

ARTIFICIAL INTELLIGENCE: A Modern Myth

John Kelly

£34.95, 253 pp. ISBN 0-13-338559-0 (hb)

The title of this work is a slight misnomer, in that Kelly is not suggesting that artificial intelligence (AI) as a whole is a myth, rather the quest for the Holy Grail, a thinking machine, is as mythical as that of King Arthur's Knights. Like the Arthurian legend, it is a myth designed to inspire; the numerous practical AI systems at work in the real world proof of its potency.

Recently there has been a plethora of books questioning the feasibility of strong AI, including works by Penrose (1989), Bringsjord (1992), Moody (1993) and this title from Kelly. All these works have much to offer the interested agnostic, but Kelly's remains the most comprehensive in scope and the most thorough in content. What gives this work its Trojan edge is that it is written by a Computer Scientist with a comprehensive knowledge of what has been, and is being, achieved by the AI community.

Kelly begins his treatise by outlining the main thrust of his argument - to call into question the base assumptions of AI and in particular the use of misleadingly anthropomorphic language. The result of his examination is a trenchant expose of the superficial nature of the attribution of human qualities and behaviours to computers.

Kelly first demonstrates how the classical definition of AI:

“AI is the part of Computer Science that is concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behaviour...”, (Barr & Feigenbaum, pp: 3)

... implies a behaviourist approach to authentication and evaluation of AI systems. Kelly asserts that, using Ryle's terminology, the phrase *Artificial Intelligence* is ‘systematically misleading’ and a ‘category mistake’. In a second reference to Ryle, Kelly also points out an inherent dualism of the strong AI perspective - that intelligence may be realised independent of the underlying material in which it expresses itself: flesh; silicon; light - a Ghost in the Machine, dismissed by Ryle in *The Concept of Mind* (1949). Kelly surveys attempts to define human intelligence and is particularly critical of the use of lists (*the ability to reason; use heuristics etc*). The use of such a collection of properties to define intelligence he defines as the **List Fallacy**:

“No mere enumeration of properties, conditions, manifestations, exemplaries, etc. can capture the coherence, the integral unity of ‘natural’ concepts”, (Kelly, pp: 68)

ie. Lists can be used to describe properties of intelligence, but the inclusion of a particular set of features is not sufficient to construct an intelligence :- the properties remain...

“... logically disjointed unless brought into relationship and integrated by the unifying power of human apperception”, (ibid)

If the essence of human intelligence is difficult to pin down, what about the essence of machine? Before the advent of computing a simple mechanistic definition would suffice.

“A machine is an assemblage of moving parts, constructed for the purpose of transmitting motion or force, and of modifying, in various ways, the motion or force so transmitted.” (Goodeve, pp: 1)

But Turing extends the concept to include computers and even more radically, to computer software. Interestingly as Kelly points out, Godel's Incompleteness Theorem defines limitations to such ratiocination, although he resists Penrose's temptation to say that these results bring to a close the attempt to express human reasoning in machines.

Kelly defines machines as having: a specific function; an explicit design; deterministic operation; automacity and mediacy. He contrasts this to man, highlighting the fact that emotion is fundamental to the human condition, but is not instantiated, Sloman-like, in any machine. Although Kelly acknowledges that there are AI researchers looking into this possibility (eg. Simons, Sloman et al). Without solid results he dismisses their claims as another idle reification, one further example of the **Existential Fallacy** that underpins much of AI:

“The assumption, in the absence of experience or encounter, that an entity conceived under a certain description actually exists.”, (Kelly, pp: 118)

To compound this situation, Kelly also highlights the implicit behaviourism of the most infamous test for machine intelligence, the Turing Test. This can be summarised as:

“If a computer behaves intelligently then it is intelligent”, (ibid)

This, he claims, is yet another smokescreen used by AI researchers to avoid serious analysis of what it actually means to think!

Kelly, alongside Newell and Simon (1976), asserts that symbolic processing is central to intelligent thought, but in a core semantic definition contrasting Newell and Simon, he refutes the idea that computers are symbol processors. With reference to Cassirer, Kelly states that *symbols cannot be reduced to signals*:

“A symbol has no actual existence as a part of the physical world; it has a ‘meaning’ ”, (Cassirer, pp: 57)

Thus computers are not the *symbol-processors* that they are widely believed to be, but *signal-processors* (in the sense of uninterpreted physical event). If the term symbol-processor is used to describe the operation of a computer, then it should be noted that the entities it is processing are symbols for us, its users, not itself. As Searle illustrates with the Chinese Room (1980), formal symbol manipulations do not, *by themselves*, have intentionality or intelligence. A metaphor that Kelly uses throughout the book is that of the computer as a text which in reality cannot *do anything*, can have no autonomy or originality (aka. Lady Lovelace's objection to mechanized thought).

“In essence then, it amounts to a ‘category mistake’ to consider computers intelligent. They are simply not the kind of entities of which it is valid to predicate knowledge or understanding or thinking or intelligence”, (Kelly, pp: 193)

Kelly concludes his work with an examination of connectionist systems, which in the sense of adjusting towards a solution, are more in line with the ideas of later Wittgenstein and Heidegger. But fundamentally such networks can be viewed as texts in the same way as symbolic AI systems, and hence are subject to the same criticisms.

In summary, the strength of Kelly's treatise, and one that makes the book difficult to review, is that AI is not attacked on any one front - there is no single argument to prove that Strong AI is impossible. Rather, Kelly meticulously examines the claims, achievements and underlying philosophy of the proponents of AI, highlighting the limitations of each theory, with the effect that reading the book is rather like being confronted by a slow moving traction engine. By the end all but the most religious practitioners of AI will have moved towards apostasy.

**Dr. J.M.Bishop,
Department of Cybernetics,
University of Reading,
Whiteknights, Reading, Berkshire,
United Kingdom.**

References

- Barr, A. & Feigenbaum, E.A., (Eds), (1981), *The Handbook of Artificial Intelligence Volume 1*, Pitman.
- Bringsjord, S. (1992), *What Robots Can And Can't Be*, Kluwer Academic Publishers.
- Cassirer, E., (1944), *An Essay on Man: An Introduction to the Philosophy of Human Culture*, Yale University Press.
- Goodeve, T.M., (1880), *The Elements of Mechanism*, Longmans, Green & Co.
- Moody, T.C., (1993), *Philosophy and Artificial Intelligence*, Prentice Hall.
- Newell, A. & Simon, H.A., (1976), *Computer Science as Empirical Enquiry: Symbols & Search*, CACM, **19**.
- Penrose, R., (1989), *The Emperor's New Mind: concerning computers, minds and the laws of physics*, Oxford University Press.
- Ryle, G., (1949), *The Concept of Mind*, Hutchinson.
- Searle, J., (1980), *Minds, Brains & Machines*, The Behavioural and Brain Sciences, **3**, pp. 417-24.
- Turing, A. (1950), *Computing Machinery & Intelligence*, Mind, **59**.