Connection Science

ISSN: 0954-0091 (Print) 1360-0494 (Online) Journal homepage: http://www.tandfonline.com/loi/ccos20

Reverse engineering the human: artificial intelligence and acting theory

Donna Soto-Morettini

To cite this article: Donna Soto-Morettini (2017): Reverse engineering the human: artificial intelligence and acting theory, Connection Science, DOI: 10.1080/09540091.2016.1271398

To link to this article: http://dx.doi.org/10.1080/09540091.2016.1271398

Published online: 13 Jan 2017.

Submit your article to this journal

Article views: 96

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=ccos20

Download by: [University of Newcastle, Australia] **Date:** 17 January 2017, At: 11:03

CONNECTION SCIENCE, 2017
http://dx.doi.org/10.1080/09540091.2016.1271398

Reverse engineering the human: artificial intelligence and acting theory

Donna Soto-Morettini

School of Arts and Creative Industries, Edinburgh Napier University, Edinburgh, UK

ABSTRACT
In two separate papers, Artificial Intelligence (AI)/Robotics researcher Guy Hoffman takes as a starting point that actors have been in the business of reverse engineering human behaviour for centuries. In this paper, I follow the similar trajectories of AI and acting theory (AT), looking at three primary questions, in the hope of framing a response to Hoffman's papers: (1) How are the problems of training a human to simulate a fictional human both similar to and different from training a machine to simulate a human? (2) How are the larger questions of AI design and architecture similar to the larger questions that still remain within the area of AT? (3) Is there anything in the work of AI design that might advance the work of acting theorists and practitioners? The paper explores the use of "swarm intelligence" in recent models of both AT and AI, and considers the issues of embodied cognition, and the kinds of intelligence that enhances or inhibits imaginative immersion for the actor, and concludes with a consideration of the ontological questions raised by the trend towards intersubjective, dynamic systems of generative thought in both AT and AI.

ARTICLE HISTORY
Received 19 June 2016
Accepted 8 November 2016

KEYWORDS
Embodied cognition; artificial intelligence; acting theory; swarm intelligence; imaginative immersion

> Theatre actors have been staging artificial intelligence for centuries. If one shares the view that intelligence manifests in behaviour, one must wonder what lessons the AI community can draw from a practice that is historically concerned with the infusion of artificial behaviour into such vessels as body and text . . . . Therefore, acting methodology may hold valuable directives for designers of artificially intelligent systems.

In his paper, *On stage: robots as performers* (2011), human–robot interaction (HRI) specialist, Guy Hoffman takes as a starting point that actors have been in the business of reverse engineering human behaviour for centuries. In other words, actors work from observable behaviour backwards, to discover motivation, intention, desire, etc. Of course an actor cannot consider imagined intentional states without an accompanying consideration of many varied external factors, such as the social/human forces that will affect decisions, character-specific psychology or contextually specific social "display rules" that might govern just how much of a character's "inner state" can or will be expressed. Still, for all this complexity, the actor is in the business of analysing human intelligence and in manifesting intentional

CONTACT Donna Soto-Morettini d.soto-morettini@napier.ac.uk Edinburgh Napier University, 10 Colinton Road, Merchiston, Edinburgh EH10, 5DT, UK

© 2017 Informa UK Limited, trading as Taylor & Francis Group

states through behaviour, and this makes the area of acting theory[1] (AT) of interest in relation to theories and practice in artificial intelligence (AI).

Specifically, Hoffman narrows his interest in AT down to two areas: continuity and responsiveness. He references acting practitioner/theorists (Stanislavski, Sonia Moore, Michael Chekhov, Augusto Boal, Sanford Meisner) and makes specific recommendations for robotics design based on AT, outlining a number of applications to robotics, including programming human-interactive robots with an "inner monologue" that might lead to more fluid responses, and exploiting "Meisner 'responsiveness'" to interactive robot design to create a more anticipatory response.

In this paper, I want to look at three primary questions:

(1) How are the problems of training a human to simulate a fictional human both similar to and different from training a machine to simulate a human?
(2) How are the larger questions of AI design and architecture similar to the larger questions that still remain within the area AT?
(3) Is there anything in the work of AI design that might advance the work of acting theorists and practitioners?

In order to consider these questions, I'd like to look closely at two areas that Hoffman's brief paper addresses: (1) embodied cognition (psycho-physical unity) and (2) the location of responsiveness/action choice and the problems of 'single agent' design.

The imitation game

Perhaps these days when thinking about the field of AI, most people think of the pioneering work of Alan Turing, whose 1950 article, *Computing machinery and intelligence* opened with "I propose to consider the question: can machines think?" Turing carries on to describe his now well-known "imitation game", which was an early challenge to the design of pure AI. Turing's game is played by a man, a woman and an interrogator who, without seeing or hearing them, must determine which is the man and which the woman. This opening game then leads to his main interest:

> We now ask the question, "What will happen when a machine takes the part of [the man] in this game?" Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, "Can machines think?" (Turing, 1950, p. 433)

In the 60-plus years since he posed his challenge, this question remains at the heart of the evolving field of AI. But it is the more direct problem that Turing poses here – the relationship between the "real" and the simulated – that has bedevilled actors for centuries: can an actor A convince auditor C that his imitation of Hamlet is 'real'? And perhaps in that sense, we might say that much theatre criticism, and even the experience of watching an actor perform, are forms of Turing's test.

Actors, of course, have a great advantage over robots in playing the imitation game. Being human, they already have a vast store of human experience to draw upon. They move and think and behave in human ways, have immediate access to their own intentional states, can draw quickly on abstract symbolic processing systems, are equipped with natural speech function and the ability to avoid obstacles when moving through space.

Still, when encountering Hoffman's brief paper, I was struck by how many of the things AI designers face are things that still challenge those who are in the business of directing or training actors. It is striking how often the way in which we describe an unconvincing actor is to say that their work is robotic or has a robotic quality. This sometimes refers to the way in which actors can be physically stiff and uncomfortable in their bodies; sometimes to the ways in which actors seem to be delivering their lines without a requisite (illusionary?) depth of thought behind the delivery, and sometimes it refers to the ways in which a performance can seem 'calculated' or pre-determined. It is common to hear all of these types of performance described as 'robotic'. And although it may be no more than an old actor's joke to recall the experienced actor's advice to the novice: "don't bump into the scenery", it still reminds us of the way in which learning actors face many of the AI designer's problems.

Conversely, there are critical accounts of how actual robots have gone wrong, and these descriptions may remind us of the way that acting teachers often describe the difficulties their students are having. A case in point might be the description (Brighton & Selina, 2003) of William Gray Walter's robot, Elsie:

- "She did not have knowledge of where she was or where she was going
- She was not programmed to achieve any goals
- She had little or no cognitive capacity." (p. 68)

Most actors will recognise in this short list the basics of Stanislavski's critique of the acting he wished to transform with his ideas, and his emphasis on the importance for individual actors to focus their preparation by paying close attention to their character's objectives, given circumstances, and the development of the character's thought process. Consequently, if it is the case that we describe bad actors as robotic, then perhaps under-performing robots may be described as bad actors.

Evolving theories of AI and robotic human simulation

AI has evolved much in regard to theories about the relationship between brain, body and mind, and it is fair to say that what has been come to be called GOFAI (Good Old Fashioned Artificial Intelligence) was largely concerned with pure AI. Turing's interest was in whether machines could think, and thinking in terms of the imitation game was about a computational intelligence. Of course here is where all the challenges for AI begin, because there are so many ways of defining intelligence and ways of talking about what AI means in relation to creating artificial forms of intelligence.

Without straying too far from our interests, I think it is important to note the distinction between 'strong' and 'weak' forms of AI research. Strong AI more or less holds the opinion that humans are elaborate computers, and that if one could crack the codes within that elaborate structure, and identify its constituent parts, then human consciousness – thoughts, feelings, emotions – could be wholly replicated. In that sense, 'strong AI' is probably of some interest to the actor, since we could argue that the labour expended by the actor in rehearsal – analysing the components of and the influences on Macbeth's behaviour – allows the actor to replicate/simulate the human consciousness of Macbeth. That replicated/simulated consciousness is, of course, metaphorically parasitic[2] on the actor's own human consciousness, but it is still a process of replication/simulation. Always cognisant

of ontological difficulties here (and which I will address in conclusion), I would like to pose some questions arising in terms of the actor's working conceptions of intelligence and ask: does this suggest that actors have some sympathy with the strong AI community's belief that humans are elaborate biological, computational entities, whose behaviour is *generated* by complex (but ultimately replicable/comprehensible) thought processes? Do actors balance their representations of human behaviour on a belief that by understanding and identifying certain intelligent components of behaviour we execute decisions that replicate the human behaviour of a character other than our own character? Is it the case that we disaggregate human behaviour (reverse-engineer the observables), following links back to the unobservable, discrete intentional states and desires and then create an executable cognitive programme? While many make the case that Stanislavski's early theories were largely concerned with this kind of activity,[3] we know that even Stanislavski himself recognised the limitations of this 'top-down' model. And while there are varieties of AI research that follow this 'top-down' model, many evolving theories of AI are working in the other direction.

The limits of 'top-down' modelling have become as clear to AI researchers as it did to Stanislavski in his later writings, and the function of the body as an element of cognitive processes has become an area of increased interest both in AI and AT. Much current AI is concerned with 'bottom-up' models that proceed from the idea that in order to create a model of human cognition, it is necessary to think in terms of how humans produce action in response to specific environments. This is one area where the relationship between AI and AT grows particularly interesting, and we can follow the theories simultaneously.

Traditionally, most AI research has worked from a SENSE → THINK → ACT (S-T-A) model: a linear paradigm of perceptual sense/sensor → leading to computational programme → leading to choice of motor action. It is fair to say that traditional (Stanislavski) AT has worked in this way too: perceptual information about environment and given circumstances → leading to a cognitive process that considers choices → leading to a specific action. But both AI and AT have developed models that have moved away from S-T-A to reconsider the complex relationship between each part of this tripartite linear model, resulting in a 'bottom-up' approach in both areas. This relocation of interest has occurred alongside (or is perhaps a result of) the growth of interest in the notion of embodied cognition, or a re-evaluation of the relationship between action–perception systems and semantic processing.

Hoffman (2012) explains the way that a bottom-up process operates in his work in robotics:

> [T]he role of action and motor execution in robotics has traditionally been viewed as a passive "client" of a central decision-making process, and as such at the receiving end of the data and control flow in robotic systems. Even in so-called Acting Perception frameworks (Aloimonos, 1993), the influence of action on perception is mediated through the agent changing its surroundings or perspective on the world, and not by internal processing pathways. Instead, we suggest that action can affect perception and cognition in interactive robots in the form of symmetrical action-perception activation networks. In such networks, perceptions exert an influence on higher level associations, leading to potential action selection, but are also conversely biased through active motor activity. (p. 5)

Hoffman's description suggests that perception (which belongs, traditionally, to the first part of SENSE → THINK → ACT)[4] is not part of a discrete, tripartite model, but is permeable and subject to *bias* through motor activity. This idea emerges from relatively new ways to conceive the ACT part of the S-T-A formula, and more specifically from recasting action

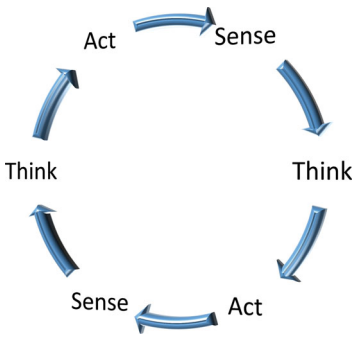
itself as something other than "passive client". The subsequent challenges in attempting to locate, precisely, the locus of thought/action is as interesting and critical for AI as it is for AT. For AI, and for Hoffman in particular, the question is how to theorise and realise the complex relationship between the sensory/perceptual and action choice.

That complex relationship may be seen differently, of course, depending on the role of THINK in our linear structure, particularly since many branches of cognitive science and AI research significantly devalue (or at least reformulate) the ways in which the traditional top-down model has worked. Hoffman's embodied alternative rests on a seminal article by Brooks (1991), in which Brooks lays out his belief that

> human level intelligence is too complex and little understood to be correctly decomposed into the right subpieces at the moment, and that even if we knew the subpieces we still wouldn't know the right interfaces between them. Furthermore, we will never understand how to decompose human level intelligence until we've had a lot of practice with simpler level intelligences. (p. 140)

Concerned largely with "low-level behaviours (such as object avoidance, walking dynamics, and so forth)", Brooks proposed looking to insects and other simple organisms for inspiration, and, as Hoffman points out (2012), for "growing" intelligence as an emergent property of increasingly complex systems. This emergent intelligence from complex interactive systems seems to me to lie at the heart of some recent developments in AT, particularly Viewpoints and Suzuki actor training systems, and we will consider this later.

But importantly, Hoffman's reformulation of SENSE → THINK → ACT is entirely concerned with the way in which perception may be biased towards motor activity, and this, in turn, may relate well to what Pulvermüller (2013) has lately proposed as a kind of "integrative-neuromechanistic explanation of why both sensorimotor and multimodal areas of the human brain differently contribute to specific facets of meaning and concepts" (p. 86). Pulvermüller's influential paper looks very specifically at research into the ways that "resources in cortical motor systems engaged by complex body-part-specific movements are necessary for processing of semantically congruent action words" (p. 102). In the studies that Pulvermüller refers to, it was demonstrated that involving, for example, the arms or legs in complex movement patterns resulted in some memory impairment of arm- or leg-related words. This suggests that the modalities required to process the complex movement itself created a deficit in the action-language processing, and this is turn has significant influence on the way that we see the relationship between action–perception and semantic processing. Ultimately, perhaps, we must accept the limitations of our linear model and consider the relationship between action, perception and cognition to be more accurately illustrated as:



The complex questions surrounding the ways in which we can conceptualise the workings of embodied cognition are as challenging for acting theorists as they are for those involved in the area of AI, and raise a number of questions – only some of which can be answered by reference to scientific literature. I would suggest that we need to think of embodied cognition in very specific ways with regard to AT, because it is one thing to acknowledge the interwoven complexity of perception/cognition, but that acknowledgment does not suggest a separate, 'bodily cognitive' process. In other words, there are ways in which we can talk about embodied cognition that is meaningful for actors but we can't, of course, turn to our knee to ask for advice on how to play Hamlet. That is, of course, a simplification, but there are some implied models of embodied cognition in AT that seem to suggest that there is a kind of brute, unmediated intelligence that inheres in the body and that if we can access this unmediated bodily intelligence we can create conditions for better, more spontaneous performance. For all their success in terms of training, I think that some contemporary acting training systems are focused somewhat in this way, and I want to look specifically at Anne Bogart and Tina Landau's Viewpoints[5] system and Tadashi Suzuki's actor training system.[6]

Like AI, these versions of AT are tinkering with the many possibilities in reframing a SENSE → THINK → ACT model of human behaviour. Also like AI, these methods represent theories that have evolved from a kind of 'top-down' early Stanislavskian model (which, as we have noted, Stanislavski himself amended by stressing the importance of a psychophysical approach to the analysis and playing of a role), to contemporary 'bottom-up', action/physically based training. In both systems, actors work first from the physical body in a very specific physical environment, and the general aim of both is towards redirecting the actor's attention, in the hope of finding greater freedom and expressiveness in performance. The question is, how do such systems do that?

Viewpoints is a combination of a number of ideas from earlier practitioners (drawing particularly on things like Rudolf Laban's efforts or Michael Chekhov's psychological gesture), and they are very clear about the non-intellectual basis of their approach, suggesting that their method is "a tool for discovering action, not from psychology or backstory, but from immediate physical stimuli" (p. 125). The Viewpoints system comprises a set of specific physical exercises and explorations of space, time and environment. It is particularly focused on creating a kind of 'collective consciousness' of creativity that can help the performer to allow inspiration or perhaps subconscious thought to flow freely. Students are advised to keep a "soft focus" in the room, by which they mean visual perception that isn't aligned with an intentional state, such as "want" or "desire", and they describe a situation in which participants can "learn to listen with our entire bodies and see with a sixth sense" (p. 20). Some of the exercises described in the first part of their book are not dissimilar to older training methods whereby actors "mirror" each other's actions, or attempt to "feel" the moment that a whole group will jump into the air, but their system is much more codified and philosophically distinct. The writers, Bogart and Landau, are working towards their stated aim of creating a responsive ensemble which takes precedence over an intellectual, 'single agent' approach to performing. This collective way of working means that "pressure is released from any one person who feels that they have to create in a vacuum. Emphasis is placed on the fact that the piece will be made by and belong to everyone in the ensemble" (p. 122).

Students are advised against expecting any pre-determined outcomes, and are instead encouraged to "let information come" to them, and to let "something occur onstage, rather than making it occur. The source for action and invention comes to us from others and from the physical world around us" (p. 19). The desired state is one in which "when something happens in the room, everybody present can respond instantly by bypassing the frontal lobe of the brain in order to act upon instinct and intuition" (p. 33). In short, I would argue that the system outlined in their book emphasises a '4E' approach to actor training that is embodied (lead by and discovered through physical interaction), embedded within an ensemble process, extended (in the sense that cognitive choices are mutually created and shared), and enactive (in that the locus of all generative thought and action is intersubjective).

Bogart and Landau write in the style of creative artists, and certainly one would not expect to hear them describe their work in the language of philosophy or cognitive science. But perhaps there is some sympathy between their ideas of bypassing the frontal lobe, and the work of those non-representational theorists like Brooks, or Andy Clark, neither of whom work through a centralised system of modelling environments. Brooks, for example, proposes a system of multiple parallel activities that do n'ot rely on any central representation. These multiple activities avoid a "representational bottleneck" which Clark (1998), in turn, describes as an impedance to real-time responses within an immediate physical environment. In his book, *Radical embodied cognitive science* Chemero (2011) lays a philosophical defence for a theory of computational embodied cognition that generates action-oriented representations (a term first used by Clark), which he sees as constituently different from traditional systems that hold representation to be neutral: "Action-oriented representations differ from representations in earlier computationalist theories of mind in that they represent things in a non-neutral way, as geared to an animal's actions, as affordances" (p. 26). Like Bogart & Landau, Chemero sees cognition as ecological and only understandable in terms of action in a specific environment. It is likely that Viewpoints methods have more to do with ecological psychology and "action-oriented" representation than in the direct suppression of thought when they speak of "by-passing". By emphasising the discovery of action as an environmental response, their method is probably better understood as a systematic approach to redirecting attention. That sounds extremely simple, but in practice, redirecting attention in a sustained way can be quite challenging and have surprising results. One finds a similar theme in the writings of Suzuki (1990), who explains that "our psychological and our physical movements do not naturally blend together" (p. 38), which reminds us that theoretical dualism is often located in the ontological assumptions of the worrying theorist, an area I have explored previously (Soto-Morettini, 2010). Suzuki writes extensively in *The Way of Acting* about the importance of the body in space. He is convincing in his idea that actors must direct attention to the body (particularly the feet) and its relation to environment, but of course before we can put his ideas into practice, we must consider the statement that the psychological and the physical do not naturally blend together.

Suzuki makes this statement specifically in the context of contemporary acting practices in relation to the *nō* theatre practice, which he describes as having begun with

> the expulsion of any expression based on psychology and individual personality; it exists on a communality that risks no dispersion . . . .In order for the *nō* to have developed in all its

> grandeur, there had to lie behind it the existence of a fixed, decidedly communal playing space. (p. 39)

Suzuki's worry mirrors Bogart and Landau's: the individual brain seems to be getting in the way of our need to respond spontaneously within a specific physical environment. Their mutual solution seems to be a relocation of interest from the 'single agent' brain to the emergent happenings from a kind of swarm model of the type advocated by Brooks and others.

But what might AT and AI gain from this turn to swarm modelling? For AI, the larger picture is related to the possibility of building, through multiple layers of parallel activity, a genuinely real-time responsive environmental intelligence. I believe it also represents a way of working towards complexity by practicing with simpler levels of intelligence (Brooks, 1991) as a necessary starting point. What AT gains, and what lies behind the desire to explore in these 'bottom-up' models, is less the kind of afferent/efferent exchange that Hoffman talks about and much more about attempting to help the actor suppress or redirect the thinking brain before it can interfere with ("bottleneck"?) something that feels as if it arises intuitively and spontaneously from the body and the body's experience.

The larger question for AT is whether, by practicing with simpler levels of interaction and re-directed attention (or "bypassing the frontal lobe"), we can ever hope to build the multiple layers of complexity needed to perform Rosalind. And herein lies some irony. The human–robotic interaction wing of AI is trying to model behaviour that reflects genuine human thinking processes and AT is trying to model behaviour that suppresses, or redirects some of that genuine human thinking process. And yet, they are in some cases both using a similar 'swarm intelligence' model.

Of course, for AT the 'swarm' model is only concerned to suppress or redirect some kinds of thinking and that is the issue. We are in rather value-laden territory here and that means we must come clean (and, I would suggest, become clearer) about the kinds of cognitive activity we worry about in AT, and the ways in which we judge those cognitive activities. To demonstrate this, we need look no further than a quotation from John Lutterbie's excellent book *Towards a general theory of acting* (2011), in which he considers dynamic systems theory to be of use in "resisting the tendency to differentiate between qualities that are seen as valuable (emotional availability) and qualities that are denigrated (intellectualism)" (p. 103).

Lutterbie's reference to denigrated intellectualism reminds us of Viewpoints's desire to bypass the frontal lobe, and immediately poses the question for us: what is "intellectualism" in acting, and in what way is it denigrated? Lutterbie talks about a "binary of intellect and emotion" that arises when interviewing actors about their practice and concludes that "Rational thought needs to be put on hold so that the experience of images and emotional responses can play freely across and through the body . . . .' (p. 5) There is much that can be said about this – perhaps the first is that it is difficult not to intuit here a rewriting of Cartesian dualism (rational/emotional standing in for mind/body) and the second is that this dualism is a prevalent part of the actor's language, most commonly articulated as "being-in-the-moment" (externally focused attention) vs intellectualising about the action (internally focused symbolic/semantic processing).

For both Lutterbie and Viewpoints, when internal rationality or psychology are involved, the result is the inhibition of external spontaneity or intuitiveness. That inhibition disrupts

the balance of imagination and 'belief' – both of which are elusive and fundamental qualities of the actor's work.

Being-in-the-moment and the process of imagination

The aim of Viewpoint's practice is in redirecting attention outward, towards environment, creating a stronger connection between empathic physiology and constantly shifting circumstances, thereby allowing the performer to remain spontaneous and "in the moment". But however desirable that idea is (and in a common-sense manner we can see that if a performer can remain wholly in the moment – attending only to the circumstances, adapting and responding organically and seamlessly – a performance is likely to feel spontaneous and fresh), but as many would argue, consciousness is not really structured this way. The defining quality of consciousness for neuroscientist Gerald Edelman is not simply the present, but "the remembered present". In a particularly provocative conclusion to his influential book, he writes: "Consciousness may be seen as the haughty and restless second cousin of morphology. Memory its mistress, perception its somewhat abused wife, logic its housekeeper, and language its poorly paid secretary" (Edelman, 1989). Edelman's point is that consciousness is not, by definition, an attendance to the here-and-now; not "in the moment" without its household retinue, described above. Rather, "recursive activation of neuronal systems allow organisms that act in the moment and in a particular problematic situation, to redeploy patterns of behaviour that remain continuous with past forms" (p. 196).

For philosopher Dennett (1993), the problem of being 'in-the-moment' would require that we define a "high point" of consciousness which, in his view, never comes. Instead, he describes consciousness as a dynamic process, during which we continually revise through "multiple drafts" whatever we are currently in the process of experiencing. In both Edelman and Dennett's metaphors, we can intuit something about the ways in which consciousness not only works synchronically in past–present–future remembrance/anticipation, but also the ways in which natural human thought moves dynamically in various states of informational input, adjusting, responding and acting as environment evolves. This idea has much resonance with Guy Hoffman's suggestion that robot–human interaction could be strengthened by more anticipatory programming, in which robots might begin to respond with only partial information, which they can revise as new information becomes available. Endowing robots with a kind of 'inner monologue' (a term that Hoffman borrows from AT) that constantly monitors environment would allow them to achieve a more fluid interaction than standard command-and-response behaviour:

> A possible implementation is an opportunistic action selection mechanism, in which robot activity is continuous . . . .The robot can, at any point in time, choose to fully produce one of several implied actions. The actual action selection occurs opportunistically based on current perceptual processing (2011, p. 3).

Of course, for the actor, more than just spontaneous, fluid response is required. The actor works within an imagined world and all of the events within that world must be endowed with a plausible simulation of belief. And if the actor's work is largely a labour of imagination, we are faced with some questions: what are the defining the qualities of imagination

in practice, and does the practical imagination gain as much through an embodied environmental context as it might in the kind of isolated imaginative immersion that we sustain in reading a novel? These questions are bedevilled by the very fact that, as Kaag (2008) points out, "The imagination is difficult to define. More often than not, it is not defined at all – only invoked as a placeholder by philosophers when they are unable to define particular cognitive processes" (p. 183). In an article drawing heavily on the work of Gerald Edelman (amongst others), and he defines the neurological basis of imagination as resting largely on the "ecological" processes of mirror neurons in terms of action/perception:

> The research on the mirror neuron system is significant in our investigation of the imagination in the sense that it begins to point to a physiological process that allows organisms to be in touch with their local situations, make generalizations from partial observations, and to adapt to their particular circumstance in the continuous flow of inquiry, learning and adaptation. (p. 194)

Kaag's descriptions remind us of the balance between spontaneous response to environment and the intricate "neuronal architecture" at work in the actor's business of replicating a 'parasitic' intelligence. But if Kaag is right about the physiological basis of imagination, then imagination itself works environmentally, and navigating even an imagined environment requires externally directed attention. If this relationship between physiological processes and environment is the key to responding to particular circumstances, either real or imagined, in an immediate and fluid way, then we must admit the importance of training attention to external circumstances for both performers and robots. In a sense, this is a de-centralising move in terms of developing theories of simulating behaviour. The reward of that de-centralising is an environmental, responsive performance that is greater than one might expect from a performance generated through a single-agent approach. In his description of neural circuitry, Kaag deploys the example of jazz musicians who have played together for many years and Edelman uses the same analogy, concluding that the surprising integration heard in a good jazz improvisation is "a kind of mutually coherent music that each one acting alone could not produce" (p. 197). Perhaps when it comes to AT one of our most challenging difficulties is that the very origin of AT was articulated in terms that has and sometimes still does address acting as a single-agent activity. The legacy of titles like *An actor prepares,* or *Building a character* might have been entirely different if those titles had been "*Actors prepare*" or "*Building characters*".

Edelman's reference to jazz improvisation brings us back to Hoffman's work with robots, which has recently focused on improving audience reception of robots improvising with jazz musicians. In that work, he describes the importance of focusing the robot's attention on the human musician, and points out that the first theorist of a relocated space of an actor's focus was Sanford Meisner, whose work at some points encouraged the actor to focus much more on what Viewpoints would call "letting information come to you". The point of this relocation of focus is what drives work in multi-agent system design and its advantages are well articulated by Eberhart and Kennedy (2011), looking beyond single-agent theories in AI:

> Real social interaction is exchange but also something else, perhaps more important: individual exchange rules, tips and beliefs about how to process the information. Thus a social interaction typically results in a change in the thinking processes – not just the contents – of the participants. (pp. xiv–v)

In his article on musical improvisation involving a robot (2011), Guy Hoffman proposes that "musical performance is as much about visual choreography and visual communication as it is about tonal musical generation" (p. 20), which makes clear the way in which creating a robot that can convince an audience of its "proto-human" abilities requires programming that is highly interactive. Hoffman's work with the marimba-playing robot focuses heavily on anticipation in order to create "real-time musical coordination" (which echoes the actor's desire to work "in the moment"). He found that by locating much of the robot's attention onto visual communication between the robot and the human player, the audience "rate the robot as playing better, more like a human, as more responsive and as more inspired by the human" (p. 20).

Conclusion

In philosophical terms, for actors and robots, it feels as if we have travelled some distance now from either our linear or our circular S-T-A models – both of which suggest that the sensing, thinking and acting are all located within a single agent. Effective performance it seems, both for robots and for humans, is located/generated somewhere in the intersubjective space between robot:human and actor:ensemble. The idea of the intersubjective space in practice between actors is a particularly intriguing one. Perhaps what Viewpoints practitioners are suggesting is something more profound than an actor representing (and therefore comprehending) the mental state of a performing partner; perhaps their ideas are closer to an "enactive intersubjectivity" (Fuchs & DeJaegher, 2009), wherein the process itself generates common meaning:

> According to our concept, social understanding is primarily based on intercorporality; it emerges from the interactive practice and coordination of the persons involved. We do not need to form internal models or representations of others in order to understand and communicate with them. Social cognition rather develops as a practical sense, a musicality for . . . rhythms and patterns . . . (p. 485)

This idea leads us in turn to some of the tougher questions surrounding multi-agent systems and suggests that in querying the location of generative thought we are naturally raising some difficult ontological questions. This is an area of much growth and study currently, well outlined in Petrov and Scarfe (2015), which addresses dynamic aspects of being and also "the process-relational character of being itself" (p. viii). While these deeper ontological questions are beyond the scope of this essay, it seems appropriate to raise the development of work in this area, and to suggest that if there are similarities in ontological considerations that apply both to human and robot multi-agent systems, there are also many differences. Seibt (2016) rejects entirely the "temptingly easy strategy of treating HRI as fictionalist analogues to human-human interactions", and suggests that what is required where human:robot interaction is concerned is "an entirely new 'classificatory framework' for simulated social interactions" (p. 5).

I began this article looking at Guy Hoffman's assertion that AI could learn from AT, but I hope that a brief look at similarities in both areas will lead us to an inverted and balancing assertion: that AT has much to learn from AI, and I would suggest that such learning can only come about through a continued interdisciplinary dialogue and exchange about how human behaviour is structured as part of a dynamic social environment and how various approaches to simulated models work within specific contexts. This in turn would require

that acting theorists and practitioners enter the debate in more rigorous terms regarding the kinds of epistemological and ontological questions we raise in our practice.

If my look at the similarities in the developmental trajectories of AT/practice and AI is in danger of falling into the "temptingly easy" danger Seibt describes, I hope that it does at least raise some interesting questions for both areas. While much of the earlier works of AT writers has centred in one way or the other on a kind of "inner programming", newer writing in the area seems to concern itself heavily with the kind of knowledge for actors that emerges without the "inner" or central control. And if the ontological differences are acknowledged, it is still safe to say that the aims of the HRI wing of AI – as developed by Guy Hoffman and others – are closely aligned to those of the acting community: to create embodied, enactive experiences that convince and engage their audiences completely.

Notes

1. I employ the term 'acting theory (AT)' here in the way that Hoffman does – in relation to acting practitioners who have written about their practical methodology of training/directing actors.
2. I use 'parasitic' here as a simple metaphor for the character/actor relationship, and not in reference to 'parasitic speech acts' or specific language 'performativity'.
3. Most recently, John Lutterbie speaks of Stanislavski's early works as not addressing the body, "other than to acknowledge that the work of the actor must be given a form in order to be communicated to the spectator, which work requires bodily techniques. Beyond this acknowledgment, Stanislavski understands the body to contain the work done by the actor." He goes on to describe the change in Stanslavski's writings, in which he "rethought the role" of the body entirely. 31–32.
4. The difficulties of course with 'SENSE → THINK → ACT' models lie in the very linearity that is described in the model, and many contemporary researchers in AI are pursuing, amongst other things, 'loosely coupled' or, 'parallel processing' systems which attempt to solve the limitations of linear models with more connectivist, networked or parallel models. These things are beyond my work here, but we must recognise that for all its ease of description, there are, of course, problems with any simple, linear model of cognition.
5. The references to Viewpoints throughout this article refer specifically to *The viewpoints book* by Bogart and Landau (2005), and not to workshop participation.
6. References to Suzuki's training refer specifically to his book, *The Way of Acting*, and not to workshop participation.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Aloimonos, Y. (Ed.). (1993). *Active perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Bogart, A., & Landau, T. (2005). *The viewpoints book*. New York, NY: Theatre Communications Group.
Brighton, H., & Selina, H. (2003). *Artificial intelligence: A graphic guide*. London: Icon Publishing.
Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*(1–3), 139–159. doi:10.1.1.12.1680
Chemero, A. (2011). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
Clark, A. (1998). *Being there (putting brain, body and world together)*. Cambridge, MA: MIT Press.
Dennett, D. (1993). *Consciousness explained*. London: Little, Brown.
Eberhart, R. E., & Kennedy, J. (2011). *Swarm intelligence*. San Francisco, CA: Morgan Kaufman Publishing.
Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

Fuchs, T., & DeJaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, *8*(4), 465–486. doi:10.1007/s11097-009-9136-4

Hoffman, G. (2011). *On stage: Robots as performers*. Unpublished paper. Retrieved from http://guyhoffman.com/publications/HoffmanRSS11Workshop.pdf

Hoffman, G. (2012). Embodied cognition for autonomous interactive robots. *Topics in Cognitive Science*, *4*(4), 759–772. doi:10.1111/j.1756-8765.2012.01218.x. http://guyhoffman.com/publications/HoffmanTopiCS12.pdf

Hoffman, G., & Weinberg, G. (2011). Interactive improvisation with a robotic marimba player. In J. Solis & K. Ng (Eds.), *Musical robots and multimodal interactive multimodal systems*. Berlin: Springer Tracts in Advanced Robotics. Retrieved from http://guyhoffman.com/publications/HoffmanAuRo11.pdf

Kaag, J. (2008). The neurological dynamics of the imagination. *Phenomenology and Cognitive Neuroscience*, 8:2. Springer Press (Winter). doi:10.1.1.471.2279

Lutterbie, J. (2011). *Toward a general theory of acting: Cognitive science and performance*. New York, NY: Palgrave Macmillan.

Petrov, V., & Scarfe, A. C. (Eds.). (2015). *Dynamic being: Essays in process-relational ontology*. Newcastle-Upon-Tyne: Cambridge Scholars Publishing.

Pulvermüller, F. (2013). Semantic embodiment, disembodiment, or misembodiment? In search of meaning in modules and neuron circuits. *Brain and Language*, *127*(1), 86–103. doi:10.1016/j.bandl.2013.05.015

Seibt, J. (2016). Draft unpublished article (in press), *Towards an ontology of simulated social interaction–varieties of the 'As If' for Robots and Humans*. Retrieved from http://www.academia.edu/19415782/Towards_an_Ontology_of_Simulated_Social_Interaction–Varieties_of_the_As_If_for_Robots_and_Humans

Soto-Morettini, D. (2010). *The philosophical actor*. Bristol: Intellect Books.

Suzuki, T. (1990). *The way of acting*. New York, NY: Theatre Communications Group.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *49*, 433–460. Retrieved from http://phil415.pbworks.com/f/TuringComputing.pdf