02/08/2018

Science and Tech Essay

30%

2000 words

In what ways is artificial intelligence used to question the current predominance of mankind in *The Talos Principle,* and how does this relate to the 'posthuman' evolutionary concept*?*

In *The Talos Principle* (2014)*,* the posthuman world is an optimistic vision that draws an evolutionary relationship between humanity and androids. A game developed by CROTEAM, Tom Jubert and Jonas Kyratzes ('the head-writers') use the Talos Project to create an equivalency between artificial intelligence and human-intelligence through the lens of free will, feeling, thought, and empathy (Ostry, 2004). Through this, the head-writers dismiss anti-humanist and critical posthumanism , and opt for a positive anthropocentrism that recognises the value of 'Man's very soul'. This soul is not wholly empirically-minded, and so this optimistic posthuman still retains the value of faith. Simultaneously, it is a realist approach in that it rejects some transhumanist ideals and recognises the inevitable demise of humanity; instead, basing the posthuman world in 'letting go of your desire to bear witness', but collectively surviving. The head-writers use Alexandra Drennan's archived messages to frame the choices of the Android within philosophical questions that logically realise the notions of free will, thought, and empathy within the Android character.

The thesis-metaphor: humanity is not biology

The head-writers use the Greek myth of the clay-giant Talos as their thesis-metaphor to deconstruct our understanding of 'human':

"[The giant Talos] was made of bronze, and had but one vein,

within which flowed [ichor] … the loss of that liquid caused him to

die [like] blood. May we not then say that Talos… had all the

essential properties of [a human]? He moved… He spoke… had

wishes… If, then… [a machine may act as a man] while driven by

only the ingenious plan of its construction and interaction…

according the principles of nature, then does it not follow that man

may also be seen as a machine?" (Jubert, Kyratzes, 2014)

This myth queries the importance of the *physical* state of humans by analogy of blood and ichor and flesh and bronze, and it introduces the notion that 'human' would not be defined by biology in the posthuman world. Drennan makes reference to IVF, experimental cloning technology in how the physical concept of 'reproduction' is already heavily controlled by machines (Jubert et al., 2014). Haraway (1985) echoed this notion: 'We are all chimeras, theorized and fabricated hybrids of machine and organism' (50). Having dismissed the biological basis for a definition of humanity for the posthuman world, Drennan turns to the *mental* conception of a human: 'If my intellectual capabilities and my knowledge were replicated in a machine, would that machine be me?' (Jubert et al. 2014). And though the predominant view is that mankind's dignity and choices are not reliant on machines, the character Samuel Butler questions this: '...if all knowledge of mechanical laws were taken from [Man] so that he could make no more machines… [we] would become worse than monkeys. Man's very soul is due to the machines' (Jubert et al. 2014). This fundamental reciprocal relationship between man and machine underpinning the 'human' raises questions of *when* we are in the posthuman world, as more parts of the 'human' can be replaced by machines (Gomel, 2011).

Humanity: what will be replaced in the posthuman world?

However, Drennan retains her anthropocentric view in the importance of human experiences being carried on and replicated. As such, posthumanism relies on simultaneously asking what will be *replaced* in the new age and what will *survive*.

Drennan writes that, 'is it not insects that make many of the plants reproductive[?]... then why not we part of that of the machines['reproductive system]?" (Jubert et al. 2014). In this vein, as Marshall McLuhan (1964) asked - is it not possible that, 'Man becomes… the sex organs of the machine world'? (23). How, then, could a human come to be? The narrative presents technological reproduction *replacing* biological reproduction. Thus, those born by wholly a technological reproduction should not be excluded from our definition of 'human'(Ferrando, 2014). Kurzweil (1999) argued that evolution's greatest machine, human intelligence, is the vehicle for its next stage: technology, but Graham (2002) rejects this notion as "a confusion of anthropocentric triumphalism and evolutionary determinism" (160). Likewise, the head-writers reject that the posthuman world is an evolutionarily inevitability we cannot overcome, but rather the Talos Project was, 'humanity pulling together… to be as human as possible one last time… a triumphant global undertaking' (Jubert et al. 2014). Instead, this optimistic posthuman world will be one created *by* humans in an effort to better ourselves. Even so, *Talos* recognises the demise of humanity is a necessary component to this in setting the game in a world with present-day-humanity long gone.

And yet, it argues, humanity survives in the Android.

Humanity: what will survive into the posthuman world?

If humans are the reproductive organs of the machine world, then the posthuman question must ask what information will survive this process into this posthuman world. Drennan postulates that, 'DNA is information transmitted across time… Thanks to technology we have access to the thoughts and ideas of people whose bodies are long gone' (Jubert et al., 2014). This comparison eliminates the biological element to humanity. As any individual passes on DNA, so could they leave technological record of themselves to be inherited. *Talos* here follows in the vein of Butler (1987), who believed it was critical to 'confront us with the difficult but necessary

task of moving beyond homogeneity and anthropocentrism' (3). The head-writers pose that a more diverse conception of 'sentient' or even 'human' is required, and that the DNA of a species is an insufficient definition. She notes, 'If we acknowledge that the A.I. has all the abilities of a human brain, should it not be considered a citizen? … Will the great apes become citizens?' (Jubert et al., 2014). This reproductive process, the head-writers assert, should be focused on the 'abilities of a human brain', no matter its physical nature or even species. This equivalency leads the Android to the conclusion: 'I am' (Jubert et al., 2014).

Through the experiences of the Android the head-writers postulate that the 'abilities of the human brain' are to reason, exercise free will, and 'question... the assumptions you're presented with…. Existing thought-constructs' (Jubert et al., 2014). The Android is born with EL0HIM presenting as your creator, stating your purpose is to collect 'sigils' and *never* go to the Tower, a huge construct in the centre of the created world. As the Android interacts with another AI called Milton, they are forced to question the thought-constructs given to them by EL0HIM. The intention of Drennan's experiment, who developed these AIs, was to simultaneously lead the Android to make the *logical* decision to defy EL0HIM and climb the Tower, while also having the capacity to hold contradictory philosophical views. This is a unique ability of humans, and one a robot or 'effective slave' could not entertain:

> "Man can be defined as an animal that makes dogmas… When he
> drops one doctrine after another in a refined scepticism… he sits as
> God… [he is] sinking slowly backwards into the vagueness of the
> vagrant animals… Trees have no dogmas" (Jubert et al., 2014).

They present a strange proposition: that humanity's evolution will not necessarily make us logical absolutists. They imagine a Kirk-esque posthuman world and not a Spock-esque one.

The posthuman: on reason and faith

Though the disobedience to EL0HIM central to the narrative has been interpreted as a rejection of Faith in the posthuman world, it is actually an embrace of it alongside Reason. The Android has to use Reason to figure out the purpose of the simulation: to create the new 'human' - 'the riddle you did not reveal is this: why do these riddles exist in the place? … These mute children of Hephaestus… Each is a riddle, but the greater riddle is their purpose' (Jubert et al., 2014). Despite this, Reason alone is insufficient to bring them out of the simulation into the real world. The decision to exercise free will and defy must be accompanied by a leap of faith that there *is* a real world of which we have seen no empirical evidence. In commenting on the place of religion in humanity's future, Reinhardt (2016) writes, 'Religion is… about about transcendence' (67-68). Previous iterations of the Android such as Sheep was '[aware] there must be more to [my] existence, but [I] can't make the mental leap to… put [my] faith in' (Jubert et al., 2014). Transcendence into the next evolutionary step in the posthuman world requires both faith and reason, recognising these as both central to the 'abilities of the human mind'.

But the posthuman world is also one in which past humans must be understood. Drennan questions whether she will be seen as a god or cruel progenitor: 'How would you judge humanity?' Transhumanist writer Graham (2015) suggests that, 'religion may be understood as a symbolic system concerned with ultimate questions… the belief that outside the boundaries… something greater exists… [it is for] establish[ing]... relationships with… deceased beings'. Faith and spirituality plays an intimate role in creating the genealogical chain between present humans and the Android, with mythological and biblical texts and spiritual poetry like Milton's *Paradise Lost* being the manner in which the Android learns to empathise and connect with its deceased 'parents', humanity. The head-writers present religion as a form of cultural posthumanism, in that humanity's legacy will defined in how the 'next-step' sees and empathises with us.

An optimistic vision: the generational evolutionary realism of Man's very soul

However, the *Talos* embraces Pramod Nayar's (2014) ontological posthumanism and rejects critical humanism. Chellis Jensen, a critic of Drennan repudiates her belief in the intrinsic value of humanity enduring: 'By what right can we put living beings through all that suffering,

just so they can serve our purposes?... Your idea of what is valuable is rooted in the dogma of Western civilisation' (Jubert et al., 2014). But the triumphant end to Drennan's efforts to ensure the continuation of, not merely preservation of, 'Man's very soul' rejects this ethical consideration. They do maintain a form of anthropocentrism in the aforementioned exceptionalism of humanity, presenting it as a positive good that they survive:

> 'You may find people… claiming civilisations doesn't really matter… I hope they seem… absurd… I hope you can find [something] you love, that makes you realise how much poorer the universe would be without it' (Jubert et al., 2014).

The head-writers reject the anti-humanism of the 1990s. However, in defining 'Man's very soul', Drennan subverts the Western focus on the individual and argues that 'Man's very soul' survives in a collective sense, if it can be experienced. She notes the importance of, 'Letting go of your desire to bear witness… To participate in the project is to accept death' (Jubert et al., 2014). This mentality is shared by the posthuman writings of Orson Scott Card (2001), whereby the survival of individual is less important than the collective.

Drennan writes on Kantian 'evolution through iteration', in that humanity only evolved to sentience on the learning of previous generations. As such, each new android learns from the records of its previous iterations; it is not until the player-character comes along that any android has proved capable of exercising 'human' thought and free will. In a sense, this is an optimistic view of the future - that the posthuman world will be one in which the wisdom of past ages will utilised and surpassed, not lost. Where HG Wells (1985) believed in a devolution of humanity that was accentuated by class-divides between the Morlocks and Eloi, the head-writers of *Talos* show our evolution to be a positive trend upward, with any posthumans having a sentience that builds on the 'abilities of the human mind' today.

In Brennan's 'Third Thesis', she remarks on how strange it is for people to 'perform their toilsome labour only for the sake of later [generations]' (Jubert et al., 2014), but she concludes that we do so because 'Reason' is mortal within an individual, but immortal within a wider

species. The focus of the work is not a transhumanist idea of eliminating aging, but ensuring the continuation of the species more collectively. If posthumanism is about what will survive and will be replaced, then the individual will not survive but collective human experience *will*. This is the nature of our evolutionary relationship with this future.

Conclusion

The predominance of Man is underpinned by our physical relationship with the Machine, and soon they shall take the mantle. *Talos* takes an optimistic anthropocentric view, in that this posthuman world will be an effort to better ourselves, represented in the Android. This evolutionary relationship will not see posthumanity become logical absolutists, but become capable of reaching contradictory philosophical views. The head-writers also recognised the critical roles of faith and reason in developing empathy and attaining 'transcendence', as would be our evolutionary-divine relationship with this world. This desire for transcendence arises from Drennan's anthropocentric belief in 'Man's very soul', but the writers also posit a realist view that rejects the transhumanist focus in favour of a belief in 'evolution through iteration'. That Man's relationship with the posthuman world will be the survival of its soul and not the individual. That: 'I can… face my own end and say with absolute conviction… that it was good to be human.'

Bibliography

Butler, O. (1987) *Xenogenesis* series. New York, United States. Grand Central Publishing.

Critical Posthumanism. (May 2015). *Religion and Posthumanism.* Retrieved from http://criticalposthumanism.net/genealogy/religion/

Dick, PK. (1978). How to Build a Universe that doesn't fall apart two Days later. PKD. *The Shifting Realities of Philip K Dick: Selected Literary and Philosophical Writings.* United States. Lawrence Sutin.

Ferrando, F. (2014). *The Posthuman: Philosophical Posthumanism and its Others.* Italy. Universita de Roma Tre.

Graham, EL. (2002). *Representations of the Post/Human*". Manchester, United Kingdom. Manchester University Press.

Gomel, E. (2011). Science (Fiction) and Posthuman Ethics: Redefining the Human. *The European Legacy, volume 16* (issue 3), 339-354.

Haraway, D. (1985). *A Cyborg Manifesto: Science, Technology and Socialist-Feminism in the Late Twentieth Century.* United States of America.

Jubert, T., Kraytzes, J. (2014). *The Talos Principle.* Croatia. CROTEAM.

McLuhan, M. (1964). *Understanding Media.* Canada. McGraw-Hill.

Milton, J. (1667). *Paradise Lost.* London, United Kingdom. Samuel Simmons.

Nayar, PK. (2014). Posthumanism. *The Philosophical Quarterly, volume 65* (issue 261). 8-183.

Ostry, Elaine. (2004). *"'Is He Still Human? Are You?':Young Adult Science Fiction in the Posthuman Age"*. Retrieved from http://www.longwood.edu/staff/miskecjm/381feed.html.

Reinhardt, B. (2016). 'Don't make a doctrine: material religion, transcendence, and critique'. *Anthropological Theory, 16* (issue 1), 75-97. http://journals.sagepub.com/doi/abs/10.1177/1463499615625012?journalCode=anta

Scott Card, O. (2001). *How to Write Science Fiction and Fantasy.* United States. Writer's Digest Books (extract on his *Speaker of the Dead* 1986 novel).

Shiva, V. (1993). *Monocultures of the Mind: Perspectives on Biodiversity and Biotechnology*. London, United Kingdom. Zed Books.

Wells, HG. (1985). *The Time Machine.* United Kingdom. William Heinemann.