

# Philosophy, Scene XVII: Enter the Automaton<sup>1</sup>

Hanoch Ben-Yami

I describe the influence of clockwork automata on modern philosophy and science, particularly on the conceptions of life, soul and mind.

Living creatures, be they animals or plants, are evidently different from nonorganic nature: they grow, they reproduce, animals move themselves and have sensations and desires. Throughout history, people have tried to account for these characteristics of life, and in Ancient, Medieval and Renaissance philosophy the explanation was provided by the soul, the principle of life. Here is Plato on life and soul:

Tell me, then, what is that of which the inherence will render the body alive?

The soul, he replied.

And is this always the case?

Yes, he said, of course.

Then whatever the soul possesses, to that she comes bearing life?

Yes, certainly. (*Phaedo*, 105cd)

And Plato's disciple, Aristotle, whose views were very different from his master's on very many issues, agrees on that:

What has soul in it differs from what has not in that the former displays life. (*de Anima*, 413a22)

As mentioned, life has many distinct characteristics. But scientists and philosophers, who tried to find something essential to all of them, saw the ability to move oneself as such. Here is Plato again:

Self-motion is the very idea and essence of the soul ... The body which is moved from without is soulless; but that which is moved from within has a soul, for such is the nature of the soul.

(*Phaedrus*, 245c–246a)

From Plato and Aristotle to the Renaissance, this view of the soul remained unchallenged. The further, different capacities of living creatures were explained by seeing the soul as tripartite: it has a vegetative part, responsible for growth and reproduction; a sensitive part, responsible for sensation and motion; and a rational part or mind, unique to humans.

---

<sup>1</sup> The article was published in Italian translation as 'Filosofia, scena XVII: entra l'automa' in *La Voce di Hora* 42, June 2017: 67–74. It is based on a talk delivered at a symposium on Janello Torriani, on 3 Dec 2016, organised as an accompanying event of the exhibition, *Janello Torriani, Genio del Rinascimento*, in the Museo del Violino, Cremona.

We move now to the sixteenth century and to the invention of clockwork automata. I shall not present the history of the invention or the automata's mechanisms, but discuss the impression they left on people of the period.

Although sixteenth century automata could do only very simple things compared to later ones, such as Vaucanson's or Jaquet-Droz's in the eighteenth century, not to say those of our times, for the people of the period they were a new, exciting and amazing thing. Artefacts that contained within themselves the power to move themselves were unprecedented. There wasn't even a word to describe such things: the word 'automaton' was adopted from the Greek, in which it had a different albeit related meaning, only during the sixteenth and seventeenth centuries. It's first recorded use in English, for instance, is as late as around 1625, in the play *The Bloody Brother* (IV.i). The excitement that automata produced is documented in the writings of the period.

Here is a passage purporting to describe the automata of Janello Torriani and the response they evoked:

For often, when the Cloth was taken away after dinner [from the table of Emperor Charles V], [Janello] brought upon the board little armed figures of Horse and Foot, some beating Drums, others sounding Trumpets, and divers of them charging one another with their Pikes. Sometimes he sent wooden sparrows out of his chamber into the Emperor's Dining-room, that would fly around, and back again; the Superior of the Monastery, who came in by accident, suspecting him for a Conjuror.

The passage is taken from Famiano Strada's book, *De Bello Belgico* (Book I, p. 7), published in 1632. It is unreliable: we don't have any evidence that Janello made any automaton of the kinds ascribed to him by Strada, and the flying sparrows were far beyond his capacities. But the passage is interesting precisely because it is unreliable. We see that in less than fifty years after Janello's death, people already imagined automata as capable of any manner of self-motion, including flight.

The passage is interesting also because of the response to the encounter with automata it ascribes to the superior of the monastery. He took Janello's automata to be not machines, but things brought to life by magic. Although the story about the superior might be fabricated, the reaction ascribed to him is reliable: while people of the period were not familiar even with the possibility of an automaton, they believed in fairies, dwarfs, or even homunculi created by alchemy, which Paracelsus was believed to have created. Consequently, when the uninitiated first saw one of these automata, they often took them for living creatures.

Descartes (1596–1650) turned things around: he looked at living creatures and thought they are automata. A passage in his *Meditations* documents how he was impressed when first seeing automata:

According to my judgement, the power of self-movement ... was quite foreign to the nature of body; indeed, it was a source of wonder to me that certain bodies were found to contain faculties of this kind. (AT VII 26)

These bodies were of course the automata. Early in his life, Descartes also became familiar with automata's technology. In his early *Cogitationes Privatae* (~1620) he suggests in outline constructing an automaton of a tightrope walker moved by magnets (AT X 231). Poisson, one of his acquaintances,

writes that Descartes planned or tried to build automata, and in his plans surpassed Torriani (AT X 231–2, footnote b) – this is surely an exaggeration, but it does show that Descartes studied automata.

Inspired by the possibilities of automata, Descartes concluded that the ability to move oneself does not require a principle which does not exist in inanimate nature. And since philosophy and science have characterised life as primarily this ability, he similarly concluded that life does not involve any principle of that sort. Animals are therefore natural automata, superior to manmade ones because made by a superior designer, God:

This will not seem at all strange to those who know how many kinds of *automata*, or moving machines, the skill of man can construct with the use of very few parts, in comparison with the great multitude of bones, muscles, nerves, arteries, veins and all the other parts that are in the body of any animal. For they will regard this body as a machine which, having been made by the hands of God, is incomparably better ordered than any machine that can be devised by man, and contains in itself movements more wonderful than those in any such machine. (*Discourse*, Part Five, AT VI 55-6)

Since the soul was conceived as a special principle of life, the conclusion from the mechanisation of life was that there is no special principle, a soul, responsible for life. And Descartes therefore writes, when discussing all the functions of the human body:

I should like you to consider that these functions follow from the mere arrangement of the machine's organs every bit as naturally as the movements of a clock or other automaton follow from the arrangement of its counter-weights and wheels. In order to explain these functions, then, it is not necessary to conceive of this machine as having any vegetative or sensitive soul or other principle of movement and life, apart from its blood and its spirits, which are agitated by the heat of the fire burning continuously in its heart – a fire which has the same nature as all the fires that occur in inanimate bodies. (*Man*, AT XI 202)

Human beings do have a soul or mind, Descartes thought, but it is not the source of life. And to conclude the rejection of a two-millennia tradition, he asserts, contrary to Plato, that we die not because the soul leaves the body, but that the soul leaves the body because we die:

It has been believed, without justification, that our natural heat and all the movements of our bodies depend on the soul; whereas we ought to hold, on the contrary, that the soul takes its leave when we die only because this heat ceases and the organs which bring about bodily movement decay. So to avoid this error, let us note that death never occurs through the absence of the soul, but only because one of the principal parts of the body decays. The difference between the body of a living man and that of a dead man is just like the difference between, on the one hand, a watch or other *automaton* (that is, a self-moving machine) when it is wound up and contains in itself the corporeal principle of the movements for which it is designed, together with everything else required for its operation; and, on the other hand, the same watch or machine when it is broken and the principle of its movement ceases to be active. (*Passions* I 5-6, AT XI 330-31)

As seen in these passages, to convince his readers of his view Descartes recurrently compares the living creature to a clockwork automaton. This is the first time in history in which such a comparison to an

automaton was made. And the comparison was powerful and effective. The philosophers and scientists adopting Descartes' view, and these were the great majority in the generation following his, also argued for the mechanical conception of life by comparing living creatures to automatic artefacts with which they were familiar. I shall give two examples.

Sir Kenelm Digby (1603–1665), the English courtier, diplomat and natural philosopher (among other things), was one of the first to take Cartesian ideas across the channel. In his *Two Treatises* of 1644 he describes living creatures, namely those organised bodies that contain within themselves the principle of their own motion, as natural engines (204–205). God is compared to a clockmaker, who, once he has laid 'the foundations of his designed machine', namely the living organism, lets it operate on its own according to mechanical principles, he being 'an improvident clockmaker, that should have cast his work so, as when it were wound up and going, it would require the master's hand at every hour to make the hammer strike upon the bell' (pp. 226–227). And to convince his readers in the validity of his comparison of living creatures to automata (a term he does not use) he describes two 'engines' he saw in Spain,

the one at Toledo, the other at Segovia: both of them set on work by the current of the river, in which the foundation of their machine was laid. That at Toledo, was to force up water a great height from the river Tagus to the Alcazar (the King his palace) that stands upon a high steep hill or rock, almost perpendicular over the river. In the bottom, there was an indented wheel, which turning round with the stream, gave motion at the same time to the whole engine... (205-207)

The engine is of course the artifice of Janello. It is remarkable that two very different achievements of this greatest of Renaissance engineers, clockwork automata and hydraulic machinery, independently inspired and convinced different philosophers in the validity of the mechanical conception of life. We witness here the deep influence of technology on human thought.

Giovanni Borelli (1608–79), in his posthumous *De motu animalium* (1680–1681), endeavoured to explain mechanically all the phenomena of life that do not involve perception or volition. Echoing Descartes, he compares living creatures to automata in arguing for the mechanization of life phenomena:

An automaton seems to present some shadowy resemblance with animals in that both are organic self-moving bodies which comply with the laws of mechanics, and both are moved by natural faculties. Let us, then, see whether it is possible to investigate the properties of natural [living] things in some way by using our knowledge of artificial ones. (2.8, Proposition 116)

We next find the Cartesian comparison of the living creature to a clock (Ibid.). Unlike Descartes, Borelli *did* claim that animals have a soul, but his animating soul influences only voluntary movements. Consequently, after powerfully arguing that the motion of the heart – the source of life – is involuntary, he concludes that 'the movement of the heart can result from an organic necessity, as an automaton is moved' (2.6, Propositions 78–79).

At this place, one can raise a question: Descartes was himself convinced and convinced others, by means of an analogy with clockwork automata and other mechanical artefacts of his age, that *life* is a mechanical phenomenon. But then, why didn't he or others generalised even further and concluded that *thought or mind* as well is a mechanical phenomenon? Descartes eliminated the soul, the principle

of life, as an immaterial principle: why did he keep the mind, the principle of thought, as such a principle? Why did Descartes become a mind–body dualist, and not a materialist? – The answer is again provided by the automata of his age, but this time by their *limitations*.

Descartes thought that human beings, and generally all creatures that have a mind, have two intellectual capacities which no machine or automaton can emulate. One was the ability to act appropriately in indefinitely many circumstances, the other a kind of use of language. These non-mechanical capacities show, he thought, that there's an immaterial principle in humans responsible for intelligence, a mind.

It is instructive to see how he argues for these limitations of automata. Suppose, he says, that one built machines in the attempt to imitate humans, and suppose these machines 'bore a resemblance to our bodies and imitated our actions as closely as possible for all practical purposes, we should still have two very certain means of recognising that they were not real men' (*Discourse*, AT VI 56).

Even though such machines might do some things as well as we do them, or perhaps even better, they would inevitably fail in others, which would reveal that they were acting not through understanding but only from the disposition of their organs. For whereas reason is a universal instrument which can be used in all kinds of situations, these organs need some particular disposition for each particular action; hence it is for all practical purposes impossible for a machine to have enough different organs to make it act in all the contingencies of life in the way in which our reason makes us act. (Ibid. 57)

Descartes thinks that every particular response of an automaton requires what we today call a *dedicated mechanism*. He is obviously thinking of the automata with which he was familiar: for the automaton on the exhibition to turn right, you should insert into its body a specific contraption; for it to move its eyes sideways, yet another; to hit the tambourine, a third; and so on. Since the automaton has a limited volume (and since mechanisms cannot be indefinitely small), only a finite number of actions can be programmed into its body. And so, unlike an intelligent creature, it can respond appropriately only in a finite number of circumstances. Descartes did not envisage a programmable artefact that can learn and adjust itself to new circumstances. His argument is based on the principles of construction of the automata *of his age*, and not of those of any possible artefact.

Descartes' second reason for making the mind immaterial relies on language use:

Machines could never use words, or put together other signs, as we do in order to declare our thoughts to others. For we can certainly conceive of a machine so constructed that it utters words, and even utters words which correspond to bodily actions causing a change in its organs (e.g. if you touch it in one spot it asks what you want of it, if you touch it in another it cries out that you are hurting it, and so on). But it is not conceivable that such a machine should produce different arrangements of words so as to give an appropriately meaningful answer to whatever is said in its presence, as the dullest of men can do. (Ibid. 56-7)

The dullest of men can respond meaningfully to questions, no machine will ever be able to do that, so man is more than a machine: he has something immaterial which enables him to respond in this way, namely a mind.

But why does Descartes think that it is inconceivable that a machine shall have such a capacity? This is not argued for in this passage, nor anywhere else in his writings; it is merely asserted. The assertion is again based, apparently, on the capacities of the automata with which Descartes was familiar. It was conceivable that automata built on such principles will be able to move in any possible way, and Strada, Descartes and others did imagine that; but the principles of construction of these automata did not provide even the beginning of an idea of a mechanism that will enable them to respond meaningfully to questions. Thinking of *such* automata, Descartes couldn't conceive an artefact which responds meaningfully to questions. Three hundred years later, reflecting on his digital computers, Turing will propose a different answer.

The automata convinced Descartes and later generations that life doesn't require an immaterial soul but that thought required an immaterial mind. In this way, this technological breakthrough shaped our worldview for centuries to come, both through its achievements and through its limitations. For the next great shift, we had to wait, as already hinted above, for another epoch-making technological breakthrough, the invention of the digital computer.<sup>2</sup>

---

<sup>2</sup> I have discussed the subject of this article in much more detail in my book, *Descartes' Philosophical Revolution: A Reassessment*.

References

Aristotle, *De Anima*

Ben-Yami, Hanoch, *Descartes' Philosophical Revolution: A Reassessment*. London, Palgrave Macmillan, 2015.

Borelli, Giovanni Alfonso, *De motu animalium*. Rome, Angeli Bernabo, 1680–1681.

Descartes, René, *The Philosophical Writings of Descartes*, Volumes I–III, translated by J. Cottingham, R. Stoothoff, D. Murdoch and A. Kenny, Cambridge University Press, 1984–1991.

Digby, Kenelm, *Two Treatises*. Paris, Gilles Blaizot, 1644.

Plato, *Phaedo, Phaedrus*.

Strada, Famiano, *De bello Belgico*. Rome, Francesco Corbelletti, 1632.