought in a being not made in the image of God may exhibit mechanical similarities to processes of thought in beings not so made – here a

18 I plead the right to omit discussion of angels.

19 Donald M. Mackay, *Brains, Machines and Persons* (Grand Rapids, MI: Eerdmans, 1980), 62-65. I am not committed to the overall scheme within which Mackay advances this position.

comparative study of humans with non-human animals is relevant. However, I have argued that, if the human brain is wired to receive and reciprocate communication from and to God, then the thinking of a created non-human entity, if it happens at all, will be of a very different kind from the human. I am not theologically agnostic there. Of course, the claim that brains think at all is contentious, to say the least: some will insist that it is persons, rather than brains, who think. They may be right, but, if so, this is surely a scientific or philosophical, and not a theological, proposition, immensely significant as it may be.

My theologically non-committal position on thinking machines is open to at least three objections. Firstly, there is the one which Alan Turing stated. Is it theologically possible that it *is* the soul that thinks; that machines cannot think because they have no soul; but that Turing's response to that objection fails? I believe that Turing's response does fail, because divine omnipotence is a subtler, far more involved matter than he assumed, and he seemed to think of the soul in Christian thought as a kind of substance which can be slapped onto any entity as God pleases, such as an elephant. The question of omnipotence and creation lies outside the bounds of our present discussion, but it does not need to be addressed anyway, for what is pertinent is the basic question about the nature of the soul. It would be theologically very demanding, to say the least, to demonstrate that we should regard the human soul as the organ of human thought. This is partly indicated by earlier remarks on the soul as a disembodied entity, from a biblical point of view. But even if we concluded that the soul is such an entity, that would not of itself demonstrate that it is the organ of human thought. More narrowly, even if persons are capable of disembodied existence in some interim state between death and resurrection, and we speak of them just like that - as persons and not as souls – that surely does not tell us much about how human thought is carried on in this world. To say the least, it is theologically unsafe to eliminate the possibility of thinking machines by appealing to the role of the soul in thinking. Of course, discussion of this question is rather bedevilled from the outset, because the concept of the soul in our discussion is liable to slide around in accordance with different uses of the word - biblical or theological, as the case may be, and in accordance with the particular modern language in which we are carrying on the discussion.

Secondly, suppose we turn in an alternative and opposite direction - to the body, and away from the soul. Does Christianity not understand humans to be *essentially* embodied in a biological way different from machines, and so does not Christianity regard the very *possibility* of thinking as depending on our possessing bodies of flesh and blood in this life and resurrected bodies in the next? Again, I doubt if we can arrive at definite conclusions about the possibility of machine thought by exploring the Christian understanding of the body. Of course, to be human in this life is, by theological definition, to be embodied. And although the apostle Paul strongly emphasises the discontinuity between our future, resurrected bodies and our present ones of flesh and blood (1 Cor-

inthians 15:35-50), life on the new earth is apparently not disembodied life in his or his apostolic colleagues' thought. Scientifically or philosophically, we may say that it is impossible to simulate or replicate the operations of the human brain under silicon, rather than carbon conditions (and I have noted that many in the field of AI are not trying to do so). However, that is precisely a scientific or philosophical point. It is not clear to me that Christian faith as such can be prescriptive in the matter. Thus, while we may reiterate our theological reasons for the conviction that machine thinking amounts to something fundamentally different from human thinking, it is surely hard to rule out machine thinking as such on theological grounds to do with embodiment. Of course, what embodiment or disembodiment actually means where AI is concerned is a question in itself. Quantum computers would doubtless have to be hauled into the discussion, with accompanying talk of 'quintelligence'.[20]

Finally, we must ask whether Christians are not bound to take a theological position on the possibility of robotic thinking which entails robot consciousness, because of current discussion of robot rights. If it is even possible that robots are sentient, as well as thoughtful, creatures, should they not have certain legal rights, such as the right not to be harmed? Can Christians afford to be theologically agnostic on the possibility of machine thought where such questions of practical ethics are concerned? Lawyers are occupied with them; we are not dealing with science fiction.[21] In response, we must agree that theology certainly has the *indirect* task of determining the possibility of robotic thought in order to formulate a position on robot ethics. It is always obligated to take on board what may be scientifically known, philosophically argued, or legally proposed. Christians will agree that pain should not usually be caused to any being which can experience pain and that there should usually be legal safeguards against that. However, it does not follow that this is directly a task for theology - that, theologically, we need to take a definite position on the question of whether AI can think or be conscious just in order for us to come up with a position on whether there should be robot rights. An accumulation of insights from different areas or disciplines, especially Christian ethics with its discussion of rights, might lead us to a theological conclusion on thinking machines. But I am not yet convinced that we can move to that conclusion directly from theological premises.

The advantage of adverting to the question of robot rights is that it shows

20 Paul Davies contextualises Frank Wilczek's ideas here in Brockman, *What to Think*, 29-30. Wilczek himself fleetingly puts quantum AI in context at the end of his essay on 'The Unity of Intelligence' in John Brockman (ed.), *Possible Minds: 25 Ways of Looking at AI* (London: Penguin, 2019), 64-75.

21 Long before citizenship was granted to a robot (Sophia) in Saudi Arabia in 2017, there had been international discussion at high political and legal levels of the status of robots or electronic persons.

that our exercise in asking the question whether machines can think has not been an exercise in shadow-boxing. It is of practical importance. Speaking generally, AI marks out yet another sphere which advertises our collective social uncertainty, at least in developed countries, about what it is to be human. Some of the most pressing questions forced on us by AI – such as concern about autonomous weapons, surveillance or employment – are independent of the question of whether or not a machine can think. It is certainly true that the question of whether or not machines can think can be a speculative distraction from the range of social and cultural issues which AI raises. Yet, the issue of robot rights is just one indication that the question of thinking machines is not an abstraction and need not be a distraction, even if some of us will take the issue as evidence that we have now reached a social point where we no longer know or feel what is real. We may conclude this for other reasons, of course. We talk of a post-truth society; much of the talk surrounding AI reveals that we are also living in a post-reality society. The advent of metaverse technology, which will enable us to hook up to, access, and indwell multiple virtual realities with accelerated ease, may just compound that conviction.

If we are to grasp the significance of the question of thinking machines, in the form I have raised it, we must notice a confluence of streams. There is the age-old fascination with the artificial animation of the inanimate; there is the pre-Christian philosophical tradition which favours rationality as the mark of humanity; there is the scientific revolution of the seventeenth century which drove European thought to explore the possibility that humans are machines – thinking machines already exist: you and me! There is the Darwinian or neo-Darwinian outlook that generated a new sense of what it is to be human; there is the technological urge to all-round improvement, at least of those of us in socially and economically privileged situations; there is the arrogation to ourselves of the right – even responsibility – to change our bodies or to construct intelligences. A Christian response to the question of thinking machines has to take all this into account and chart all these waters. I have not even dipped my toe in them.

As a matter of the most fundamental professional principle, scientists have long asserted that the work of science must proceed untrammelled by unscientific considerations – under no constraint by religious beliefs, for example. This is an assertion that has to be dissected (itself a scientific exercise) for us to understand its ramifications and consider it from all angles. Speaking generally, there is no need to quarrel with the spirit or the instinct which informs this conviction, as long as that spirit or instinct is properly refined. At the same time, it is a scientifically unimpeachable truth that anything we know or believe with good reason in any sphere must, if possible, cohere with what we believe or think we know about anything else. If Christians have reason to espouse Christian faith, it would be unreasonable for their Christianity not to affect their science at all, just as it would be unreasonable for their science

not to affect their Christianity at all. This is admittedly a bland and general observation, but elaboration is simply not possible here. What I have sought to argue is simply this: Christians in science are surely free to explore in theological freedom the possibility that machines are able to think, but they are also bound to hold that any thinking which is not ultimately unto God is infinitely less grand and glorious than the thinking to which humankind is summoned in unity with Jesus Christ in his Church.

Stephen Williams is Honorary Professor of Theology at Queen's University, Belfast