

Scripting for Agency

Introduction

Scripting for Agency addresses questions concerning personhood by examining how disciplines such as art, literature, complexity & life sciences and AI & machine learning all imply assumptions about the prerequisites that make a person, or an agent. A key cross-disciplinary feature of these points of view is that a person may be thought of as 'written' or inscribed, albeit in a manner that results in unpredictable, autonomous behaviour. But how are we to intuit this? To what extent is something as close to home as the human mind, written – inscribed by ancient grooves of genetic memory, articulated by chains of DNA text and scripted by a culture of stories?

As a storyteller, working variously in writing, theatrically performing, or making short films, the aim of my thesis is to focus on this 'scripted' or algorithmic aspect of personhood, and in particular on that peculiar duality of a code and its 'expression' - for example, the existence of a code and 'running' it, a genotype and its phenotype, 'codes of conduct' and behaviour. Each of these examples manifest as a dual system of writing/reading, and they also characterise those parameters typically believed to form a person. In this view, we can imagine that an individual is a script that is 'playing out' its cultural and genetic designations, along with any creative deviation implied by iteration and mutation. But - within this algorithmic mechanism already appears a ghostly element - what does it mean to 'run' or 'play out' a script, anyway? To actualise, perform or invoke a code?

The textuality of persons is a familiar notion in philosophical enquiry, but one that can be practically examined in the laboratory of art practice. I propose to undertake a comparative study of characters born under vastly different circumstances, be they genetic, literary or computational - addressing how art practice can offer a unique and significant contribution to contemporary studies of agency and personhood. This PhD by practice and written dissertation will examine the virtual properties of text, performance and narrativity and how these may be manipulated to probe and exacerbate the trembling boundaries between 'authentic' and 'fictional' beings in light of contemporary AI research.

Research Background and Question

In the budding era of artificial intelligence, storytelling may seem a surprisingly ancient technology to consult on the matter of what exactly we mean by such a thing as a 'person'. After all, it is precisely the advent of machines that think, see

and create to increasingly uncanny degrees which reveals so unavoidably, perhaps painfully, the extent to which personhood is so little understood amongst persons themselves. Tech company spokesmen tend to soothe us from the impending identity crisis with voices of reason: machines may impress us with their prowess for calculation, pattern-spotting and inferences from statistical data, but they will never truly have what humans have; that special-something, that ingredient-X, that emotional investment in the world, a consciousness, a self-awareness. This argument is proliferated widely in the media and all too often prematurely satisfies the debate on the potential personhood of machines, especially given that it is based on an unexplained confidence and belief in 'the ghost in the shell', revealing a pervading predilection for dualistic thinking despite millennia of philosophical controversy surrounding that very assumption. Thinkers such as AI ethics scholar Joanna Bryson are quick to spot the elephant in the room: There are as yet no clear diagnostics as to what constitutes a person and the 'special ingredient' required for subjectivity, or indeed whether the search for such an ingredient is itself misleading as to the mechanics of personhood (Hofstadter, Dennett). While AI dramatically emphasises the gaps in our understanding of the parameters that constitute personhood, storytelling presents itself as an age-old yet ongoing precursor to modern endeavours to produce artificial agents, whereby I propose to study fictional characters as prototypes of 'general' or subjective AI, as well as gain insight from them as to what is meant by 'person'. As such, my project takes departure from the observations laid out in narratologist John Frow's *Character and Person* (2015), who, noting that 'social personhood functions as a kind of fiction' (legally, medically, politically, culturally), questions exactly how fictional characters are ontologically different from real persons. Not only do they move us and become familiar to us like persons in their own right, and not only do we engage with them in cognitively similar ways to real persons (Zunshine); fictional characters teach us how to be persons. They are social constructs that enter our social purview and construct us in turn, and already at this primitive turn in the history of human technology, the fictional characters that populate our stories begin to live among us and raise us in their image. Humans live with the ghosts of fiction and become ghosts of fiction themselves.

Whether the textual phenomenon underlying autonomous systems as diverse as oral narrative tradition, persons, social norms, organisms and prospective general AI is merely an amusing metaphor, a coincidence; or whether it carries with it a fundamental, mechanical performativity worth implementing into a definition of personhood, selfhood or agency - as well as a reassessment of the distinctions of such terms and their relevance - remains to be seen. But if the role of text proves to be significant, then it would deliver a mechanical account of agency that preserves its ghost, and writing could become a potent experimental tool.

We know that prescription and autonomy are not as contradictory as they seem. Complexity sciences have proven that a system as complex as our own

universe is technically programmable (Wolfram), while in the life sciences the notion of scripted organisms has long been accepted. Meanwhile, 20th century philosophical revelations (in quantum mechanics, social critical theory, New Materialisms) contest the assumption that our world is distinct from its conceptualisations, arguing that epistemological positions co-create ontological ones and that the one cannot be apprehended independently of the other (Barad).

We learn that the artwork's intentionality extends far beyond the artist-agent that summons it, and elopes with its complexity to act rather as an agent in its own right (Gell). Artworks, authored by a fractal trajectory of layer upon layer of aesthetic reference and semiological ingenuity, far exceed the intentions of the singular individual, the 'artist', that happens to be there first at the site of a new artwork coming to be. Like any agent then, the artwork has magical potential, some reckless agenda inscribed beyond the stuff and history of its making. Art is not merely 'about' something, but constructively meddles with the stuff of the world and newfangles it - newfangles the world.

It is these theories of authenticity and agency themselves that compel me as an artist to question whether that which feels most real and substantial to us can be written and rewritten by aesthetic and textual tampering. By playfully preying on the growing confusions in the space between persons and characters, autonomy and automation, authenticity and reproduction, artwork and world, I find that I am able to address these theoretical distinctions experimentally.

Research Methods and Outcomes

The project will result in (1) a scholarly book on the performativity of text in relation to scripting for agency in narrative (2) *Anomaline*, a novel about an antiperson and (3) a series of one-woman theatrical seances. The thesis (1) will attempt to expand the 'general' AI discourse to include not only computers, but also the agency of fictional characters and the programmability of organisms; cataloguing and cross-examining a history of characters that transgress their virtuality and the artistic innovations that enable this. In writing the novel (2) I am exploiting the comparable agential force my protagonist and I exert in order to coax somebody strangely alive from the "text that becomes a machine" (Derrida). The novel is already partially developed (70,000 words) and acts as a case study for the paradoxical role of authorship in the case of writing a text with its own independent agenda, questioning whether this is in fact paradoxical at all. By dramatising the relationship between author and protagonist I am using this research to heighten the reader's awareness of their own performative role in realising the agency of the character. Meanwhile, a video storytelling practice (3) provides a running commentary of the research, airing unruly tangents, abusing theories in ways often inaccessible to

stricter disciplines and stretching them malleably in the maverick language of improvised personas.

An integral outcome of this interdisciplinary project will also be collaboratively convening these ideas with AI, Robotics & Cognition researchers from the Computing department at Goldsmiths in the form of joint publications and seminars (4). What would it be like to invite an AI engineer, a geneticist, a philosopher, a novelist and actor to the same séance? In this respect my primary supervisor, Michael Archer will be able to offer guidance on the philosophical, critical and literary aspects of my research and Grace Schwindt, my co-supervisor, will have the expertise to advise me on matters relating to developing practice-based works for screen and stage. The training and networking resources provided by Goldsmiths and CHASE are therefore vital to the rigour and innovation of my project: through exchanges with researchers, published papers, conferences and links to the research networks throughout the consortium, I hope to collaboratively develop my storytelling practice into a site of AI experimentation and agent coaxing.

Schedule of Work

Year 1 *Defining the Field & Reviewing Literature*

- Term 1
 - Developing basic evolutionary algorithms with collaborator (programmer), observing autonomous processes in computational scripting
 - Flashpoint seminar: present working paper and performance
 - Develop full paper from seminar: 'Accommodating Others: The Role of Authorship as Medium'
 - Informal discussions with scholars from Computing department
 - Define field more narrowly and review current literature
 - Planning remaining structure of novel

- Term 2
 - Screening & Performance at Elephant West, London (Jan 2020)
 - Installation (Goldsmiths): participatory experiment on 'predictability coefficient' in perception of agency in animated, abstract figures
 - Develop video performances
 - Review of literature & potential interdisciplinary collaborations
 - 2 novel chapters

- Term 3

- Learning to use Google Tensorflow
- Write paper: 'The Blind Methodology'
- Synthesise output from first year into draft chapters of thesis
- Complete first draft of novel

Year 2 Finalising more complex artworks

- Term 1
 - Feedback on novel draft
 - Write third paper
 - Develop one-woman play from shorter video performances
- Term 2
 - Rehearse play
 - Rework novel
 - Upgrade: synthesise output so far to include current state of novel draft, documentation of new performance works, experimental results and thesis composed of 3 edited papers
- Term 3
 - Feedback on second novel draft
 - Perform play for a live audience
 - Write fourth paper
 - Update thesis as a whole, to be 70% complete.

Year 3 Synthesising projects into comprehensive artworks and thesis

- Term 1
 - Finalise novel
 - Complete first draft of thesis
- Term 2
 - Exhibit shorter video works and perform play live
 - Rework and fine-edit thesis
- Term 3
 - Finalise minor details in thesis (1)
 - Novel complete and ready for publishing (2)
 - One-woman play complete and ready for more ambitious staging (3)
 - Documentation of collaborative projects synthesised as part of thesis or separately as appropriate (4)

Bibliography

Barad, K. (2007). *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham, North Carolina: Duke University Press.

Brooke-Rose, C. (1999). *Subscript*. Manchester: Carcanet.

Bryson, J. (2012). 'Structuring Intelligence: The Role of Hierarchy, Modularity and Learning in Generating Intelligent Behaviour', in McFarland, D., Stenning, K. and McGonigle-Chalmers, M. (eds.) *The Complex Mind*. Palgrave MacMillan.

Burke, S. (1995). *Authorship: From Plato to the Postmodern, A Reader*. Edinburgh: Edinburgh University Press.

Dawkins, R. (1991). *The Blind Watchmaker*. London: Penguin Books.

Dennett, D.C. 1991, 'Real Patterns', *The Journal of Philosophy*, vol. 88, no. 1, pp. 27-51.

Dennett, D.C. (2004). *Freedom Evolves*. London: Penguin Books.

Derrida, J. (1994). *Specters of Marx: The State of the Debt, the Work of Mourning and the New International*. New York: Routledge.

Derrida, J. (2002). "As If It Were Possible, 'Within Such Limits'...", in Rottenberg, E. (ed.) *Negotiations: Interventions and Interviews (1971-2001)*. Stanford, CA: Stanford University Press, pp. 343-370.

Frow, J. (2014). *Character and Person*. Oxford: Oxford University Press.

Gell, A. (1998). *Art and Agency: An Anthropological Theory*. Oxford: Oxford University Press.

Hebb, D. O. (1949) *The Organization of Behaviour: A Neuropsychological Theory*. New York: Wiley

Hofstadter, D. (2007). *I Am a Strange Loop*. New York: Basic Books.

Johnson, G.A. (ed.) (1993). *The Merleau-Ponty Aesthetics Reader, Philosophy and Painting*. Evanston, Illinois: Northwestern University Press.

Strathern, M 2012, 'The Whole Person and Its Artifacts', *Annual Review of Anthropology*, vol. 33, no. 1, pp. 1-19.

Ward, M. (2018). *Seeming Human: Artificial Intelligence and Victorian Realist Character*. Ohio: Ohio State University Press.

Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL: Wolfram Media.

Zunshine, L (ed.) (2010). *Introduction to Cognitive Literary Studies*. Baltimore: Johns Hopkins University Press.