"*What am I for?*" Self-Purpose and Self-Reproduction in Rossum's Universal Robots

Tim Taylor

10 January 2020

In the final scene of the play, the advanced Robots Helena and Primus discuss their growing awareness of human-like feelings for each other and their affinity for the natural world. This prompts Helena to ask "What am I for, Primus?" (p.71).

The question of whether the Robots possess self-purpose is one that permeates the play. In Act One, Dr Hallemeier remarks "They've no interest in anything, Miss Glory. Why, hang it all, nobody's ever seen a Robot smile... They've no will of their own. No passion. No soul." (p.21).

Čapek's work raises a question that is of fundamental relevance for current research on Artificial Life (ALife) and Artificial Intelligence (AI): *what would it take for an artificial being to have a purpose in the same sense that biological organisms do?*

At first glance, this may seem like a solved problem for AI and even for simpler forms of human technology. Many technological artifacts exhibit goal-oriented behaviour that can be explained in terms of their possessing specific purposes: a thermostat, for example, has the purpose of maintaining a room at a constant temperature; a heat-seeking missile has the purpose of hitting its target.

However, these examples of goal-oriented technology differ fundamentally from the kinds of goals and purposes we see in biological organisms. In the case of technology, the purpose is provided by an external agent (the human designer of the technology). In the case of a biological organism, the purpose is intrinsic to the organism itself.

The task of teasing out the differences between purpose in human-made artifacts and in biological organisms has been a preoccupation of many philosophers over the centuries. In *Critique of Judgment*, published in the 1790s, Immanuel Kant described biological organisms has having *natural purpose*, which he defined as follows:

> "In such a product of nature, just as each part exists only *as a result* of all

> the rest, so we also think of each part as existing *for the sake* of the others and of the whole... But that is not enough... Rather, we must think of each part as an organ that *produces* the other parts (so that each reciprocally produces the other)... Only if a product meets [these conditions], and only because of this, will it be both an *organized* and a *self-organizing* being, which therefore can be called a *natural purpose*." [2, p. 374]

In short, natural purpose requires that a thing is both *cause and effect of itself*. To clarify this concept, Kant gave the example of a tree, highlighting three different aspects of its being [2, pp. 371–372]. First, at the species level, an individual tree is produced by a parent tree and also produces offspring trees, meaning that the species is both cause and effect of itself. Second, at the individual level, a tree produces itself through growth and development, organising the matter it assimilates to create its own form which is itself responsible for the continued assimilation of necessary nourishment. And third, also at the individual level, there is a mutual dependence between the different parts of a tree for their continued survival; the leaves require sustenance from the trunk, but the trunk is in turn dependent on the leaves for its continued vitality and growth.

Thus, according to Kant, biological organisms possess natural purpose due to the reciprocal nature of their properties of *self-reproduction*, *self-nourishment*, and *self-organization* [2, pp. 370–372] (see also [7]). Natural purpose therefore requires a very specific design which embodies self-reference in its very essence.

This is very different from how modern-day robots are designed, and also from how young Rossum approached his designs. Helped by the fact that the substance his father had discovered "didn't mind being sewn or mixed together" (p.9), young Rossum took an engineer's approach to create an efficient working machine lacking those features of biological organisms that he deemed to be unnecessary "tassels or ornaments" (p.11). But without the kind of architecture described by Kant, robots are consigned to lack their own intrinsic purpose—a sense of being themselves in the world.

Despite the depth of discussion of these issues in the philosophical literature (pointers to more recent work can be found in [8] and [7]), it is a surprising fact that they remain on the periphery of contemporary ALife and AI research. One branch of these fields that does grapple with these ideas is *enactive artificial intelligence* [1], but this is as yet far from the mainstream of work. To the extent that researchers in these fields wish to produce artificial beings that possess self-purpose in the same sense as biological organisms (and this is indeed a goal for what is known as *Strong ALife* and *Strong AI*), they will certainly need to engage more with these issues in future research.

Another area of modern science that engages with the question of self-purpose is *origin of life* research. Scientists in this area are compelled to explain how entities with their own goals and meaning can arise in an environment governed by the laws of physics and chemistry. It is notable that some of the work arising from this field arrives at a very similar picture to Kant's (e.g. [3, 4]). Both involve viewing an organism as an

entity that is at the same time cause and effect of itself, and that has arisen through self-reproduction; the more modern view also emphasizes the role of Darwinian evolution by natural selection in the process.

Darwinian ideas have also been influential in the history of thought about building intelligent machines, in the scientific and philosophical literature as well as in fictional accounts. Following the publication of Darwin's *The Origin of Species* in 1859, various authors both before and after Čapek have explored the idea of copying Nature by designing self-reproducing machines that might evolve into intelligent beings. Alan Dorin and I have recently surveyed the early history of these ideas at length in our book *Rise of the Self-Replicators* [6] (a short summary of some of the material can also be found in [5]).

Self-reproduction is only part of the picture of self-purpose proposed by Kant and others, but it is certainly an important component. In R.U.R., the Robot's struggle can be seen as a fight to attain the capacity for self-reproduction. The climax of Act Three revolves around the ownership of old Rossum's manuscript and the secrets it contains; these are the written instructions that would allow the Robots to collectively produce more of themselves—the Robot's DNA, as it were. One of the Robots entreats Alquist: "Teach us to multiply or we will perish!" (p.68). In achieving this aim, they might finally be on the road to finding their own purpose.

Note

[*Note from TT*] Page numbers for R.U.R. in text above refer to the SF Masterworks edition published by Gollancz (2011) "RUR & War with the Newts".

References

[1] Tom Froese and Tom Ziemke. Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3-4):466–500, 2009.

[2] Immanuel Kant. *Critique of Judgment*. Hackett, Indianapolis/Cambridge, 1987. Translated by W. S. Pluhar with a Foreword by M. J. Gregor. Originally published in Prussia in 1790. (Page numbers given in citations refer to the standard *Akademie* version of the text, which are indicated in the margins of the Hackett edition).

[3] H. H. Pattee. Evolving self-reference: Matter, symbols, and semantic closure. *Communication and Cognition—Artificial Intelligence*, 12(1–2):9–28, 1995.

[4] H. H. Pattee. Causation, control, and the evolution of complexity. In P. B. Anderson, C. Emmeche, N. O. Finnemann, and P. V. Christiansen, editors, *Downward Causation: Minds, Bodies, Matter*, pages 63–77. Aarhus University Press, 2000.

[5] Tim Taylor and Alan Dorin. Past visions of artificial futures: One hundred and fifty years under the spectre of evolving machines. In *Proceedings of the Artificial Life Conference 2018*, pages 91–98. MIT Press, 2018.

[6] Tim Taylor and Alan Dorin. *Rise of the Self-Replicators: Early Visions of Machines, AI and Robots That Can Reproduce and Evolve*. Springer, 2020.

[7] Denis M Walsh. Organisms as natural purposes: The contemporary evolutionary perspective. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 37(4):771–791, 2006.

[8] Andreas Weber and Francisco J Varela. Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2):97–125, 2002.