

A Review “A Jurisprudence Of Artefacts”

Kane Oliver*

University of Sussex, England

Last year, Blay Whitby and I wrote a paper entitled “How To Avoid A Robot Takeover: Political and Ethical Choices in the Design and Introduction of Intelligent Artifacts”. (Whitby & Oliver, 2000). It is quite a short paper with a very long title. Frank Sudia’s paper should remind us that the title of any revision will need to be even longer as we should have included ‘Legal Choices’. Our paper was a response to those doom mongers (Warwick (1997) in particular) prophesying the inevitable destruction of humanity by intelligent machines. The message was simple - ‘robot takeover’ may be possible but it is not inevitable. There are many, many choices to be made. “A Jurisprudence of Artefacts” goes further and sketches in more detail the form that some of those decisions might take.

To those who are already convinced that future artefacts should be accorded rights (the Liberals), Sudia’s paper is a consideration of how to bring about that accordance. To a Conservative, however, it would appear that Sudia has provided a *reductio* against the notion that corporations should be treated as persons (as if one were really required). On the Conservative reading the argument runs something like this:

Corporations can be treated as persons legally. Therefore non-humans could be treated as persons legally. That which is very like a human person and is a person legally should be treated as a person in a moral sense.

The slide from machine to person occurs in two stages 1) the slide from machine to legal entity, 2) the slide from legal entity to moral entity. There are two key points in the slide from machine to legal entity:

* Many thanks to Frank Sudia for Comments on a previous version of response that did much to make my own unwarranted assumptions clear.

Correspondence: the Philosophy of Cognitive Science at the School of Cognitive and Computing Science (COGS), at the University of Sussex, England.

Email: Kaneo@cogs.susx.ac.uk

1. The claim that machines can act on behalf of or as the agent of a human person.
2. The enactment of permission for the artelect to serve as its own sole officer and director.

As these two points are doing so much work in the argument I would have liked to see them discussed in greater detail. In what way can machines act on our behalf or as our agent? Are there cases where the law distinguishes machines from people in cases such as these, and if so, why? With regards to the second point I have two questions. Firstly, why would the measure ever be taken (other than as a way of tidying up legislation following a Liberal victory in the debate)? Secondly, it is not clear that such a measure could be taken in the absence of a Liberal victory, since an artelect might not be the kind of thing that could serve as an officer or director.

The slide from legal entity to moral entity is not argued for. Instead the paper is about how a Liberal could set about making the legal position match up with their moral judgement. But of course the current or future legal position can have no bearing on the outcome of the philosophical arguments. The paper, nonetheless, raises some fascinating points, 1 but it is difficult for a non-Liberal to comment on the paper since it is not written with them in mind. The text is littered with language that begs the big questions. Words such as 'belief', 'desire', 'dignity', 'agency', and others all remain subjects for strong debate in the philosophy of mind and cognitive science. What is difficult to take is not that Sudia does not engage with those debates (that would be a very different paper), but rather that those debates are not even acknowledged. Talk of humans being "an obvious waste of resources" and the disclosure of non-human status being an unjustified discrimination appear as presuppositions rather than being argued for. A Conservative would argue that humans are a necessary component of the corporation because a person must be there to take responsibility. Humans can only be dispensed with if we can claim that the artelects are persons in a moral sense, and no argument is given for such a claim. 'Disclosure of non-human status' would be unjustified discrimination only if the artelects were persons a moral sense. The question at issue is whether disclosure of non-human status equates to the disclosure of non-moral-person status and whether that invalidates any contract.

There is also a question about the constitution of artelects and corporations - the person status of corporations seems to bottom out in the person status of its parts, i.e. employees and directors (this is the only way anyone could make sense of Santa Clara on anything other than the rather

strange grounds of convenience). When a charge of corporate manslaughter, say, is brought against a corporation, the chief executives or senior managers are surely punished on the grounds that they are responsible by commission or omission for the failings of the system. If the executives were not the kinds of beings who could be punished in this way then the charges against them would be senseless. So we have to ask whether artefacts are the kinds of thing that can be punished in the required way. Even if we claim that artefacts could have experiential states (a difficult claim to make without argument) the question of whether they could be responsible for their actions or that of the corporation would remain.

So, could there be a machine that could be responsible for its decisions (the claim that they can make decisions is perhaps only slightly less controversial) - not in the sense of being merely causally implicated in a strong fashion but in the sense of requiring praise or blame for its actions? As a non-Conservative I think that there could be such machines. However, as a non-Liberal, I disagree with Sudia on the range of artefacts which might qualify. Sudia appears to allow current, easily available AI (GOFAI for want of a better term) as well as the more recent paradigms of connectionism and evolutionary computing. This is more liberality than his argument actually requires (see Bechtel 1985 for a treatment of the difficulties in attributing responsibility to GOFAI machines).

One final remark: my instinct is that the battle for robot rights will not be sponsored by the corporations. One does not have to be a rabid anti-capitalist to ask how often have corporations attempted to instigate a change in the law in favour of the rights of their workers? Moreover the philosophical driving force behind the claim for artefact rights comes from the possibility of incredibly powerful machines which have not been designed (in any strong sense). One has to ask just how keen corporations will be to put on the market powerful products that are inherently unpredictable and trust that insurance companies will pick up the cheque in the event of any difficulties? Much of Sudia's main argument still stands. The way is still open for some other party to invoke Santa Clara (or something like it) in a fight to secure rights for artefacts. The road to emancipation might travel through the courts in an argument over the moral status of artefacts produced accidentally and not for profit.

Note

1. The question of intellectual property is particularly interesting. Once again I believe that it is still an open question whether or not artefacts would be the kinds of thing that could have rights. However, leaving that to one side for now, my intuition is that the assumption of localisation is doing a lot of work in motivating Sudia's claims. The paper does raise the issue of IP in the context of multi-national corporations and, as an expert in IP law, Sudia is best placed to discuss the impact of new technologies in this context. But as the assumption of localisation might be strongly analogous to the situation prior to the rise of multi-nationals I would have liked to see more on what would happen to IP law if the assumption of localisation were to be relaxed.

References

- Bechtel, W. 1985. "Attributing Responsibility to Computer Systems", *Metaphilosophy* Vol.16 No.4 October.
- Warwick, K. 1998. *In the Mind of the Machine: The Breakthrough in Artificial Intelligence*. London: Arrow Books.
- Whitby, B. & K. Oliver. 2000. "How to Avoid a Robot Takeover: Political and Ethical Choices in the Design and Introduction of Intelligent Artifacts." *AISB Quarterly* No. 104 Autumn.