How Many Robots Does It Take? Creativity, Robots and Multi-Agent Systems

Stephen McGregor and **Mariano Mora McGinity** and **Sascha Griffiths**[1]

Abstract. This paper seeks to situate computational creativity within the context of ongoing theoretical and practical investigations of environmentally situated and dynamic systems. Beginning with a consideration of the evidently goal directed nature of creativity, the problem of how teleological behaviour emerges in a fundamentally physical world. Creativity is reassessed as a search for goals in a dynamic environment rather than as a pursuit of a fixed goal in a stable and finite space of possible actions. A significant consequence of this evaluative shift is the impossibility of considering truly creative systems as anything other than embodied agents deeply entangled in an environmental situation. Two fields are discussed as potential habitats for such systems: robotics and multi-agent systems. Creativity from the perspective of ongoing research in these areas is considered, and some preliminary thoughts for future directions of enquiry are offered.

1 Introduction

This paper will address the question of the relationship between goals and creativity. Notions of purpose are so deeply ingrained in the standard view of creation that creativity itself is often defined in terms of the accomplishment of some expressive objective. Implicit in the problem of modelling creativity, however, is the emergence of end directed action in a reductionist world: how can something that is not in a physical sense present nonetheless contribute to the operation of a physically supervenient system?

Having posed the question of how a creative agent views its own objectives, the paper will turn to an exploration of the related problem of causality. In particular, the emergence of absent causes – which is to say, the influence of possible worlds, both historical and futuristic, removed from present reality – is addressed. This etiological inquiry is couched in terms of evolution by natural selection, with a brief consideration of this well researched process as a model of evidently goal directed and therefore potentially creative behaviour. A general hypothesis regarding the viability of explaining goals as emergent properties of complex systems, grounded in contemporary theoretical investigations of dynamic systems, will be put forward. *Contra* the idea that computationally creative agents must necessarily be handed a well defined goal by an external designer, dynamic processes are proposed as a basis for models that can discover their own goals through collaboration and environmental interaction.

This theoretical consideration is followed by a preliminary exploration of two compelling areas of research that move beyond what has been the *de rigueur* constraint satisfaction approach to computational creativity. First the topic of robotics will be considered from the perspective of the modelling of creativity, with particular attention to the problem of how a robot obtains, represents, and adapts its own goals. Robots are importantly embedded in a physical environment, and this situation opens the door to the possibility of the emergence of dynamic attractors that might be construed as new and unexpected goals outside any representation of an objective built into a robot's programming. The conclusion of this investigation will be that it seems reasonable to at least consider the possibility of an adequately flexible robot formulating goals that can be considered as evidence of its own creativity.

Next multi-agent models are considered, with particular attention to the ways in which complex patterns of activity with the trappings of intentionality can emerge from interactions within a population of agents individually following very basic sets of predetermined rules. As with robots in their environmental entanglements, swarms of agents have some prospect of generating collective behaviour that can be interpreted as being directed towards ends outside of the simple constraint satisfaction requirements programmed into the functioning of each independent agent. In the case of multi-agent models, the model becomes the environment, with the attractors that arise in the course of interactions becoming the handles for assessing the system in terms of the formulation and pursuit of goals. On the one hand, interpretation of action in a simulation of such a system presumably still falls back on the analysis of an external observer. On the other hand, the emergent properties of such systems potentially offer models of the parallel emergence of cognitive phenomena such as creativity in a physically grounded universe. Again, there seems to be scope for considering the implementation of multi-agent systems as a form of computational creativity that begins with a traditional programming task but that subsequently goes beyond mere constraint satisfaction.

The ideas proposed in this paper are at this stage indications of direction for future research. Exciting work is currently being done in the fields of both robotics and multi-agent systems, with some applications specifically towards modelling creative systems [1, 17]. This paper is intended to serve as a bellwether for further research in this direction, with the objective of moving beyond a constraint satisfaction approach to computational creativity. Traditional rule based implementations of creative agents have accomplished much in recent years, but reconsidering the emergence of goals within highly dynamic environments offers the grounding of a more robust argument for creative autonomy arising within the systems themselves. This reconsideration of the relationship between creativity, goals, and the environmental situation of creative agents can furthermore become a platform for extended discussions of interesting philosophical questions about causation and cognition.

[1] Queen Mary University of London, UK, email: s.e.mcgregor, m.mora-mcginity, sascha.griffiths@qmul.ac.uk

2 Creativity and Goals

> What redeems it is the idea only... and an unselfish belief in the idea – something you can set up, and bow down before, and offer a sacrifice to...
>
> JOSEPH CONRAD, *The Heart of Darkness*

If an agent cannot choose its own goals, can its "creative" behaviour really be considered creative? This question is familiar to anyone who has attempted to construct a creative agents, especially agents that produce works of art. "*Why* did the computer choose this word, this note, this color, and not another?" is generally asked by people confronted with an artistic piece produced by a computer. If the same question were put to an artist the answer would most likely be as unsatisfactory as that provided by the computer, yet it is the computer's response that is most troubling. The assumption is that art made by computers is something that needs to justify itself, whereas it is accepted that artists produce according to their inspiration, perhaps because works of art are seen as reflections of the mind or the personality of the artist who made them, and it is troubling to think of a computer as having a "personality". However, the fuzziness of the artist's response seems to reflect something essential in the creative process, something which can and should be exploited in the design of artificial creative systems: the goal is not fixed; it shifts and drifts and wanders; it might not even exist *a priori* but rather will emerge from the creative process itself.

2.1 Causality and Teleology

Aristotle's "Doctrine of the Four Causes" presents a framework for understanding the relationship between actions and motivations in a world of physical events [2]. Starting with the essentially reductive premise that the material nature of entities is at the root of actions in the world, the Doctrine outlines a hierarchy of relationships, culminating in the theory of how teleological – which is to say, goal directed – behaviour stands as the "final cause" that explains the regularity with which functional effects are produced in the natural world. Aristotle's four causes can be enumerated:

1. Material Cause - the behaviour indicated by the physical properties of matter
2. Efficient Cause - the consequences of the manipulation of physical material
3. Functional Cause - the reasoning regarding efficient causes that informs actions on materials
4. Final Cause - the goal that motivates functional planning

Conventional approaches to creativity generally descend Aristotle's causal ladder: there is a goal, a plan for achieving this goal, a set of actions carried out to realise that plan, and a world of physical relationships in which those actions have consequences. Indeed, a fundamental principle of a certain approach to aesthetics is that the perception of beauty involves the recognition of a function that defines an artefact and an appreciation of the creative process employed in the achievement of that functionality [46]. An alternative theory, rooted in the philosophy of Kant, considers aesthetic experience to unfold in a perceptual domain of its own, involving a detachment from any practical consideration of an object of beauty [20]. Even in this latter case, though, beauty, from the perspective of a creator, becomes an objective unto itself, with the elicitation of an aesthetic response in principle indicating the achievement of this goal. So, regardless of the theoretical grounding adopted by an analyst, creativity seems to be bound up in an end directed process.

Computational creativity has tended to adopt a similar line. Ritchie has characterised the creative behaviour of an information processing machine in terms of the identification of a class of existing artefacts that qualifies as a target domain, subsequent generation of artefacts that are expected to fall within this domain, and then evaluation on the part of the system of whether the creative goal has been achieved [32]. Output produced without some sort of goal criteria has been described as "mere generation", a ramble through a state space that, regardless of its consequences, cannot be properly considered as creativity [11]. This lines up well with Boden's description of levels of computational creativity, with high level transformations of state spaces trumping lower level recombinations of elements within a predefined space [10]. It is in this transformational degree of symbol manipulation, involving the delineation of a state space above a traversal of a known space, where the complexity of the goal directed aspect of creativity becomes evident. A fundamental challenge for a computer scientist interested in designing autonomously creative systems is therefore to understand what it would mean for computers to make decisions about the definition of their own search spaces.

But it is not even clear how teleological processes arise in the material world, reducible, as it is, to the interactions of physical fields on a very small scale. Deacon has taken Aristotle's Doctrine as a starting point for his own exploration of the emergence of goal oriented behaviour in material reality, beginning with the premise that modern philosophy has sometimes tended to use dualism and homunculi to obscure the hard question of final cause [13]. For Deacon the first step up from the tumult of pure physics is a consideration of thermodynamic processes, by which a tendency that is so reliable it has a nomic aspect emerges from the random interaction of particles. In fact, despite the regularity implicit in the terminology "laws of thermodynamics", there is no principle that requires systems to move towards entropic arrangements; it is just the overwhelmingly likely outcome of a stochastic process. The kernel of teleology might be discovered in the apparent lawfulness of entropy that arises in systems that are actually just complex and unpredictable.

Like Deacon, Kauffman recognises the seeds of emergence in the way that order can spontaneously come about in a dynamic system, giving rise to interpretable attractors [19]. The contemporary case for emergence maintains that nested hierarchies of interacting attractors can be extrapolated into apparently teleological behaviour. At the higher end of the scale exist cases like evolution by natural selection, which, while it has been grasped through an astounding act of reductionist interpretation, can nonetheless only really be understood as a process directed towards the goal of fitness—and in fact it has been argued that evolution itself should be treated as a creative process. To put it simply, an evolved organism is a confluence of functions that result in their own perpetuation. Taking an example offered by Millikan, the biological operation of an animal can only be understood in terms of the functional role that the creature's various organs play in sustaining life, and these functions have been determined through an assiduous process of evolutionary trial and error: a lion's heart exists in order to pump blood through the lion's body, even though the genetic and developmental process that resulted in the existence of the organ cannot have been somehow aware of that outcome [24].

But, Millikan asks, what happens if a fully developed lion comes into existence spontaneously? While the lion might be considered an operational organism, it is tempting to conclude that its organic components have no function in the sense of having been selected because of a goal they accomplish. An evolved lion inherits properties of goal directedness from the generational history of organisms that has contributed to its fitness. This extension of the lion's emergent identity into the past corresponds to a converse projection of the functional properties of its components towards the accomplishment of future goals, specifically the goals of the lion surviving and reproducing. The spontaneous lion, on the other hand, while it also has some hope of coincidentally surviving and replicating, has simply happened: it cannot be interpreted as the fulfillment of a goal that has emerged in the unfolding of events in a complex and unpredictable environment. In terms of Aristotle's efficient cause, the lions are identical, but in terms of final cause they seem to be completely different.

Bickhard has responded to Millikan's case for a connection between causal history and functionality, however, by arguing that the history of a system cannot be a part of its ongoing operation [9]—history is, presumably, a contextualised interpretation of a present situation. Instead, Bickhard proposes, function should be understood in terms of the contribution a functional component makes to its system's persistence in a state that defies the entropic tendencies of the universe [8]. It is the case that the dynamics of complex and chaotic systems result in the emergence of processes that, in their regularity, seem to have a sense of following some kind of rule. This shift away from the basic laws of physics begins with processes such as thermodynamics, where the regularity lies precisely in the predictable breakdown of structure in systems, and moves out towards the further from equilibrium states that characterise the process of evolution, or more explicitly cognitive apparati such as representational symbols.

So by the emergentist account, causation is understood in terms of nested layers of dynamically coupled, intricately entangled processes, with each emerging attractor becoming an element in a higher level of interactions. This view escapes the paradoxes of trying to incorporate some representation of the system's past into its current operation, and at the same time seeks to explain the evident gravity of future outcomes in the workings of higher order complexes. The upshot of this is that teleological processes are necessarily associated with systems that are highly non-linear on several levels, an insight that sits well with the enactivist world view of Varela, Thompson and Rosch, who suggest that a mindful agent – which is to say, one capable of the planning and execution inherent in creativity – must be situated in a deeply interactive relationship with a dynamic and unpredictable environment [42].

There is a gravely concerning ramification to this conclusion from the perspective of a computer scientist interested in designing autonomously creative agents, however: if teleological processes only emerge in the context of complex interaction with a chaotic environment, it is difficult to imagine how a symbol manipulating machine could hope to creatively flourish in its rule based domain. Considering that even computational processes modelled nondeterministically can be reduced to linear operations, the case for a strictly algorithmic system producing output that would be judged even basically creative seems doomed. Two possibilities immediately present themselves as the beginning of a solution to this challenge: the modelling of dynamic interactions between rule following agents, and the physical construction of environmentally situated robots.

3 Robots

The classic intelligent agent concept [34] entails that an agent should be able to use actuators to manipulate its environment, which it monitors via perception. The agent has goals which it is trying to satisfy via its actions. Whether these goals have been reached is subject to an evaluation which the agent achieves by applying a metric. In classic AI, the agent's environment is understood much less literally than it is in robotics. Robots exist in a physical world that they actively manipulate and that directly affects their actions. Also, they share this world with humans. What follows offers a brief survey of contemporary approaches to robotics as they relate to creativity, followed by some thoughts on the future exploration of robots as creative agents.

From the perspective of the description of creative systems, the great appeal of robots lies in their situation in the same highly nonlinear environment from which human creators have emerged. As a first approximation, robots might be considered to have goals that are handed to them by a designer, grounded in external observations: in this case, the robot becomes an expression of its own creator's stance towards the world, and even in this basic instance a dynamic emerges where the robot's behaviour can become an element in a larger creative system, with the designer responding to the robot's successes and failures through subsequent design decisions. In what follows, this scenario will be case in terms of robots as a form of creative expression. More complexly, robots might be modelled as adaptive agents involved in a feedback loop with their own environments. In this case, while there may be overarching goals handed to a robot by a designer, it is the behaviour that emerges in the pursuit of this goal that may be considered creative. Ultimately, it is conceivable that robots or perhaps even more compellingly networks of robots might be involved in processes with unpredictable outcomes that can be interpreted as the emergence of truly goal oriented causation.

It has recently been argued [27] that real progress in natural language processing will depend on a more human-like machine which has a situated knowledge embodied in its own physical form. This presence [26] is necessary for a cognitive architecture which is more human-like and therefore capable of a human-like command of natural language. This may just as well be just as true for other cognitive abilities.

Feldman [16] sees two possible ways in which a robot can fully understand human subjective experience. One is a full simulation of the human body to gain insights into human experience. The other would be a new type of grounding that builds up an understanding of the world through the robot's own sensors and bodily experience.

Creative automata and machines which exist in the physical world have been built for centuries. There is, for instance, the case of von Kempelen's speaking machine [43, 15], which was a hybrid between a research project on the human vocal apparatus and an entertainment tool similar to a musical instrument.

Creativity in the domain of robotics can be conceptualised in terms of creative activities that are performed by intelligent agents capable of performing a full action-perception loop which takes the environment into consideration. Within this action-perception loop the, agent must have some "creative goal".

3.1 Agents and Embodiments

In order to understand what it means for a robot to be creative we will now describe a few systems which do in some way fulfill the criterion of being "deemed creative" if they were "performed by a human" [44]. Creative robots come in two flavours currently: they are either presented as being creative themselves or they are used as tools for expressing a human's creativity. We will fist deal with the later kind of robot for creative tasks.

Robots as a form of creative expression are teleoperated, which is to say their actions are determined by the perceptions and decisions of a human performer. Ogawa et al. [30] report on a teleoperated robot called the "Geminoid" [36] being used for a task in which the android and an actor performed a play live on stage together. This robotic agent had the following properties:

- The android takes the shape of a physical body which is modelled on an actual female human. The body has 12 degrees of freedom (DoF). These are mainly used for its facial expressions which closely copy the operator's facial movements. It also has loudspeakers which transmit the operator's voice to the audience.
- Perception is accomplished through a camera system which lets the operator see the machine's view of its environment.
- The machine's processing of the environment is realised by feeding the video back to the operator, and its actions are hence based on receiving "commands from the human operator".

So the robot's body itself is used for artistic expression. The authors conclude, based on experiment, that the robot actually improved the audience's sense of immersion in the performance. This is a surprising result but shows that the embodiment through the artificial agent can actually generate a different level of "meaning", as the authors suggest. It is actually the human-like but not-human body that generates this added meaning.

Robots as creative agents are autonomous to a certain extent. Tresset and Deussen [40] report on a robot, named e-David, which creates visual art through painting on a canvas. This agent had the following properties:

- e-David is not anthropomorphic (human-like). It is an industrial robot that only consists of an arm. The arm is also its actuator, with which it manipulates a pencil or brush.
- The perceptive apparatus is a camera system.
- The system performs the action-perception loop by creating an image it intends to paint and then monitoring its output by perceiving the painting as it emerges through its own actions applied to a canvas.

Embodiment is crucial in the case of e-David. The authors list thirteen ways in which e-David's embodiment has a direct impact on the final result of the visual art it produces. These include the velocity at which the arm moves, the pressure it applies to the painting, and control of the amount of paint on the brush. All of the factors have a direct effect on the visual appeal of the product which e-David produces. Thus, this robot demonstrates the importance of considering the physical presence of an artificial creative system in the creation of visual art.

Both the Geminoid and e-David illustrate how important the actual physicality of an intelligent agent it is and how their individual embodiments shape their creative output. However, the processing system in each case is actually quite different. Whereas e-David is autonomous in its actions to a large extent, the Geminoid is operated by a human. Thus, these two specific robots have different levels of autonomy and one needs to debate what "responsibilities", in the sense of Colton and Wiggins [11], they take on within the creative process.

3.2 Goals

As already illustrated, robotic agents that use their physical appearance and structure to pursue creative objectives can differ in their goals. Whereas the Geminoid in the study discussed above tries to evoke emotional response in an audience, e-David monitors its own output on a canvas via a visual feedback system.

Similarly, musical robots have goals which they pursue. In this case, the environment is typically the musical instrument with which the robots interact physically.

A robot coordinating its own body in a creative process

Batula and Kim [6] present a system which plays the score of a two-finger piece on a piano. The robot in this case is a small humanoid. Its environment is a keyboard. Its perception relates to the monitoring of its own motions and audio-feedback.

The robot's goal is to play the piece it has been assigned correctly. The authors frame their research as an investigation into the motorics required for musicianship. The robot's goals are simple: it detects mistakes in its own playing. This is very much in line with our argumentation. The system's physicality comes from the control of its own limbs in relation to the velocity of its playing. The robot controls its own motion, and the decisions of how to play rely solely on its own bodily control.

A robot coordinating with another body in a creative process

A contrasting approach is presented by Mizumoto et al [25]. In their approach the focus is on ensemble performance. The goal is for the agent to ask: "Am I creating the same output as another agent?"

This is a different question because the machine is no longer in control of the speed at which the product is created. The robot plays a theremin while the human plays drums. The robot's perception is used to actually calculate the action of the actuators, in contrast to the actuators acting independently to exert force on the physical environment. The required processing relies on a coupled-oscillator model.

3.3 Environments

What kind of environments do robots encounter in the course of creative processes? The comedic robot is one recent concept which has been implemented. Thus far, these robots are the only agents which actually treat an audience as their environment. They do exactly what an intelligent agent does by monitoring what effects their actions have on the environment.

Audience Monitoring

Other agents with which robots interact may be artificial (see section 4) or human audiences. Knight [21] analyses the impact of embodiment on performances in robot theatre. Knight et al [22] present a system which tells jokes to an audience.

In the system described a small humanoid robot is the comedic agent. Its goal is to make the audience laugh. It monitors levels of audience interest and attention (more precise methods are further described in [23]). The robot presents jokes and will choose the sequence of jokes in accordance with the audience's reaction. This is a direct application of the action-perception loop. The quality of the creative output is measurable in the sense that the audience reaction is the operationalisation of what the output should achieve.

Interacting with the Audience

Katevas at al [18] also use a humanoid robot as a stand-up comedian. In their performance, however, the robot actively engages with the audience and directly addresses individual members of the audience. In this way, the robot influences the outcome of the creative process. The goal is an active audience reaction, so the robot tries to improve the outcome and generate more laughs by engaging the audience.

As such the robot is not only relying on its output in the form of jokes, but also actively and preemptively shapes the audience's reaction and hence its environment's reaction to the jokes. This can be considered a different approach. If joke telling is considered an artistic and creative process, then the audience's reaction is the measure by which one can tell whether the result of the activity is of good quality. The robot here imitates the practices of human stand-up comedians by actively inducing a reaction in the audience. It does not just rely on the humorous value of the verbal stimuli it presents to the audience.

3.4 Creative Robots

In line with the theoretical points raised above, entertainment robotics is a growing market [7, 33]. The potential here is vast. A robot can use the principles outlined above to become an active companion [12, 3], giving itself an advantage over static media such as television broadcasting or film.

The three principles addressed here, embodiment, goals, and environments, will play a crucial goal in developing systems that can be deemed creative. This section has illustrated differing approaches to all three of these topics. In designing creative robotic systems, the human designer will have to think carefully about how the agent will pursue its goals within the given environment.

In line with the argument in this paper, for a robot to be truly creative it must be able to show adaptive behaviour. Embodiment will obviously be given from the outset in a robotic system, influencing the system's actions, perception, and interaction with the environment in a non-trivial way. However, real adaptivity for creativity will arise only if the robotic agent is able to define its own goals. An approach to robotics which includes this kind of behavioural autonomy is evolutionary robotics [29]. This approach assumes that the agent has some kind of overall goal such as playing a musical piece or amusing an audience via comedic practices. The sub-goals upon which the system operates would have to be adaptable. One way of implementing such a strategy would be to devise methods that allow the robot to choose between the goals outlined above (see section 3.2), or, with respect to interacting with the environment, choosing between the two strategies of, for instance, interaction with an audience as described above (see section 3.3).

4 Multi-Agent Systems

It is sometimes easy to forget that artists are not totally isolated from their environment: they come into contact with other artists who are tackling problems and trying to reach goals very similar to their own. Artists, scientists, chess players, normal people trying to make ends meet—creative people are influenced by other people, and they themselves influence other people, very often people with whom they are in no direct contact. Think of the generations of musicians influenced by Beethoven or of mathematicians working on problems formulated by Gauss.

In fact, one would be justified in thinking that creative processes are never the work of one individual alone, no matter how visionary and illuminating her thinking might be: every creator stands on the shoulders of giants. The intention here is to discuss how this interaction might be modelled through artificial agents, and how such an interaction might influence the behaviour of the agents towards, ultimately, determining the goal of the creative process itself.

As they relate to the imperative of creative goals as behavioural causes, the appeal of multi-agent systems is their potential for producing emergent attractors which cannot be understood as components of any single agent's behaviour. Agents themselves may be goal oriented – indeed, their processes are typically modelled in terms of the satisfaction of very basic criteria – but these goals are simplistic, whereas the operation of the overall system is nuanced. The power of simple agents collaborating to develop and realise complex goals can be observed in various real-world contexts, from the swarm behaviour of certain insects to the efficacious productivity of financial markets and indeed the homeostatic condition of entire ecological systems. This paper considers the question of how computers might be used to model multi-agent systems and then to analyse the potential for considering these systems as generators and executors of creative objectives.

4.1 Interacting agents

Multi-agent systems have been used extensively to model the origin and evolution of an impressive array of different social constructs [14, 31, 35], from ant or termite colonies to computer networks to economic markets. Agents are assigned a more or less rigorous set of beliefs, desires and intentions which determines their interaction. Agents are goal-oriented: their actions are determined by a desire to maximise a reward function, and it is through their interaction that the system evolves.

Most interestingly, multi-agent systems can show emergent properties: interaction between the agents allows the self-organisation of system properties that were not originally part of the system. Self-organisation, i.e. the lack of a centralised element imposing structure on the emergent property, is an important characteristic of such systems, revealing how organised properties can arise from simple interactions alone. These kinds of systems have been used, for instance, to model the self-assembling of biological complex structures[28], or to model the origin and evolution of language [37, 5, 4, 38]. In Steels' work, agents create and agree on a lexicon to name a series of objects in their environment. Their interaction follows a protocol specified in a "language game", similar to the language games described by Wittgenstein [45]. Van Trijp [41] shows that the "Naming Game" will converge towards a stable lexicon if certain requirements are met.

According to Tomasello [39]:

> The current hypothesis is that it is only within the context of collaborative activities in which participants share intention and attention, coordinated by natural forms of gestural communication, that arbitrary linguistic conventions could have come into existence evolutionarily...

This hypothesis seems to validate the modelling approach. An effort to understand creative processes as an attempt at collaborative behaviour by intelligent agents might prove to be very fruitful.

4.2 Creative processes as collaboration: a thought experiment

We propose a thought experiment which could help to illuminate the relationship between goal-seeking behaviour and creativity. Agents of different physical or cognitive characteristics are placed in an environment and forced to collaborate in order to achieve a series of tasks. To simplify things, we propose the following interaction rules:

1. All interactions are one-to-one: two agents are chosen and made to interact.
2. Agents are chosen at random: the system does not show a topology, i.e. it is a mean-field system.

3. One of the agents adopts the role of the demonstrator; the other is the observer.

Both agents have a clear idea of the task that is to be carried out. However, their different physical and cognitive skills require them to adapt their own actions to the task: some agents are better equipped to carry out the task in one way, whereas others must find efficient ways to carry out the task. During the interaction, the demonstrator performs the task in the most efficient way it can. Obviously, this way depends on all the previous experience of the agent. More particularly, it depends on what it has learned from all its previous interactions with other agents.

Following this demonstration, the observer must decide whether it is fit to perform the action in the same way. It does this by attempting to imitate the demonstrator. If it cannot, it must try to find a way to perform the action in a way that will resemble the demonstrator's actions, only adapted to its own abilities. If it succeeds in carrying out the action, then the observer will include this action into the set of actions it is capable of carrying out to perform the assigned task; the game is successful and two new agents are chosen to play new game. If, on the other hand, after a fixed number of attempts the agent is incapable of performing the action in a satisfactory way, then the game starts again, only now a new task is chosen: the goal changes. The new task should be similar to the previous one, if possible, so that agents might be able to identify properties of the task that are difficult for them, and perhaps learn to avoid them or find a way around them.

The hypothesis offered here is that such a system would become stable, i.e. it would reach a point after which all interactions would be successful. At this point all agents would have learned how to behave when forced to carry out a collaborative task. Every agent would have learned to adapt its own goal, according to its capabilities, to fulfill the task in a cooperative manner. Every agent would have learned how to work around what it cannot do.

5 Conclusion

This paper has examined the idea that creativity can be understood in terms of a process of adaptation on the part of agents attempting to accomplish a set of goals in complex and unpredictable environments. The hypothesis presented here is that agents dynamically coupled with their environments might become involved in the instigation of higher level emergent features that can be interpreted as potentially surprising and valuable new goal directed behaviours. There is scope for hoping that a network of multiple environmentally situated agents, each independently working towards their own micro-goals, will remit a systemic shift that in turn can become a target for discovery of new possible goals available to agents. From an external perspective, such a system offers the overall impression of being directed towards goals that are not in any way present in the programming that defines the behaviours of its components. In the physical universe, definable as it is in terms of a few simple rules of interaction, has nonetheless become a cauldron for such complex emergent systems evolution and cognition. In the same sense, a system of simple, interactive, environmentally oriented computational agents might have a chance of developing patterns of behaviour that can collectively be considered creative.

Existing work in the pertinent fields of robotics and multi-agent systems has been briefly discussed. The embodied situation of robots invites a consideration of the development of goal directed behvaiour in an unpredictable environment. And the dynamics of multi-agent systems present a platform for investigating the possibility of treating the attractors that emerge unexpectedly in the course of interaction as unanticipated creative objectives. The juxtaposition of these two topics in the context of computational creative naturally suggests an amalgamation: a potential project developing swarms of individually adaptive robots, treating their own community of robotic co-agents as an environment embedded in the physical world, with each robot adapting its behaviour based on its interaction with its peers. On an individual level, the robots would update their procedures based on observations of other robots and with the pre-programmed objective of accomplishing simple goals. On a collective level, the robotic system as a whole might very well take on an emergent aspect, with unexpected intimations of higher level organisation. The question raised by such a model is whether the system's proclivity for organising itself in a surprising and potentially effective way can be considered the creative discovery of a new objective.

ACKNOWLEDGEMENTS

Griffiths is supported by ConCreTe: the project ConCreTe acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET grant number 611733. McGregor is supported by EPSRC grant EP/L50483X/1.

REFERENCES

[1] Mohammad Majid al Rifaie, John Mark Bishop, and Suzanne Caines, 'Creativity and Autonomy in Swarm Intelligence Systems', *Cognitive Computation*, **4**(3), 320–331, (2012).

[2] Aristotle, *The Works of Aristotle Volume II: Physica*, The Clarendon Press, Oxford, 1930. Translated by R. P. Hardie and R. K. Gaye.

[3] Ruth S Aylett, Ginevra Castellano, Bogdan Raducanu, Ana Paiva, and Mark Hanheide, 'Long-term socially perceptive and interactive robot companions: challenges and future perspectives', in *Proceedings of the 13th international conference on multimodal interfaces*, pp. 323–326. ACM, (2011).

[4] A. Baronchelli, V. Loreto, L. Dall'Asta, and A. Barrat, 'Bootstrapping communication in language games: Strategy, topology and all that', in *The Evolution of Language, Proceedings of the 6th International Conference (EVOLANG6), edited by A. Cangelosi, A. D. M. Smith & K. Smith, World Scientific Publishing Company*, (2006).

[5] J Batali, 'Computational simulations of the emergence of grammar', in *Approaches to the Evolution of Language: Social and Cognitive Bases*, eds., J R Hurford, M Studdert-Kennedy, and Knight C., 405–426, Cambridge University Press, Cambridge, (1998).

[6] Alyssa M Batula and Youngmoo E Kim, 'Development of a mini-humanoid pianist', in *10th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2010)*, pp. 192–197. IEEE, (2010).

[7] George A Bekey, *Autonomous Robots*, MIT Press, Cambridge, MA, 2005.

[8] Mark Bickhard, 'The Emergence of Contentful Experience', in *What Should Be Computed to Understand and Model Brain Function?*, ed., Tadashi Kitamura, World Scientific, (2001).

[9] Mark Bickhard, 'The Dynamic Emergence of Representation', in *Representation in Mind, Volume 1: New Approaches to Mental Representation (Perspectives on Cognitive Science)*, eds., Hugh Clapin, Phillip Staines, and Peter Slezak, Elsevier, (2004).

[10] Margaret A. Boden, *The Creative Mind: Myths and Mechanisms*, Weidenfeld and Nicolson, London, 1990.

[11] Simon Colton and Geraint Anthony Wiggins, 'Computational creativity: The final frontier?', in *Proceedings of 20th European Conference on Artificial Intelligence (ECAI)*, eds., L. De Raedt, C. Bessiere, D. Dubois, P. Doherty, P. Frasconi, F. Heintz, and P. Lucas, pp. 21–26, Montpellier, France, (2012). IOS Press.

[12] Kerstin Dautenhahn, Sarah Woods, Christina Kaouri, Michael L Walters, Kheng Lee Koay, and Iain Werry, 'What is a robot companion-friend, assistant or butler?', in *Intelligent Robots and Systems,*

2005.(IROS 2005). 2005 IEEE/RSJ International Conference on, pp. 1192–1197. IEEE, (2005).

[13] Terrence W Deacon, *Incomplete nature: How mind emerged from matter*, WW Norton & Company, New York, NY, 2011.

[14] Mark D'Inverno and Michael Luck, *Understanding Agent Systems*, Springer Verlag, 2nd edn., 2004.

[15] Homer Dudley and Thomas H Tarnoczy, 'The speaking machine of Wolfgang von Kempelen', *The Journal of the Acoustical Society of America*, **22**(2), 151–166, (1950).

[16] Jerome A Feldman, *From Molecule to Metaphor: A Neural Theory of Language*, MIT Press, Cambridge, MA, 2006.

[17] Petra Gemeinboeck and Rob Saunders, 'Creative Machine Performance: Computational Creativity and Robotic Art', in *Proceedings of the Fourth International Conference on Computational Creativity*, (2013).

[18] Kleomenis Katevas, Patrick G.T. Healey, and Matthew Tobias Harris, 'Robot stand-up: Engineering a comic performance', in *Proceedings of the 2014 Workshop on Humanoid Robots and Creativity at the 2014 IEEE-RAS International Conference on Humanoid Robots (Humanoids 2014)*, Madird, Spain, (2014). Available: $http://cogsci.eecs.qmul.ac.uk/humanoids/Katevasetal\_2014.pdf$.

[19] Stuart Kauffman, *At Home in the Universe: The Search for Laws of Complexity*, Oxford University Press, 1995.

[20] Gary Kemp, 'The Aesthetic Attitude', *British Journal of Aesthetics*, **39**(4), 392–399, (1999).

[21] Heather Knight, 'Eight lessons learned about non-verbal interactions through robot theater', in *Social Robotics*, eds., B. Mutlu, C. Bartneck, J. Ham, V. Evers, and T. Kanda, 42–51, Springer, (2011).

[22] Heather Knight, Scott Satkin, Varun Ramakrishna, and Santosh Divvala, 'A savvy robot standup comic: Online learning through audience tracking', in *International Conference on Tangible and Embedded Interaction*, Funchal, Portugal, (2010).

[23] Heather Knight and Reid Simmons, 'Estimating human interest and attention via gaze analysis', in *2013 IEEE International Conference on-Robotics and Automation (ICRA)*, pp. 4350–4355. IEEE, (2013).

[24] Ruth Millikan, *Language, Thought, and Other Bilogical Categories*, MIT Press, Cambridge, MA, 1984.

[25] Takeshi Mizumoto, Takuma Otsuka, Kazuhiro Nakadai, Toru Takahashi, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G Okuno, 'Human-robot ensemble between robot thereminist and human percussionist using coupled oscillator model', in *, 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1957–1963. IEEE, (2010).

[26] Roger K Moore, 'Presence: A human-inspired architecture for speech-based human-machine interaction', *Computers, IEEE Transactions on*, **56**(9), 1176–1188, (2007).

[27] Roger K Moore, 'From talking and listening robots to intelligent communicative machines', in *Robots that Talk and Listen – Technology and Social Impact*, 317 – 336, De Gruyter, Boston, MA, (2014).

[28] Radhika Nagpal, Attila Kondacs, and Catherine Chang. Programming methodology for biologically-inspired self-assembling systems, 2002.

[29] Stefano Nolfi and Dario Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-organizing Machines*, Cambridge, MA, 2000.

[30] Kohei Ogawa, Koichi Taura, and Hiroshi Ishiguro, 'Possibilities of androids as poetry-reciting agent', in *RO-MAN, 2012 IEEE*, pp. 565–570. IEEE, (2012).

[31] Liviu Panait and Sean Luke, 'Cooperative Multi-Agent Learning: The State of the Art', *Autonomous Agents and Multi-Agent Systems*, **11**(3), 387–434, (2005).

[32] Graeme Ritchie, 'Some Empirical Criteria for Attributing Creativity to a Computer Program', *Minds and Machines*, **17**, 67–99, (2007).

[33] Florian Röhrbein, Sascha Griffiths, and Laura Voss, 'On Industry-Academia Collaborations in Robotics', Technical Report TUM-I1338, Technische Universität München, (2013).

[34] Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall International, Harlow, third edn., 2013.

[35] Yoav Shoham and Kevin Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press, 2008.

[36] S.Nishio, H.Ishiguro, and N.Hagita, *Geminoid:Teleoperated android of an existing person*, chapter 20, 343–352, I-Tech, Vienna, Austria, 2007.

[37] L. Steels, 'A Self-Organizing Spatial Vocabulary', *Artificial Life*, **2**(3), 319–332, (January 1995).

[38] L. Steels, 'Self-organization and Selection in Cultural Language Evolution', in *Experiments in Cultural Language Evolution*, ed., L Steels, John Benjamins, Amsterdam, (2012).

[39] M. Tomasello, *Origins of Human Communication*, MIT Press, Cambridge, MA, 2008.

[40] Patrick Tresset and Oliver Deussen, 'Artistically skilled embodied agents', in *Proceedings of AISB2014*, Goldsmiths, University of London, (1st - 4th April 2014).

[41] R van Trijp, 'The Evolution of Case Systems for Marking Event Structure', in *Experiments in Cultural Language Evolution*, ed., L Steels, 169–205, John Benjamins, Amsterdam, (2012).

[42] Francisco J. Varela, Evan Thompson, and Eleanor Rosch, *The Embodied Mind*, MIT Press, Cambridge, MA, 1991.

[43] Wolfgang von Kempelen, *Über den Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*, J.V. Degen, Vienna, 1791. Facsimile re-print from 1970, Stuttgart: Frommann-Holzboog.

[44] Geraint A Wiggins, 'A preliminary framework for description, analysis and comparison of creative systems', *Knowledge-Based Systems*, **19**(7), 449–458, (2006).

[45] L. Wittgenstein, *Philosophische Untersuchungen*, Joachim Schulte Wissenschaftliche Buchgesellschaft, Frankfurt, 2001.

[46] Nick Zangwill, 'Aesthetic Functionalism', in *Aesthetic Concepts: Sibley and After*, eds., Emily Brand and Jerrold Levinson, Oxford University Press, Berlin, (2001).