

ALAN McGILL

The Robot's Redemption: the Role of Artificial Intelligence in the Salvation of the Cosmos

A view of creation as fallen prior to the fall of humanity implies that the non-human portion of the cosmos can exercise freedom of a kind, having in some sense strayed from its original holiness, while, at the experiential level, the reality of innocent suffering wrought by inanimate forces is difficult to reconcile with the will of a benign deity. Against this backdrop, the present paper proposes that artificial decision-making agents would represent a new phase in non-human creation's response to the divine. While these impersonal moral agents may lack any consciousness of God, their ability to select between options in a calculated manner could implicitly accept or reject grace along the lines envisaged by Karl Rahner, who views the fundamental option for or against God as played out in relation to decisions made in relation to created things and persons, not requiring explicit consciousness of the Creator.

Keywords: Artificial intelligence, freedom, moral choices, creation, salvation, consciousness, the Fall, grace, fundamental option, supernatural existential.

Introduction

This paper raises the question as to whether the moral decisions of autonomous artificial moral agents (AMAs) could represent the emergence of a new phase in the relationship between the inanimate portion of creation and its Creator, whereby components of the inanimate creation respond to God's offer of salvation with a new degree of freedom made evident in their ability to select between options, exercising discretion that is free from direct human oversight.

The paper argues a thesis that the emergence of AMAs reflects the evolving nature of creation's response to God, constituting a confluence of personal, systemic and natural forms of freedom. This combination of the personal, the systemic and the natural includes yet ultimately exceeds the influence not only of individual humans, but even of humanity as a whole. It is a prospect that underscores the importance of thinking of salvation in terms broader than those that pertain only to individual humans or even humanity as a species, challenging our soteriologies to acknowledge the freedom of inanimate forces in accepting or rejecting grace and the interplay of personal, social, and technological factors as it pertains to both sin and salvation.

A universe in need of salvation

St Paul envisages creation as 'groaning in labour pains' (Rom. 8:22) and in need of a restored relationship with its Creator. Gabriel Daly characterises Paul's vision as essentially hopeful, declaring that the "The cosmos itself is taken into the sweep of God's salvific purpose."¹The point is that, in this Pauline view, creation must possess some measure of freedom in order to have departed from the divine plan and in order to crave restoration.

The universe's groaning for salvation is indicative of a problem – or at the very least, of incompleteness. The doctrines of the Fall and of Original Sin signify the reality that the world as we find it falls short of the plan of a loving God – such as we might conceive it – and suggest that even inanimate creation possesses some measure of freedom to deviate from original holiness.

The freedom of the inanimate creation can be conceptualised without resorting to unwarranted anthropomorphism when conceived as an alternative to a blue-print model of the cosmos wherein the forces and processes of nature are micromanaged by the deity. Viruses, rain-showers and tsunamis are not, in this view, manipulated by the Creator but driven by natural forces at a degree of remove from the deity.

Two theological assumptions regarding the God-world relationship

It should be acknowledged that not all Christian thinkers regard the non-human component of creation as either fallen or capable of exercising any degree of agency in its relationship with the Creator. This paper proposes the implications of AMAs for those strands of the Christian tradition that regard the non-human portion of creation as (i) fallen and (ii) endowed with non-human forms of freedom. The strands of Christian thought that embrace, or at the very least engage with, these two theological positions seem significant enough for the present discussion to be worth considering.

(i) non-human creation as fallen

While Jewish thought does not feature a doctrine of the fall or of Original Sin as such, the tradition speaks of a concern for *tikumm olam*, or 'repair of the earth', thus acknowledging that the present condition of the earth falls short of the Creator's ideals.

Within the Christian tradition, there is also a strong precedent, evidenced by thinkers from St Gregory of Nyssa to St Augustine for the position that the

1 Daly, G. *Creation and Redemption*, Dublin: Gill & MacMillan (1988), p. 90.

physical component of the non-human creation is fallen.² The Second Vatican Council reflects such thinking when it teaches 'Christians believe that "the world has been established and kept in being by the Creator's love [but] has fallen into slavery to sin. . ."'³

The position that the non-human component of creation is fallen is reflected in the doctrine, or motif, one might say, of the fall of the angels as inhuman, free-willed creatures. The Patristic thinkers Tatian and Athenagorus classified the fallen angels as demons and numbered Satan among them.⁴ Whether understood as inhuman spirits or as impersonal agents, the doctrine of the fall of the angels suggests one sense in which the non-human creation may be considered fallen.

A related argument for the fallenness of the non-human creation relates to the corruption of the wider non-human creation as finagled by fallen angels. C.S. Lewis proposes that prior to the fall of humanity, the fallen Satan corrupted the material universe.⁵ Gregory Boyd endorses the position that, before the evolution of humanity, evil spirits corrupted the processes of evolution so that the non-human creation can be thought of as fallen prior to the Original Sin of humanity,⁶ and Alvin Platinga, with Augustinian underpinnings, considers the possibility that natural disasters are caused in some part by the agency of demons who continue to manipulate natural processes.⁷

Pope John Paul II in a general audience of 8 October 1986 regards nature beyond human nature as fallen due to the sin of humankind. This is evident in the title of the audience: 'The State of Man in Fallen Nature', suggesting that nature beyond humanity is fallen. The pontiff asserts:

The biblical author does not hesitate to attribute to God as a sentence of condemnation. It implies the 'cursing of the ground' – visible creation has become rebellious and hostile. St Paul says that as a result of man's sin 'creation was subjected to futility', and for this reason also 'the whole creation has been groaning in travail together until now' until it will be 'set free from its bondage to decay' (Rom. 8:19-22). This lack of balance of creation has its

2 Henriksen, J.O. *Imago Dei. Den Teologiske Konstuksjonen av Menneskets Identitet*, Oslo: Gyldendal Akademisk (2003).

3 Vatican Council II, *The Pastoral Constitution on the Church in the Modern World (Gaudium et spes)*, Vatican (1965); available at: http://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat_ii_cons_19651207_gaudium-et-spes_en.html. 2, 2

4 Russell, J.B. *Satan: The Early Christian Tradition*, Ithaca: Cornell University Press (1981), p. 73

5 Lewis, C.S. *The Problem of Pain*, New York: Macmillan (1962), pp. 133-135.

6 Platinga, A. 'The free will defense', *The Analytic Theist*, Grand Rapids: Eerdmans (1998), pp. 16-48.

7 Boyd, G. *Satan and the Problem of Evil: constructing a trinitarian warfare theodicy*, Downer's Grove: Intervarsity (2001), pp. 313-317.

influence on the destiny of man in the visible world.⁸

The pontiff notes that while humanity influenced the fall of the non-human world, the unbalanced, postlapsarian non-human world in turn exerts influence upon humanity and its destiny, implying some degree of agency on the part of the non-human creation. Granted, it should be added, Catholic theology regards the postlapsarian nature as not only fallen but also infused with grace, its sacramental perspective viewing creation as simultaneously corrupted and sacred.⁹

From a reformed perspective, Calvin asserts 'the whole order of nature was subverted by the sin of man',¹⁰ a position shared by Reinhold Niebuhr, Bonhoeffer and Moltmann.¹¹ Reformed thought does not, for the most part, share the sacramental, panentheistic perspective of the Catholic Tradition, which is not to say that it does not recognise the efficacy of grace. Nonetheless, it may be fair to argue that, from a reformed perspective, fallen non-human nature is in a bleak situation.

Having noted precedence in Christian thought for the position that the non-human creation may be regarded as fallen, it might still be argued that its fall is due to corruption by fallen humans or fallen angels and that non-human creation itself does not exercise freedom in its relationship with its Creator. We turn therefore to an argument that all of creation is endowed with some measure of freedom in relation to its Maker.

(ii) non-human creation as free

The position that the non-human component of creation possesses a degree of freedom to cooperate with God's plan is explicitly evident in Process Theology. Alfred North Whitehead regards all of creation as endowed with creative freedom. Whitehead posits that 'The world is self-creative; and the actual entity as self-creating creature passes into its immortal function of part-creator of the

8 Pope John Paul II, 'The State of Man in Fallen Nature', general audience of October 8, 1986; https://w2.vatican.va/content/johnpaulii/it/audiences/1986/documents/hf_jp-ii_aud_19861008.htm

9 Himes, M. "Finding God in All Things": a sacramental worldview and its effects', in Landy, T. (ed.) *As Leaven in the World: Catholic perspectives on faith, vocation, and the intellectual life*, Franklin, WI: Sheed & Ward (2001) available online at <https://litpress.org/Products/GetSample/8271/9780814682715>

10 Calvin, J. *Commentary on Genesis*, 2 vols., King, J. (trans.) 1554, Grand Rapids, MI: Baker, (1996), 1:102; online at <http://www.ccel.org/ccel/calvin/calcom01.pdf>.

11 Bonhoeffer, D. 'Creation and Fall: a theological exposition of Genesis 1-3', *Dietrich Bonhoeffer Works, Volume Three*, DeGruchy, J.W. (ed.), Bax, D.S. (trans.), Minneapolis: Fortress (1997); Moltmann, J. *God in Creation: An Ecological Doctrine of Creation*, London: SCM (1985), p. 68; Lamoureux, D.O. 'Beyond the cosmic fall and natural evil', *God and Nature Magazine* 68 (1), 57; Roskos, N. 'Christian Theology and Fall' <http://users.clas.ufl.edu/bron/pdf--christianity/Roskos--The%20Fall%20in%20Christianity.pdf>.

transcendent world; the freedom inherent in the universe is constituted by this element of self-causation.¹² For Whitehead, the world, and each constituent part of it, is endowed with freedom to bring about novelty beyond a perpetuation of the status quo.

Process thinkers view creation as continuing to unfold through a process of concrescence wherein novelty emerges, becoming 'concrete' through an interplay of the subjective aims of entities, divine aims and the impact of historical context. In speaking of subjective aims, Process thinkers are not anthropomorphising non-sentient entities. Whitehead regarded consciousness as the 'tip of the iceberg' in creation, convinced that most of the ongoing process of creation transpires at an unconscious level, remarking 'consciousness presupposes experience, and not experience consciousness.'¹³ Whitehead and Process thinkers of his ilk suggest that there is an elasticity built into natural processes and natural forces so that they veer towards one possibility or another, albeit without cognitive reflection. They do, nonetheless, in their inanimate way, select between options.

A recognition of the impersonal model of freedom possessed by inanimate forces is not unique to Process thought. As Gabriel Daly notes 'One does not have to be a card-carrying Process theologian in order to recognize that God has freely accepted limitations on his divine power when he creates our universe.'¹⁴ The freedom of the entire creation, it might be said, results from God's voluntary self-limitation as noted by John Macquarrie.¹⁵ If God voluntarily limits divine control of the non-human component of creation, then that component of creation is, in effect, granted a degree of freedom.

In light of these ruminations, characterising the inanimate component of the cosmos as fallen, and endowed with a form of freedom with which to implicitly embrace God's gracious offer of salvation or to reject it, we now address the significance of technology in the unfolding relationship between God and creation.

The fallen nature of technology

Pope Francis regards technological developments as holding moral significance.

12 Whitehead, A.N. *Process and Reality: an essay in cosmology*, New York: Macmillan (1929), pp. 84-88.

13 *ibid.*, p.83; Cobb, J.B. & Griffin, D.R. *Process Theology: An Introductory Exposition*, Louisville: Westminster John Knox Press (1976), p. 17.

14 Daly, G. *op. cit.*, (1), p. 33.

15 Macquarrie, J. *In Search of a Deity: An Essay in Dialectical Theism*, London: SCM Press (1984), p. 180.

We have to accept that technological products are not neutral, for they create a framework which ends up conditioning lifestyles and shaping social possibilities along the lines dictated by the interests of certain powerful groups. Decisions which may seem purely instrumental are in reality decisions about the kind of society we want to build.¹⁶

In his TED talk of 25 April 2017, the Holy Father elaborates upon the morally-charged role that he attributes technology, lamenting that ‘our techno-economic systems which, without even realizing it, are now putting products at their core, instead of people.’¹⁷

On a similar note, Thomas Merton implies that technology seems to veer from its role in the right ordering of creation, decrying that ‘If technology remained at the service of what is higher than itself – reason, man, God – it might fulfill some of the functions now mythically attributed to it.’¹⁸ The implication is that technology has become self-serving.

Merton’s subsequent remark seems particularly applicable to the prospect of AMAs: ‘But becoming autonomous, existing only for itself, it imposes upon man its own irrational demands and threatens to destroy him.’¹⁹ Hence, albeit in a rather sweeping condemnation, Merton regards technology as becoming increasingly destructive of the common good as it becomes more autonomous of direct human control. Merton objects that technology tends to prioritise efficiency over humanitarian concerns. ‘Technology has its own ethic of expediency and efficiency. What *can* be done efficiently *must* be done in the most efficient way – even if what is done happens, for example, to be genocide or the devastation of a country by total war.’²⁰ In noting technology’s wayward turn from its rightful purpose, Merton attributes it its ‘own ethic’, as distinct from the ethic of its human designers and the social systems that condition it. Merton hence implies that technology can possess some degree of moral responsibility and autonomy.

The prospect of autonomous Artificial Moral Agents

The capacity of computers to leverage experience so as to make decisions

16 Pope Francis, *Laudato Si* (On Care for Our Common Home) (2015), available at http://w2.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html.

17 Pope Francis, ‘Why the only future worth building includes everyone’, Technology, Entertainment, Design (TED) conference, April 2017, Montrasio, H. (trans.), available at https://www.ted.com/talks/pope_francis_why_the_only_future_worth_building_includes_everyone/transcript?language=en#t-272053.

18 Merton, T. *Conjectures of a Guilty Bystander*, New York: Random House (1966), p. 71.

19 *ibid.*

20 *ibid.*

and improve performance is often referred to as artificial intelligence, a term coined by John McCarthy in the context of a 1956 Dartmouth conference for thought-leaders in this emerging field of inquiry.²¹

Recognising the philosophical difficulties entailed by the question as to whether a machine can think, as opposed to replicating human input, Alan Turing in 1950 devised the 'imitation game', showing that a machine could convincingly imitate human conversation.²² Turing's test was not directly concerned with whether the machine was right but rather whether it could imitate a thinking human.

The distinction between artificial intelligence and intelligence augmentation may be significant at this point in so far as the latter suggests an aid to human intelligence rather than a force that could function independently of it.²³ Still, a slippery slope may exist in this regard if automated systems are empowered to learn from their experience and to derive more effective work habits. John Markoff notes the permeable line between intelligence augmentation and artificial intelligence, remarking that machine may 'simultaneously augment and displace humans'.²⁴

The freedom of AMAs

Wallach and Allen map degrees of autonomy for AMAs, ranging from the operational morality inherent in the autopilot system on an aircraft to full moral autonomy. In order for an AMA to undertake and implement autonomous moral decisions, it would need both moral sensitivity and autonomy to make decisions. An autopilot system, for example, has significant autonomy in decision-making but no moral sensitivity. By contrast, a decision-support system in healthcare (MedEthEx) may be programmed to reflect a high level of moral sensitivity yet have no autonomy to act.²⁵

James Moor outlines a hierarchical schema of AMA's ranging including 'implicit' ethical agents such as autopilot systems that are simply programmed to adhere to a predetermined program. In contrast to these implicit moral agents, Moor classifies 'explicit' ethical agents as those programmed to reason accord-

21 Kaplan, J. *Artificial Intelligence: What Everyone Needs to Know*, New York: Oxford University Press (2016), p. 2.

22 Warwick, K. & Shah, H. *Turing's Imitation Game: Conversations with the Unknown*, London: Cambridge University Press (2016), p. 2.

23 Markoff, J. *Machines of Loving Grace: the quest for common ground between humans and robots*, New York: Ecco-HarperCollins (2015), p. 24.

24 *ibid.*, pp., 12, 16, 24, 31.

25 Wallach, W. & Allen, C. *Moral Machines: Teaching Robots Right from Wrong*, New York: Oxford University Press (2008), pp. 25, 26, 27.

ing to ethical systems such as deontic ethics.²⁶ Such systems are programmed by their human creators to adopt a particular approach to ethics and then possess sufficient freedom to apply that prescribed approach to a given situation. These scenarios suggest pre-programming according to an ethical code, whether an established school of ethics such as deontologist or utilitarian ethics, or a code customised by the system's designers.

An explicit ethical agent might be programmed according to Isaac Asimov's three ethical rules for robotics.²⁷ Asimov's first law dictates that a robot may not injure a human being or, through inaction, allow a human being to come to harm. The second law prescribes that a robot must obey orders given to it by human beings except where such orders would conflict with the first law and the third law requires that a robot must protect its own existence, as long as such protection does not conflict with the first or second law. Asimov later added the stipulation that 'a robot may not harm humanity, or, by inaction, allow humanity to come to harm'. These rules cannot by themselves, however, account for dilemmas where, for example, a robot must decide among the conflicting needs of several humans.

An explicit ethical agent could be programmed with an elaborate code far exceeding the basics outlined in Asimov's rules. Kris Hammond suggests that an artificial moral agent programmed with a utilitarian approach to ethics can make a more objective and speedy decision than a human being since the artificial agent is not hampered by emotional bias. Hammond argues that artificial intelligence will 'run the numbers and do the right thing'.²⁸ This however suggests a form of quantitative rather than qualitative utilitarianism, and arguably dismisses the role of emotional intelligence, including empathy.

While an explicit ethical agent as categorised by Moor is programmed to implement a given ethical code or approach, albeit with some discretion in the matter of application, an AMA deemed a 'full' ethical agent could go beyond the implementation of a code to learn from experience and develop moral stances of its own.²⁹ The prospect of an AMA functioning as a full ethical agent may be purely hypothetical. Wallach and Allen acknowledge that it remains contentious among scientists and among philosophers whether it is possible for a machine to be a full ethical agent.

Full ethical agency, some might argue, would require conscious freedom. Wallach and Allen, however, contend that freedom does not necessarily com-

26 Wallach & Allen *op. cit.*, (25), p. 33.

27 Singer, P. 'Isaac Asimov's laws of robotics are wrong', available at <https://www.brookings.edu/opinions/isaac-asimovs-laws-of-robotics-are-wrong/>

28 Hammond, K. 'Ethics and Artificial Intelligence: The Moral Compass of a Machine - the question of robotic ethics is making everyone tense', <https://www.recode.net/2016/4/13/11644890/ethics-and-artificial-intelligence-the-moral-compass-of-a-machine>.

29 *ibid.*, pp. 33-34.

prise a 'mystical' dimension such as consciousness and could simply consist in an ability to select from a range of options.³⁰ These authors frame in pragmatic terms the driving rationale behind moral deliberation, asking 'Why does the internal dialogue begin? We've seen how it begins with a proposer-attention codelet popping a proposal into mind.'³¹ Along these lines, Daniel Dennett posits that free will merely consists in an ability to make a selection from a range of options.³²

Also instantiating what might be regarded as a pragmatic alternative to a mystical view of freedom, Stan Franklin explores the possibility of representing moral deliberation in a mechanical fashion. Franklin has devised a conceptual framework known as a learning intelligence distribution agent (LIDA), suggesting how the various influences on ethical decisions, from feelings to rules, might be represented.³³ The LIDA framework proposes that the moral agent experiences five to ten inputs every second and five to ten behaviours are selected. In each cognitive cycle, there is a 'winner' that coordinates a behaviour response. Franklin's approach views moral decision-making as a deliberation between conflicting stimuli.³⁴ For Franklin, 'it is the perceived novelty of a given situation that leads to it being the subject of deliberation rather than simply being selected'.³⁵ The author notes that 'Past moral experiences may become nodes invoked by new but similar experiences.'³⁶

Notably, Franklin locates much moral deliberation at the subconscious level.³⁷ While modern psychology locates much of humanity's motivations at the subconscious level, the question arises as to what kinds of deliberation might take place in 'subconscious' creation, or put differently, at a level of consciousness distinct from human or even animal sentience.

Moral choices and the implicit acceptance or rejection of salvation

Karl Rahner proposes that a person's ultimate decision for or against God, that is, their acceptance or rejection of salvation, does not depend upon a conscious awareness or conceptualisation of God.³⁸ Rather, it is played out

30 Wallach & Allen *op. cit.*, (25), p. 60.

31 *ibid.*, p. 182.

32 Dennett, D. *Freedom Evolves*, New York: Viking Press (2003).

33 *ibid.*, pp. 176 -179; Baars, B. & Franklin, S. 'Consciousness is computational: the LIDA model of global workspace theory', *International Journal of Machine Consciousness* 1(1), 23-32.

34 Wallach & Allen *op. cit.*, (25), p.181.

35 *ibid.*, p. 183.

36 *ibid.*, p. 182.

37 *ibid.*, p. 174.

38 Rahner, K. 'The commandment of love in relation to the other commandments', Kruger, K.

in relation to a lifetime's decisions about created things and persons. Ron Highfield characterises Rahner's position arguing that 'The decision about God takes place in decisions about finite things, since God is unthematically present in every act of choice as its ground and goal.'³⁹ Hence, for Rahner, there are soteriological implications, for better or for worse, in the decisions of a primitive human who lacks any concept of God or gods, or a modern person completely insulated from religious concepts.⁴⁰

In Rahner's view, the fundamental option is not a one-off, discrete choice in the same category as other choices. For Rahner, it is an implicit choice that cannot be directly equated with any explicit decision.⁴¹ Joseph Fuchs regards the fundamental option as a form of freedom that is not consciously exercised in the same manner as psychological freedom, that is, the freedom to make an explicit choice.⁴² In a similar vein, John Glaser refers to this as 'core freedom'.⁴³

If the choice for or against God can be an implicit one played out through choices that are made ostensibly in relation to created things and persons rather than in relation to an explicit recognition of the Creator, this is of great significance for AMAs. Rahner's focus is highly anthropological and the present work moves beyond his scope by exploring whether an acceptance or rejection of grace, albeit 'unexpressed, unthematised, unobjectified, and unspoken' may transpire in impersonal freedom whereby a non-human agent selects between options in a calculated manner.⁴⁴

None of this is to argue that the inanimate component of the cosmos, including artificial intelligence, is to be saved in the same manner as are human persons. If the salvation of human persons takes the form of theosis, this is not necessarily to claim that all of creation is to be saved in such a manner.⁴⁵ It is however to argue that the non-human component of creation possesses a form of free will.

(trans.) *Theological Investigations* 5, Baltimore: Helicon Press (1966), pp. 439-460; Rahner, K. 'Theology of freedom' & 'Reflections on the unity of the love of neighbor and love of God', Kruger, K. (trans.) *Theological Investigations* 6, Baltimore: Helicon Press (1969), pp. 178-196, 231-252.

39 Highfield, R. 'The freedom to say "No"? Karl Rahner's doctrine of sin', *Theological Studies* 56 (3): 485-505, 487.

40 Rahner, K. *Foundations of Christian Faith: An Introduction to the Idea of Christianity*, Dych, W. (trans.), New York: Seabury Press (1978), p. 98.

41 *ibid.*, pp. 93-106.

42 Fuchs, J. *Human Values and Christian Morality*, Dublin: Gill and Macmillan (1970), p. 93.

43 Glaser, J. 'Transition between grace and sin: fresh perspectives', *Theological Studies* (1968) 29, 263-265.

44 Rahner, K. *op. cit.*, (40), p. 98.

45 McFarland, I. *In Adam's Fall: A Meditation on the Christian Doctrine of Original Sin*, Chichester: Wiley (2010), X.

Freedom without personal conscience

From a Catholic perspective, personal conscience is an internal forum wherein the voice of God resounds, guiding a person towards salvation.

Deep within his conscience man discovers a law which he has not laid upon himself but which he must obey. Its voice, ever calling him to love and to do what is good and to avoid evil, sounds in his heart at the right moment... . For man has in his heart a law inscribed by God... . His conscience is man's most secret core and his sanctuary. There he is alone with God whose voice echoes in his depths.⁴⁶

Social and organisational structures have since the dawn of tribal structures to some degree distanced decision-making from the direct influence of personal conscience. The literature of organisational theory suggests that bureaucracies tend to deliberately create a 'bounded rationality', that is, their structures constrict the stimuli that confront a human agent and hence constrict the factors that the human agent takes into account in decision-making.⁴⁷ In this manner, it might be argued, a bureaucracy distances decision-making from the discretion exercised in personal conscience. Policies, role-descriptions and protocols, in the interest of efficiency, impartiality and standardisation constrict human agents from adopting a holistic assessment of each decision with which they are confronted. Granted, bounded rationality may be morally defensible to the extent that it can serve to limit personal bias and foment impartiality. At worst, however, it can manipulate human agents to suppress their conscientious objections and follow orders.

The prospect of fully autonomous artificial intelligence would distance decision-making even further from the influence of personal conscience. Indeed what distinguishes the quest for artificial intelligence (as opposed to intelligence augmentation) from other developments is that it is, by definition, a deliberate attempt to endow automated systems with capacities associated with human intelligence so that they can make decisions independent of direct human intervention – and, in effect, at a remove from the influence of human conscience.

It might on the other hand be argued that humanity has for millennia used prosthetic devices, extending human agency through tools, systems, gadgets and structures. Use of these prosthetic devices, it might be argued, has not led to an abdication of human responsibility so much as an extension of it, effectively extending the perceived boundaries of the self. As Steven Kraftchick

46 The Congregation for the Doctrine of the Faith, *Catechism of the Catholic Church: Revised in Accordance with the Official Latin Text Promulgated by Pope John Paul II*, (Vatican City: Libreria Editrice Vaticana, 1997), § 1776, Available at http://www.vatican.va/archive/ENG0015/_INDEX.HTM.

47 Simon, H.A. *Models of Bounded Rationality*, Boston: MIT Press (1982).

notes, "The questions are actually age old, but a new wave of technology and institutional advancement has refashioned our senses of them, with our dawning recognition that technology is now part of our human make-up, and not simply, something we use."⁴⁸

Kraftchick evokes the prospective situation of the cyborg, a human-robot hybrid, proposing that a Pauline theology might allow for a very inclusive definition of personhood, not necessarily excluding cyborgs from God's salvific plan. Kraftchick suggests, with reference to St Paul

... because of his insistence that 'in Christ' distinctions based on physical features or cultural constructions of gender, race, and station are rendered inconsequential (Rom. 12-14; Gal. 3:28), if and when the boundary between the organic human and the hybrid human/cyborg is sufficiently blurred so that one cannot be recognized from the other, Paul may very well argue that these transhumans be included in the redemptive community along with the rest of God's creation.⁴⁹

Notably, Kraftchick is not simply arguing that the cyborg might, because it is partly human, be included in the community of the redeemed, but that it could be saved 'along with the rest of God's creation'.⁵⁰ The author hence regards the rest of creation, by implication including its inanimate components, as in need of redemption and potentially within its sweep.

Inanimate creation's capacity to respond to God

Christian belief in angels and in demons as fallen angels has long affirmed that free will exists in the inhuman realm. Walter Wink appropriates the motif of the angel and spirit to express his conviction that systems and structures can possess a level of awareness and agency. Wink argues that organisational and technological systems possess a 'spirit', 'power' or ethos that, although not a personal being, has the capacity to respond to God, either positively or negatively. Having defined spirit as the capacity to be aware of and responsive to God, Wink notes the contribution of process thinkers who recognise the freedom exercised by creation.⁵¹

Alfred North Whitehead and John Cobb envisage God persuasively luring creation to align its subjective aims with the divine aims.⁵² Cobb speaks of a

48 Kraftchick, S. 'Bodies, selves and human identity: a conversation between transhumanism and the Apostle Paul', *Theology Today* 72 (1), 47-69, 48.

49 *ibid.*, 69.

50 *ibid.*

51 Wink, W. *The Powers That Be: Theology for a New Millennium*, New York: Galilee (1993), p. 19.

52 Whitehead, A.N. *An Essay in Cosmology, Corrected Edition*, New York: The Free Press (1978),

'doctrine of self-determination' whereby every entity in creation interacts with divine aims and with the conditions posed by its history so as to exercise real freedom.⁵³ Robert Kark Gnuse argues that God 'provides direction for the universe to take, even though individual entities are free to deviate from that divine subjective aim'.⁵⁴ Ian Barbour posits that God acts as a 'structuring cause', providing the range of possibilities within which creatures act, and Catherine Keller speaks of a 'cosmic collaboration' within which creation is endowed with freedom to co-create with God – or not.⁵⁵

Wink never claims that the powers are conscious in the same manner in which human persons can be. Rather, he argues that they are 'aware' of God and capable of being 'responsive' to God. Wink proposes, 'We are right on the brink of discovering Soul at the core of every created thing. There is nothing from DNA to the United Nations that does not have God at its core. Everything has a spiritual aspect. Everything is answerable to God.'⁵⁶

When Wink asserts that institutional spirits are in some sense 'aware' of God and can respond to God, the question arises as to whether the consciousness, intentionality and free will in question must be of the anthropomorphic kind. Could there exist in creation, other forms of consciousness, intentionality and free will? Wallach and Martin ask whether consciousness may be 'in *some* measure a universal property of matter'.⁵⁷ If this is the case, human consciousness and machine consciousness would be subsets within this broader property of all matter.⁵⁸

Wink suggests that it is with the emergence of human consciousness that the universe becomes conscious of God as its principal of cohesion. 'The universe is late in arriving at an awareness of itself as a unity, and this awareness has come into the world for the first time with humanity'.⁵⁹ While Wink might be interpreted as saying that it is only the human component of the universe that is self-aware, we might recall his position that spirit, the capacity to be aware of God and to respond to God, lies at the heart of all things.⁶⁰ With this in mind, Wink's remarks might be construed as saying that the evolution of humanity has, in turn, facilitated the emergence of a broader self-awareness

p.187; Wiles, M. 'God the Creator' in *God's Action in the World: The Bampton Lecture for 1986*, Eugene, OR: Wipf & Stock (1986), p. 19.

53 Cobb & Griffin *op. cit.*, (13), p. 25.

54 Gnuse, R.K. *The Old Testament and Process Theology*, Eugene, OR: Wipf & Stock (2016), p. 36.

55 Keller, C. *On the Mystery: Discerning Divinity in Process*, Minneapolis: Fortress (2008), p. 62.

56 Wink *op. cit.*, (51), p. 5.

57 Wallach & Allen *op. cit.*, (25), p. 66.

58 *ibid.*

59 *ibid.*, p. 114.

60 Wink *op. cit.*, (51), p. 5.

in the universe. 'The Powers did not know, but they know now.'⁶¹ Gabriel Daly strikes a similar note, contending that humanity's conscious craving for salvation gives voice to a wider, cosmic yearning.⁶²

Teilhard de Chardin argues that the cosmos has always possessed a measure of consciousness, even before the emergence of humanity.⁶³ 'Whatever instance we may think of, we may be sure that every time a richer and better organized structure will correspond to the more developed consciousness.' The author adopts a form of panpsychism, detecting degrees of consciousness in the inanimate component of creation. Consciousness, for Teilhard, has evolved in the cosmos, reaching the point of self-consciousness in human beings.

While both Wink and Teilhard ascribe some degree of consciousness to impersonal forces, the thesis of the present paper does not ultimately depend upon any form of panpsychism. Rather, it argues that it is the AMA's ability to select between options in a calculated manner that represents a response to God who is unthematically present in and through created things. In keeping with Rahner's theology of fundamental option, a selection between options needs not reflect conscious awareness of God to order to implicitly accept or resist God.

The interplay of the personal, social and technological in the acceptance of salvation

Christian theology has not restricted its understanding of sin to the personal realm, having long recognised the reality of original sin and, in modernity, broadly acknowledged the scandal of social, systemic or institutional sin. To accept the existence of systemic sin is to recognise that impersonal forces can gather momentum so as to select options at odds with the aims of a loving deity such as that preached by Jesus of Nazareth.

Creation's resistance to cosmic salvation, it might be argued, can be understood as collusion between the personal, the institutional and the natural. In a case in point, the scandal of human starvation may reflect the most colossal suffering and evil on the face of the planet. The horror of famine, it might be argued, reflects the tragic interplay of personal, systemic and natural forces resulting, for example, from the whims of dictators, the consequence of economic systems and changes in weather patterns.

If soteriology can learn from hamartiology, then it might also be argued that what is true of sin, that is, resistance to the divine plan, may also be true of

61 *ibid.*, p. 115.

62 Daly *op. cit.*, (1), p. 129.

63 Teilhard de Chardin, P. *The Phenomenon of Man*, New York: Harper & Row (1955), pp. 30, 60-61.

creation's cooperation with the divine plan. Creation's cooperation with grace as well as its resistance to it could comprise a collusion between personnel, systemic and physical factors.

To argue that the emergence of AMAs would, in effect, represent the evolution of a new phase in the response of the cosmos to its Creator is not necessarily to claim that computers will ever be conscious in the manner in which human persons are. Rather, it could be to say that AMAs would represent a new confluence between the responses of animate and inanimate creation to God.

The moral freedom of AMAs would seem to include elements of the personal vis-à-vis the decisions of its human originators. It would include the systemic vis-a-vis the ethical and technological systems with which it is programmed, and the physical or natural in the sense that the decisions undertaken by automated systems and the freedom of matter itself all contribute towards outcomes with moral significance. This confluence of the personal, the systemic and the physical would, by virtue of its multi-faceted nature, constitute a complex response to grace.

Ray Kurzweil, influenced by Teilhard, divides the history of the cosmos into epochs of increasing complexity: (i) the epoch of physics and chemistry; (ii) the epoch of biology; (iii) the epoch of brains; (iv) the epoch of technology; (v) the epoch of the merger of biology and technology; (vi) the epoch in which the universe wakes up.⁶⁴ The emergence of AMAs could be indicative of an epoch of the merger of biology and technology, not necessarily in the form of cyborgs, but in a confluence of natural and artificial intelligences.

As such, the goodness and the harm resulting from the decision made by AMAs may reflect social or physical evil as well as the personal morality of its human designers. Somewhat paradoxically, the more operational autonomy an artificial moral agent commands, it might be argued, the more responsibility for its misdeeds is incurred by its human originators who knowingly incurred risks by unleashing the artificial agent upon the world.

Conclusion

The present paper anticipates the phenomenon of the AMA, arguing that its selection between options would either cooperate with divine aims or else resist them, and hence play a role in the cosmos's acceptance of or resistance to salvation.

The paper avoids ascribing an anthropomorphic form of freedom to inanimate creation in general, or specifically to AMAs. Rather, it characterises the freedom of inanimate creation as a pragmatic model of freedom made evident in the

64 Kurzweil, R. *The Singularity is Near*, London: Duckworth Overlook (2005), p. 15.

ability to select between possibilities in a calculated manner. Arguing that such selections hold soteriological as well as moral significance, the paper notes Rahner's position that the ultimate choice for or against salvation is played out in and through a plethora of decisions in relation to created things, not requiring an explicit awareness of God.

While personal, systemic and natural manifestations of good or evil may be evident in the consequences wrought by AMAs as they make and implement morally significant decisions, the advent of such moral agency may signal a new phase in creation's acceptance or rejection of grace. Certainly, the response of AMAs to grace can be described in terms of the extant categories of personal, systemic and natural factors, but it would represent a new interplay of these factors.

Just as human freedom evolved from the forms of freedom possessed by other species, so too the freedom of AMAs would evolve in continuity from the freedom of humanity and of the wider creation. We need not assert a radical, categorical discontinuity between the freedom of AMAs and preceding forms of freedom in order to detect the emergence of a new phase in the cosmic response to God.

Alan McGill PhD is Chair of Religious Studies at Cristo Rey Atlanta Jesuit High School, Part-time faculty in Religion at Georgia Gwinnett College, and Adjunct Instructor in Theology at Spring Hill College.
