# Of (zombie) mice and animats

Slawomir J Nasuto and J Mark Bishop

**Abstract** The Chinese Room Argument purports to show that‘ syntax is not sufficient for semantics’; an argument which led John Searle to conclude that ‘programs are not minds’ and hence that no computational device can ever exhibit true understanding. Yet, although this controversial argument has received a series of criticisms, it has withstood all attempts at decisive rebuttal so far. One of the classical responses to CRA has been based on equipping a purely computational device with a physical robot body. This response, although partially addressed in one of Searle’s original contra arguments - the ‘robot reply’ - more recently gained friction with the development of embodiment and enactivism[1], two novel approaches to cognitive science that have been exciting roboticists and philosophers alike. Furthermore, recent technological advances - blending biological beings with computational systems - have started to be developed which superficially suggest that mind may be instantiated in computing devices after all. This paper will argue that (a) embodiment alone does not provide any leverage for cognitive robotics wrt the CRA, when based on a weak form of embodiment and that (b) unless they take the body into account seriously, hybrid bio-computer devices will also share the fate of their disembodied or robotic predecessors in failing to escape from Searle’s Chinese room.

---

J. Mark Bishop
Goldsmiths, University of London, UK, e-mail: `bish@gold.ac.uk`

Slawomir J. Nasuto
University of Reading, Reading, UK e-mail: `s.j.nasuto@reading.ac.uk`

[1] In this work the term enactivism will be used to delineate theoretical approaches to cognition that emphasise perception as action encompassing, for example, Gibson’s ‘ecological approach’; Varela et al’s ‘embodied mind’; Nöe’s ‘action as perception’ and O’Regan and Nöe’s ‘sensorimotor account of vision’.

## 1 Introduction

In his 1980 paper *Minds, Brains and Programs* (MBP)[46] John Searle formulated his influential *Chinese Room Argument* (CRA) aimed at refuting the possibility of achieving the holy grail of Artificial Intelligence[2], what he termed 'Strong-AI': that is, creating a truly intelligent computational device; instantiating mind in machine.

In spite of the controversy it generated, CRA remains a hallmark argument in the debate over the possibility of instantiating mind in computing devices. In its most basic form, it addresses the most radical version of the claim as proposed by good old fashioned Artificial Intelligence (GOFAI)[3]. Nonetheless many scholars do not agree that the CRA succeeds or at least try to suggest frameworks which could circumvent its conclusions. One such area purported to escape the CRA argument is 'cognitive robotics'. The hope of its proponents is that by providing a physical body, computational operations are married to cognitive processes via embodiment and enactivism, and by virtue of the latter the CRA argument fails to apply.

This paper will briefly introduce the original argument and will argue that in its current form, cognitive robotics is more aligned with a particular form of enactivism (weak enactivism) which does not seem to offer a way out of Chinese Room.

Furthermore, there has been a nascent field of hybrid systems which blend artificial and biological systems. The question can then be extended to such hybrids: some forms of which perhaps might circumvent the CRA.

The paper will review such developments and will consider them from this perspective.

## 2 Chinese Room Argument

The CRA has been considered one of the most influential arguments in the history of philosophy of mind achieving at the same time a status of notoriety amongst the proponents of AI who aimed but failed to quash it with various counter-arguments[10][45].

In a thought experiment John Searle - who can only speak English - is locked in a room and communicates with external interlocutors via messages written on paper[4]. Searle has a rule-book with instructions in English for

---

[2] The Dartmouth Proposal, "Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it",[30].

[3] From Newel & Simon (1976), *'a physical symbol system has the necessary and sufficient means for 'general intelligent action' '.*

[4] NB. In this work we deploy an extended form of the CRA; in the original version interlocutors merely pose [Chinese] questions about a given story [also in Chinese], which Searle, using his rule-book, responds to appropriately.

manipulating strings of [Chinese] symbols received as input and subsequently formulating an output string of symbols, such that the characters appear to the interlocutors to be linguistic responses in Chinese; in this manner communication is achieved via appropriate exchange of Chinese ideographs.

Yet, in spite of being able to converse with the Chinese interlocutors in a way that for all purposes appear to them as if he can understand Chinese, Searle proposed that in fact he does not understand a word of Chinese, no matter how skilful his manipulations of the Chinese symbols is.

The CRA was intended to show that computers may one day become skilful enough to appear to process language in a meaningful way by using only syntactic manipulation, however by this process they remain incapable by themselves of giving rise to meaning or semantics.

Thus the Chinese Room Argument challenges functionalism and computational theory of mind. The latter proposes that mental states are simply computational states which are implementation-independent. As such, they can be instantiated in a computational device by mere symbol manipulation. Although John Searle did not dismiss the possibility that machines could possess intentionality and true understanding (indeed he specifically identified humans as such 'biological machines'), he did not believe these qualities could come about by sheer computational symbol manipulation alone.

## 2.1 Intentionality in computational systems?

A number of arguments have been put forward against the CRA, some of which had already been anticipated by Searle in the original paper[46]. These counter proposals can be categorised into groups purporting to refute CRA on different grounds. Various forms of systems replies try to argue that understanding is not a property of Searle alone, but of the entire system. What that system should be is the subject of particular variants of the systems reply.

Some variants of the System reply posit to give rise to true understanding the system must be effectively implementing a simulation of a brain (or at the very least, be imply,enter via some kind of connectionist architecture). Detailed taxonomies of different replies to the CRA together with rebuttals have been presented elsewhere[10, 45, 46]. Instead of providing yet another one here, we wish to focus on a specific kind of systems reply, the so called 'robot reply', which although considered in the original paper by Searle, has more recently gained particular momentum thanks to the links between cognitive robotics and a new move in cognitive science called enactivism[32, 41].

The robot reply proposes that true understanding must arise from grounding of meaning in the physical world and hence that the system must enable such grounding to take place. This is to be achieved by an appropriate rule-book enabling the robot to implement the 'right type of manipulations' and concomitant sensory motor coupling afforded by the robot's interactions with the external world. It is claimed that such an extended system (robot plus appropriate computational mechanism; the latter often proposed to be a connectionist architecture or brain simulation) can fulfil the necessary and sufficient conditions for meaning and understanding to arise.

However, as previously mentioned, the initial extensions of the basic CRA discussed by Searle in the original paper[46] explicitly addressed such a 'Robot reply' which Searle claims buys nothing, for the CRA could easily be extended by providing additional input in the form of symbolic values corresponding to camera and other sensory readings; the outputs strings Searle now produces encompassing both the robots verbal responses [in Chinese] and symbolic commands to manipulate (unbeknownst to Searle in CRA) the external objects by the robot's actuators. Such an extension would only require a more complicated rule-book; the extra syntactic inputs and different forms of response would continue to afford no real understanding if it were not there in the first place. In accord with Searle's response to the Robot reply, we similarly conclude that if we were to attribute genuine mental states/intentionality to such a computationally driven robotic device, we would also be obliged to do so also for any modern car equipped with electronic sensors and computer.

## 3 Robotic reply and enactivism

A very refreshing movement within cognitive science has gradually been emerging which rejects the computationalist view of cognition in favour of enactivism[53]. Enactivism emphasises the importance of embodiment and action in cognition and proposes that the most fundamental notion is that of embodied autonomy, which superficially at least offers renewed fundamental justification to cognitive robotics as a useful tool able to address the most fundamental questions about cognition and understanding.

Cognitive robotics itself could be viewed as a departure from the disembodied good old fashioned AI (GOFAI) as it also considers that embodiment is fundamental for cognition to arise. Moreover, various forms of cognitive robotic stress to a different degree the importance of embodiment for cognition, with some placing more emphasis on the actual body and its affordancies, than on the nitty gritty of the central 'computational' processing unit[43, 44]. In fact, this modern successor of GOFAI has been proposed to provide a fertile experimental ground for cognitive science[32]. Considered to be a radical departure from GOFAI by its enthusiasts, it has found itself in mutually beneficial symbiosis with some forms of enactivism[29, 41]. At first sight it thus seems that such an alliance may be able to provide a rebuttal to CRA on theoretical grounds.

### *3.1 Does cognitive robotics escape the CRA?*

In order to answer the above question it is important to emphasise that there are many interpretations of enactivism[5] and that cognitive robotics is particularly aligned with versions that emphasise the role of sensori-motor couplings[37]:

> our ability to perceive not only depends on, but is constituted by, our possession of ... sensorimotor knowledge

at the same time eschewing Varelian enactivism in which fundamental autonomy stems from the organisational closure of living systems[2]:

> it is somehow intuitive that cognition relates to sensorimotor interactions rather than to material self-constructing processes.

This form of enactivism embraces Gibbsonian affordances and moreover proposes that the experienced qualities[37]

> pattern[s] in the structure of sensorimotor contingenc[ies]

are sensori-motor laws[40, 41]. As Nöe put it[37]:

> for perceptual sensation to constitute experience, that is, for it to have genuine representational content, the perceiver must possess and make use of sensorimotor knowledge.

Although we agree that sensorimotor interactions are important for cognition, the move away from the organisational closure proposed by Barandiaran and Moreno[2]:

> ... as well as being somewhat awkward for cognitive robotics (since it would imply that no genuine cognitive behaviour can be expected from non-self-constructing artifacts) this thesis [that autopoiesis is necessary for cognition] is also conceptually uncomfortable

appears to us unjustified; for the above sentiment seems to be based on either expediency and handpicking the elements of enactive theory that suit a particular style of robotic approaches to cognition, or on confusion between organisational closure and autopoiesis. However, although the latter two notions seem intimately linked - with the notion of autopoiesis being the minimal organisation of the (unicellular) living systems - organisational closure is broader as it characterises further autonomous systems such as multi-cellular organisms as well as the nervous or even the social systems[18].

Nevertheless, in spite of the convenience of such an argument for current cognitive robotics, the concentration of the sensorimotor account solely on

---

[5] Our discussion specifically addresses the particular interpretation of sensory motor account derived from early works of Nöe and O'Regan on this subject, which seems to have been adopted within cognitive robotics community[29, 32, 43, 44]. It is important to note though that both authors have since developed their accounts in separate and increasingly divergent directions[38, 42].

the external world - with the reduction of the role of the body to mere instantiation of appropriate sensory motor couplings and disregard for the material self-constructing processes which also constitute the integral part of the body - does not seem to to us to afford any extra mileage over and above the original robot reply considered by Searle in MBP.

Conversely, we suggest that as long as efforts within cognitive robotics are directed only towards grounding meaning in the external world - whilst neglecting the need for concomitant grounding in internal states - all devices so constructed can ever hope to achieve are merely ever more sophisticated reflections of the relational structure of the external world in the relational structure of their internal [formal] representations, with fundamentally no account of either 'raw feel' or the genuine understanding of anything.

To illustrate this consider how Searle - merely deploying the CRA rule-book inside the Chinese room - could ever answer the following question (posed, of course, in Chinese), "Are you hungry?". We suggest that there is a fundamental difference between Searle's ability to answer questions of this form, with his ability to 'converse' about the relationships between objects *external* to the Chinese Room. In the latter case the rule-book, augmented by any of Searle's own contemporaneous notes[6], may enable him to identify symbol associations and appropriate manipulations without actually entailing any understanding on his part. In this sense he indeed would be acting (perhaps with the help of a pen and paper) as an expert system or a neural network - making associations between the symbols and the frequencies of their co-occurrences. A neural network can capture such associations between objects by tweaking its internal weights - albeit this is a mechanistic operation, itself devoid of meaning (i.e. ungrounded).

In fact, the above observation applies whether or not one considers 'the classical Chinese Room Argument' or the embodied (robot-reply) version as long as the embodiment is merely intended to provide sensory motor coupling in the sense of extra information about the possible manipulations various objects entail. This is why the Chinese Room Argument enables Searle to make reasonable responses as long as his queries are *exclusively* about the external world; the Chinese room can algorithmically capture such 'semantic webs', as this is essentially merely a statistical problem - computers already can do this.

Internalising the entire Chinese room[7] as in Searle's initial response to the systems reply to the CRA will not help either, as long as Searle is not allowed to interact with the external world directly (i.e. without the veil of the formal CRA rule-book) for, in this case, Searle would immediately start forming mappings between his own internal meanings and the new symbols

---

[6] Such 'notes' may define 'internal representations' that, for example, might highlight specific associations between strings of symbols.

[7] I.e. Searle memorises the rule-book, his notes etc. and subsequently performing all operations in memory such that there is nothing in the system that is not now inside Searle.

and their associations (this is exactly how we learn any foreign language). Meaning would therefore be transferred by association to the new symbols, which by themselves do not originally carry any meaning (to a non Chinese speaker).

Similarly, the question of whether a symbolic computational, sub-symbolic connectionist, or continuous dynamical system approach should be adopted translates into the question of formal richness of the internal relational universe or the mathematical nature of the mapping between external and internal relational spaces. Although there are very important considerations delineating some key properties of cognitive states, they pertain 'only' to necessary aspects of intentionality related to the nature of regularities in the external world (continuous and statistical or symbolic and recursive) and the best formal means to extract and manipulate them; they do not reference, and remain ungrounded in Searle's own internal bodily states. As various CRA variants elaborate, the precise nature of operations needed for the construction of internal representations (or means by which a mapping between external and internal relational structures is achieved) is irrelevant.

Some cognitive roboticists concede that current robotic platforms have been too impoverished in terms of their sensory surface to provide proper embodiment, but they insist that it is merely a matter of providing robots with more sensors in order to achieve genuine intentional states. However, adding more sensors (e.g. touch, proprioception) and actuators does not buy anything apart from larger rule-books, vectors to correlate or look-up tables.

The above considerations, important as they are, are clearly insufficient to fully ground intentional states as, for example, Searle in CRA would painfully become aware if the CRA experiment was ever actually conducted by cynical interrogators. The demonstration would be very simple, if cruel, as all that would be needed is to lock the door of the Chinese room and wait; soon enough, as the monoglot Searle remains unable to communicate his bodily needs to the outside world in Chinese, the CRA (or Searle to be precise) would be no more[8].

This is because the rule-book details purely formal associations between uninterpreted symbols. No amount of codifying associations and frequencies of co-occurrences between symbols relating to the external world will help Searle in the Chinese room communicate his internal states and desires, or to answer questions that inherently call for reference to the internal state of the 'system' (of which Searle is a part). E.g. questions such as :- 'do you believe this story to be true?', 'do you like this story?', 'how does this story make you feel?', etc. Any of the associations the rule-book could be permitted to codify (that Searle could try to use to answer such questions) will, *ex hypoth-*

---

[8] Searle, being unable to communicate his basic bodily requirements for food and water to his interrogators outside there room, would quickly die.

*esi*, relate only to external objects and hence will remain mere third person observations; none can ever detail appropriate first person associations[9].

Ironically, the inability of Searle (in the CRA system) to communicate his own internal states can be contrasted with his perfect ability, *ex hypothesi*, to communicate about the internal states of Chinese interlocutors; they are mere external states to him after all.

The strict sensorimotor account - and hence much of modern cognitive robotics - for all their claims of radical departures from computationalism/GOFAI, seem to invoke a parallel move to the implementation invariance of the latter approaches; this time a hardware implementation invariance, which in effect states that details of different embodiments do not matter as long as they afford the same sensory motor contingencies. The latter, though are assumed to amount to appropriate causal relationships between possible manipulations or actions (*how the sensation changes in response to object manipulation*) and sensations (*how the objects 'feels'*). However, because body can be memoryless, invoking hardware invariance principle, the sensory motor laws must amount to appropriate co-occurrences of activations of appropriate parts of the nervous system.

Although the sensory motor account seems intuitive and appealing in its emphasis of the fact that we understand by being in the world and acting upon it, nevertheless its account rests on some special role or properties that motor actions must have when leading to perception of their outcome. Why 'a pattern in the structure of sensorimotor contingencies'[37] is any different from patterns in sensory data? After all, both must result in (and only in) respective concomitant patterns of activity of neurons in appropriate brain structures.

If, rather than talking about sensory motor coupling we substitute another sense for acting - we also get co-occurrences and it is not easy to see why this would lead to fundamentally lesser (rather than simply different) understanding than sensory-motor coupling. At the end of the day whether it is sensory-motor or sensory-sensory coupling, both correspond to patterns of neural activations co-occurring in a coordinated manner in the brain and there is nothing in the sensory motor account that explains why co-occurrences between sensory-motor neural activities should assume such special role.[10]

The other possibility is that either hardware implementation invariance is violated (the body does count) or there is more to sensory motor laws

---

[9] Note that the original CRA argument is about Searle answering questions about a story; the questions we provide above are merely illustrations of the inherent limitation of CRA system that could be gleaned by more Searle-sympathetic interrogators.

[10] Interestingly, that the co-occurrences in the form of correlations (actually sensory-sensory correlations sic!) are indeed important is illustrated by the rubber hand illusion, in which subjects, when presented with a rubber hand in appropriate position in their field of view and observing how that hand is stroked simultaneously with their own (hidden) hand, report feel that the rubber hand is their own[5].

than action-sensation associations. Whichever is the case, both alternatives seem to point to the same conclusion, the extra “ingredients” present must be related to the biological makeup of the organism. At the most fundamental level, these will be bio-physico-chemical properties of the body (including the nervous system) induced by motor actions and sensory activations; metabolic properties of its constituents at all levels (as we can talk about metabolic needs of the entire organism, of its components - eg the brain but also about metabolic properties of individual components - the cells).

Consistently with Varelian forms of eanactivism[18, 38], true intentionality can only arise in systems which ground meaning jointly - respecting external constraints as well as internal states - a situation which, as the CRA illustrates, is impossible to achieve by a computational (or in fact any mechanistic/formal) system as they have no such physiological states at all.

What is closely related is that even though formal systems (even those instantiated in a robotic device) may be in principle rich enough to reflect the complexity of the relational structure of the external world, there is nothing in their constituent structures that will make them do so; or do anything at all for that matter. This is because of their very nature - abstraction of any mechanistic rule or formalism from any system that instantiates it. For example what the symbolic operations are should be invariant to the means by which they are accomplished. Thus, there is nothing that inherently compels an artificial agent to do anything, to perform any form of formal manipulation that could help it to map out the regularities of the external world. Turing-machine based robotic systems can at best, using Dennett’s phraseology, instantiate ‘as-if’ autonomy and teleology; in reality merely reflecting their designers wishes and goals.

In contrast, real cognitive agents have internal drives at all levels of organisation - survival, metabolic and physical - that make them act in the world, make them react to the external disturbances (information) and manipulate it in such a way that they will support immediate and delayed fulfilment of the drives at all levels. Such manipulation of information is effectively intentional as it is tantamount to the biological, biochemical and biophysical changes of real cognitive agents’ biological constituents, which are intrinsically grounded (they have metabolic, physiologic and ultimately survival values).

The intentionality comes not only from the potential mapping between the relational structures of the external world and the states of the biological constituents; but also appears as a result of external disturbances (which under such mapping correspond to information manipulation) which are also intrinsically grounded as they follow real physical laws and do not come about merely for the symbol manipulation’s sake. Systems which are based only on formal manipulation of the internal representations are thus neither intentional nor autonomous (as no manipulation is internally driven nor serves an intrinsically meaningful purpose other than that of system designer’s).

## 4 Modern embodiments

But the story does not end with robotic systems controlled by Turing-machines alone. In the recent years, huge strides have been made in advancing hybrid systems. These devices are robotic machines with both an active neurobiological and artificial (e.g. electronic, mechanical or robotic) components. Such devices start to blur the divide between the artificial and the biological. In particular, systems integrating artifacts with the nervous system may offer interesting avenues to explore new potential counter arguments to the CRA.

Indeed, there has been a long history of attempts to create interfaces between artefacts and the motor system, in the form of prostheses[34]. Interfaces with the sensory modalities include cochlear implants for improving hearing[4], as well as retinal implants, which recently have been shown to be capable, in principle, to enable reading to their users[17, 57].

Great strides made in implant technology advanced it beyond augmenting sensory modalities towards interfacing directly with the brain, with deep brain stimulation being one of the clinically approved treatments for some neurological disorders[16, 25]. Recent animal studies have successfully demonstrated possibility of creating implant replacing deep brain structure such as hippocampus for restoring existing memories[3].

In the above case the implant was trained on data recorded from the hippocampus of an animal previously trained on a spatial memory task. When subsequently the hippocampus was inactivated, the animal showed impairments on the same task, whereas the behavioural measures of task performance were restored, once the hippocampal input/output function has been replaced by the implant.

Other studies have demonstrated that implant devices could be used to lay down new associations, as was the case for classical conditioning of rats with synthetic cerebellum implants[31]. Rats with inactivated cerebella shown no ability to learn new classical conditioning responses, whereas in animals in which the input output functions of cerebella have been replaced by implants created to mimic them, the rats recovered ability to learn new classically conditioned responses.

Brain Machine Interfaces (BMIs) open new communication channels by allowing direct interface between the brain structures (typically cortex) and external devices, and may afford a seamless interface with prostheses[24, 27, 33, 36, 54]. Brain Computer Interfaces (BCIs) strive to achieve similar aims by less invasive means (typically using noninvasive EEG signals), thus extending the range of potential applications beyond the clinical realm[7, 12, 26].

Finally, animats - robotic embodiments of neural cultures grown in vitro - allow for investigation of the biological neuronal networks underlying sensory processing, motor control, and sensory motor loops[14, 28, 39, 55].

Given such considerable advances, it then becomes a very pertinent question to enquire whether some form of bio-machine hybrids could achieve what does not seem to be in the grasp of the conventional cognitive robotics. That is, whether a suitable combination of a computationally driven robotic device with a biological body can achieve a true understanding denied to its electromechanical, Turing-machine driven, cousins by the CRA.

In order to entertain such a possibility it is important though to delineate which of the types of systems outlined above might be good candidates for such consideration. It seems clear that such systems divide along the faultline defined by “who is in charge” - they can be either ‘sentient being’ - driven (these include prostheses, implants, BMI, BCI), or driven by ‘formal systems’. Extant animats fall into the latter category.

Of the hybrid advancements, the ones where the overall control of the system rests with the sentient agent are not really addressing the problem at hand. This is because any form of understanding claimed by the hybrid system would quite clearly be enabled via bootstrapping the sentient being’s ‘understanding’. Conversely the problem we wish to consider is whether a formal system with a form of biological embodiment that is not afforded by standard and recent cognitive robotics systems circumvents the CRA objections. It thus follows that out of the advancements overviewed above the animats provide a platform that is a serious contender for such a position.

## *4.1 Animats*

Recently, one of the co-authors, with a team from University of Reading, developed an autonomous robot (aka ‘animat’) controlled by cultures of living neural cells, which in turn are directly coupled to the robot’s actuators and sensory inputs[56]. Such devices come a step closer to the physical realisation of the well known ’brain in a vat’ thought experiment[11].

The ‘brain’ of the system consisted of a cultured network of thousands of neurons. The cultures are created by first removing any existing structure from cortical tissue of foetal rats and then seeding the resulting suspension containing neuron bodies on a plate and providing suitable nutrients. The plate has an array of 8x8 electrodes embedded at the base (a multi-electrode array (MEA)), which provide a bi-directional electrical interface to the cultures via appropriate hardware.

Within a short time after seeding, the neurons spontaneously begin to form new connections between each other and henceforth start engaging in communication. Given the right culture medium containing nutrients, growth hormones, and antibiotics, a culture tends to develop within a day into a monolayer with a dense network of connections, and within a week it starts to produce spontaneous activity in the form of single action potentials. The activity intensifies over the subsequent weeks developing into bursts of activ-

---

[11] For movie of an animat see `www.youtube.com/watch?v=1-0eZytv6Qk`

ity across the entire culture, which continue until culture maturation (ca 1 month since seeding).

Thus, MEAs allow for monitoring of an electrical activity of entire cultures as well as for their electrical stimulation via electrodes. This ability of bi-directional communication enabled creation of closed-loop systems between physical, and simulated, mobile robotic platforms and cultured networks. At Reading we used an off the shelf robotic platform (Miabot; Merlin Robotics, UK), because of its simplicity, accuracy of motor command encoding and speeds suitable for movement in an enclosed, custom built robot pen.

The created system was modular and consisted of several hardware and software modules including a robot (hardware or simulation), an MEA and its recording and stimulation hardware and software, a computer workstation for conducting on the fly machine learning analysis of recorded culture activity and extracting pertinent features of neural activity, another workstation for controlling the robot and delivering commands to robot actuators. The resulting signals from the robot ultrasonic sensors were translated into stimulation signals received by the culture and all the different modules were linked into an overall closed-loop system via a TCP/IP protocol.

Cultures used in our studies consisted of tens of thousands of neurons and showed complex, seemingly random pattern of connectivity and resulting activity. However, further study of the activity of our cultures has demonstrated functional basic excitatory (glutamate) and inhibitory (GABA) synapses, whose effect on the culture activity was consistent with that observed *in vivo*. Moreover, we also observed the presence of functional cholinergic synapses, both nicotinic and muscarinic, as well as presence of cholinergic neurons[21]. Both effects and developmental changes of such cholinergic system have been consistent with those reported in *in vivo* studies.

In an intact brain cholinergic input from subcortical structures innervates diffusively cerebral mantle. Combined with the very specific positioning of cholinergic synapses in local cortical circuitry, this results in coordinated changes in the mode of activity of the cortex in response to changes in the concentration of acetylcholine. This is one of the reasons why the cholinergic system has been implicated by many neuroscientists in such important cognitive functions as working memory, learning and attention[8, 22, 23]. The presence of functional cholinergic system in our cultures suggests that, in principle, they posses the biophysical capacity to support such cognitive functions in suitably embedded cultures.

This is further corroborated by studies of the functional organisation of cultures from our laboratory, as well as those obtained at Steve Potter's lab at Georgia Tech. These results show the development of functional connectivity from initially random to one exhibiting hallmarks of 'small world' networks, similarly to the functional connectivity observed in cortical networks[15, 48]. As functional connectivity is believed to reflect the organisation of a complex system, such as the brain, in ways mirroring its computational properties[49], such similarity indicates that functionally the cultures have the potential to support a range of information processing tasks performed by the cortex *in vivo*. Similarly, the presence of metastable states, which we have identified in such cultures, have been widely suggested, on the basis of numerous animal experiments, to support cognitive processing ranging from perceptual differentiation, through working memory[58].

Although, consistently with other groups doing research on animats, our platform - analogous to a simple Braitenberg vehicle - has shown relatively simple behaviours in the form of obstacle avoidance[56], nevertheless, in terms of complexity, including the number of neurons, their functional connectivity, their computational and biophysical properties etc., showed the capacity for supporting information processing functions observed in intact brains.

Moreover, cultured networks analogous to ours have been shown to respond to open loop conditioning, suggesting that the biological mechanisms present in them can also support plasticity and learning[28, 47]. One of the most interesting of such experiments was performed by Steve Potter's group, which performed a closed loop conditioning of an animat, in which the choice of stimulation patterns was a function of animat behaviour gradually leading to the animat settling on a desired behaviour, (following prespecified direction in this case[1]). This demonstrates that, in principle, such closed loop conditioning can be used to achieve any form of association and henceforth can be incorporated in training an animat to perform much more complex tasks.

Given the above results obtained in ours and other labs, it is not so obvious that the potential of 'animat' devices (for example, to behave with all the flexibility and insight of intelligent natural systems) is as constrained by the standard a priori arguments purporting to limit the power of (the merely Turing machine controlled) robots highlighted earlier. Surely, animats go way beyond conventional robots controlled by computers (i.e. virtually all cognitive robotic systems of today) if not yet in computational or behavioural sophistication, then certainly in their hybrid mechano-biological makeup and non-computational capacity.

Because the tasks the animats perform are actually achieved by embodied 'biological nervous system', they appear to be the best candidates to assuage the concerns of those who, in words of Andy Clark, "... fear that the embodied mind is just the disembodied mind with wheels on"[9]. It seems feasible that as the animat system grows in complexity and their performance becomes more autonomous and sophisticated, the powers of the embodied neural systems will eventually allow them to achieve some form of intentional behaviour, acquiring them status of sentient beings along the way. In particular, forms of closed loop conditioning, such as demonstrated in[1], could be used to train the animat such that the culture would produce patterns of activity that would amount to appropriate manipulation of Chinese symbols, if such were presented to the appropriate sensors. The resultant neural activity could easily be mapped back onto appropriate animat responses, as if the system could answer questions in Chinese with understanding.

## 5 Zombie rodent - an ultimate embodiment

In spite of the animat's obvious advance on completely lifeless robotic systems, the first objection to a specter of a sentient animat could be levelled using recent arguments from the enactivist camp. In a paper from 2011, Cosmelli and Thompson have discussed at great lengths the limitations of 'brain in a vat' setting[11],

> Suppose that a team of neurosurgeons and bioengineers were able to remove your brain from your body, suspend it in a life-sustaining vat of liquid nutrients, and connect its neurons and nerve terminals by wires to a supercomputer that would stimulate it with electrical impulses exactly like those it normally receives when embodied.

Although their imagined setup differed from an animat in that the brain in their gedankenexperiment has been embodied virtually in a simulation by a supercomputer providing appropriate inputs, nevertheless in congruence with the thought experiment an animat also enjoys the presence of biological nervous system and a compatible 'envatment'. Nevertheless we believe that even such systems cannot really possess intentionality for two primary reasons. First, the objections raised by Cosmelli and Thompson with respect to their thought experiment envatment apply equally to the robotic embodiment present in animat. This is because an animat, with all the standard robotic embodiment augmented by the MEA experimental hardware geared towards providing cultures with environment appropriate for their long term survival and function, amounts more to Cosmelli's and Thompson envatment than true embodiment. For the envatment to count as a true embodiment it, in their own words,

> .. would need to be a surrogate body subject to control by the brain. By 'body' we mean a self-regulating system comprising its own internal, homeodynamic processes and capable of of sensorimotor coupling with the outside world.

We agree with Cosmelli and Thompson that, in spite of the superiority of the physical embodiment over simulation, which parallels the difference between simulated and physical robots emphatically stressed by all self-respecting roboticists, even animat embodiment is too impoverished to provide anything more than some form of sensory motor coupling which, as we tried to argue consistently with Cosmelli and Thompson, seems necessary but not sufficient to account for intentional states.

Second, we will argue in the reminder of this section that the lack of a proper embodiment is only a part of the problem; the other equally important deficiency of animats is the mechanistic implementation of their conditioning; as long as the processing is following externally imposed constraints, which are arbitrary from the perspective of 'the brains' biology, there is little chance of the system developing true intentionality. This line of argument will ultimately extend the power of the Chinese Room argument towards properly embodied systems, which nevertheless base their functioning on formal mechanistic and externally driven operations.

The implant technology has advanced beyond creating artificial augmentation of sensory or even cognitive systems. The scientists have tapped into the biological structures in order to induce in the intact living animals specific desired behaviours. These developments offer the possibility of using the operant conditioning and inducing the behaviours in a way analogous to robotic systems, in animals with an otherwise fully intact body. From the perspective of our discussion, this offers a possibility of creating an ultimate embodiment - a system with fully functional biological body, equipped with functional brain and the normal sensory motor coupling, which nevertheless could be driven [via suitable conditioning] to perform specified associations (e.g. Turing style symbol manipulations).

For example, John Chapin and his group inserted an electrode in the medial forebrain bundle (MBF) in a rat's brain[51]. The MBF is believed to be involved in a biological reward system and in generation of pleasurable feelings, which is corroborated by behavioural animal intracranial self-stimulation (ICSS) studies, as well as human subjects reports. Other electrodes were inserted in cortical areas processing information arriving from animal's whiskers. This setup enabled Chapin's group to use operant conditioning in order to train the rat to respond with appropriate turns to stimulation of corresponding whisker areas.

Several days of training taught the animal to start turning according to remote signals without the MFB stimulation, as a remotely controlled robot would. The animals could be steered to navigate through different environments[12] or perform even more complex tasks, such as climbing, although they would not perform tasks which they perceived as 'dangerous'.

The fundamental condition for the success of such training is that for it to work, the experimenter must treat the animal as a sentient being - he must employ the natural desires and goal seeking of an autonomous biological agent and must do so by tapping into the biological machinery responsible for such behaviours. Another, equally important, condition for the success is that in order for the animal to want to follow the training (for the conditioning to take place), it must be able to discriminate consciously the options so that it can form the associations between the target options and reward.

These conditions may seem limiting from the perspective of our discussion on, both, fundamental and pragmatic grounds. First and foremost, employing an existing sentient being's teleological behaviour and conscious discrimination creates the dangerous possibility that the animal could learn to map the imposed associations on its own intentional interpretations and hence could bootstrap its own (rat-level) intentionality onto the formal Turing style symbol manipulation artificially imposed on it. Secondly and more pragmatically, electrodes do not provide sufficient discrimination in delivering stimuli to the

---

[12] See movie of a guided robot at `www.youtube.com/watch?v=D5u2IWFNFDE`

appropriate targets, hence the possibility of generating complex conditioning responses and rich patterns of co-activations that may be necessary for even the most rudimentary forms of cognition might not be possible using such technologies.

However, in recent years, an exciting technique called optogenetics has come to the fore. Optogenetics can provide a sublimely refined levels of control of brain microcircuitry and can henceforth address, at least in principle, both caveats. Optogenetics is in a broad sense a combination of optics, genetics and molecular neuroscience[13]. It uses viral vectors in order to target specific neuron types and make them express light-sensitive proteins identified in algae or bacteria. As these 'opsins' act as ion pumps or channels when activated by light of specific wavelengths, neurons that express them can be specifically and temporally precisely activated or inhibited by laser.

Using optogenetic technology it is possible to make different cell types express different opsins and hence to induce a very precise spatial and temporal patterns of activations and inhibition in the treated tissue. Optogenetics offers the level of spatiotemporal control of manipulation of neural networks activity both *in vitro* and *in vivo* not afforded by traditional chemical or even electric stimulation, thus it presents the possibility to probe, and also to control very precisely, individual targets in order to investigate and manipulate their function.

Such technology was used in a recent study that demonstrated possibility to perform operant conditioning on a mouse. When the animal, expressing activating opsin in parts of the brain involved in reward system (amygdala and the nucleus accumbens), performed a target ecologically neutral response, the researchers shone light into its brain, activating neurons, axons of which formed the path between the two brain regions,[50]. In those animals in which they transfected the same pathway with opsins that would block the activity in response to light, scientists were also able to use light to stop mice exhibiting a previously conditioned response to a relevant cue.

Although their scientific objectives and experimental technologies were quite different, the experiments performed in[50, 51] both obtained desired responses tapping into a creatures volitional systems, effectively manufacturing wilful behaviours consistent with those required by the experimenters. Thus, although both - from our perspective - are subject to the first limiting condition mentioned above, however the experiments performed by Stuber and his colleagues demonstrate the potential level of specificity and temporal precision of stimulations that may be necessary to induce very specific patterns of responses, thus addressing the second, pragmatic limitation highlighted above.

Essentially the same optogenetic techniques were used by Deiserroth group[20], which led to driving a rodent's response in a way not dependent on its willful behaviours or conditioning. The freely moving mouse exploring its surroundings started to move in a very unnatural way, turning consistently left-wise upon commencement of optical stimulation of the right motor cortex. The behaviour returned to normal willful exploratory behaviour as soon as the stimulation was turned off[13].

## *5.1 From intuition pump to physical realisation of thought experiment*

Experiments such as the ones reported above, although conducted with completely different and legitimate research questions in mind, open up a possibility of creating a zombie mouse, in which its behaviour is based on mechanistically developed patterns of activations of brain structures and not related to animal wilful behaviour or conscious perception, as attested by the stark contrast between the artificial behaviour of the animal while under the stimulation and when it is freely behaving (the stimulation is off).

In principle, such optogenetic techniques can be used to deliver very precise control of neural structures in real time with millisecond precision and in closed loop fashion, where optical control is a function of observed neural activity and the resultant behaviours; for example, in conditioning experiments such as those performed by Potter's group on neural cultures[1]. They could be used to achieve desired behaviours in animats or indeed in animals, where the associated patterns of activity need not rely on animal wilful behaviour, thus addressing the first, fundamental limitation mentioned above. Thus, such an animal's brain could be conditioned, upon pattern of activation corresponding to Chinese characters input, to go through a sequence of neural activation patterns resulting in the little murine squeaking a perfectly appropriate response in Chinese (well, not really, but it could produce instead a sequence of lever presses corresponding to such a response)[14].

However, upon inspection of the behaviours of the Deiserroth mouse from the experiments reported in[20], it seems obvious that they are alien to the animal. There is nothing in the animal's intrinsic makeup that would cause it to behave in this way out of its own accord, and it is extremely unlikely that it would ever acquire any intentionality of such externally imposed behaviours.

This is in spite of the fact that the creature would be equipped, by nature, with perfect embodiment and, by experimenter, with artificial sensory-motor couplings resulting in it experiencing the world consistent with induced actions. However, these induced couplings would not be the effect of the intrinsic animal needs (metabolic or otherwise) at any level; to the contrary, they are the cause of metabolic demands. As the animal would be driven, this would cause sequences of sensory-motor couplings, hence it would be the experimenter that would drive these metabolic demands in an arbitrary way (from

---

[13] See movies of experimental animal at `www.youtube.com/watch?v=88TVQZUfYGw`

[14] Selmer Bringsjord proposed a thought experiment surgery on Searle in[6] that was similar in spirit to our zombie mouse; we believe though that at the end of both experiments our zombie mouse would be better off than Searle.

the perspective of metabolic needs of animal or its cellular constituents) thus the casual relationship between the bodily milieu and the motor actions and sensory readings would be disrupted. However, it is the right type of such couplings and their directionality that ultimately leads to intentionality according the enactive approach.

Hence we do not expect that such a ‘zombie mouse’ would acquire any form of understanding of the presented Chinese story. A fortiori, if such formal rulebook following does not lead a sentient being to acquire an understanding, we do not expect that the analogously trained animat with its impoverished envatment for a body, will be any luckier in this respect.

## 6 Conclusions

This paper argued that neither zombie mice nor animats will escape Searle’s CRA, which we suggest continues to have force against claims of their symbol-grounding, understanding and intentionality.

Similar objections towards embodied AI have been put forward in[18, 19], however, their discussion is limited to traditional robotic systems. Our paper extends this line of argument towards hybrid systems, or even systems with fully functional body, which are driven by formal computational rules.

A zombie mouse was used as a vehicle for demonstrating that it is not a ‘trivial’ matter of providing an appropriate embodiment for effectively Turing Machine style operations that could account for emergence of meaning, grounding and teleology. Furthermore, we believe our zombie mouse argument also demonstrates that if the mechanistic account is not consistent with the low level embodiment (as was the case for zombie mouse - the information processing imposed on it is external and arbitrary with respect to the properties of the ‘machinery’ [the brain and the organism] in which it is implemented), then the result is exactly that - a zombie - with no understanding or ownership of the actions imposed on it.

We suggest that what body provides goes over and above what robotic/artificial embodiments can offer: in the right conditions both the natural body and artificial embodiments are a source of correlations of activations of different brain areas caused by different dimensions of real objects/world. As we tried to articulate in this paper, such correlations are important, they may be even necessary, but they do not seem to be sufficient for meaning to arise and this seems to hold true as much for fully artificial system as for those that blend the artificial and biological components.

Finally, we do not wish to appear as providing a wholesale criticism of cognitive robotics. Indeed, we believe that this area offers very fertile grounds for creating experimental platforms for testing information processing aspects of embodied cognitive processing[32]. However, we do remain sceptical whether such systems or their hybrid mechano-biological extensions of late, driven by mechanical formal computational rules are able to answer the most fundamental questions about the nature of intelligence and cognition. In order to achieve such a breakthrough the embodied systems yet to be developed will have to seriously take the body into account.

## References

1. Bakkum DJ, Chao ZC, Potter SM (2008) Spatio-temporal electrical stimuli shape behavior of an embodied cortical network in a goal-directed learning task. *J Neural Engineering* 5(3):310–323
2. Barandiaran X, and Moreno A (2006) On what makes certain dynamical systems cognitive: a minimally cognitive organization program. *Adaptive Behavior* 14(2):171–185
3. Berger TW, Hampson RE, Song D, Goonawardena A, Marmarelis VZ, and Deadwyler SA (2011) A cortical neural prosthesis for restoring and enhancing memory. *J Neural Eng* 8(4):046017
4. Blake W (2000) Cochlear Implants: Principles and Practices, Lippincott Williams & Wilkins; Philadelphia.
5. Botvinick M. and Cohen J.D. 1998. Rubber hand 'feels' what eyes see. *Nature* 391, 756
6. Bringsjord S (2002) Real Robots and the Missing Thought-Experiment in the Chinese Room Dialectic". In *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, eds John Preston and Mark Bishop, New York: Oxford University Press
7. Brunner P, Bianchi L, Guger C, Cincotti F and Schalk G (2011) Current trends in hardware and software for brainŰcomputer interfaces (BCIs). *J. Neural Eng.* 8: 025001
8. Caruana DA, Warburton EC and Bashir ZI (2011) *J Neurosci.* Induction of Activity-Dependent LTD Requires Muscarinic Receptor Activation in Medial Prefrontal Cortex. 31(50):18464–78
9. Clark A (2009) Review of Radical Embodied Cognitive Science by A. Chemero. The MIT Press
10. Cole D (2009) The Chinese Room Argument. In *The Stanford Encyclopedia of Philosophy* (Winter 2009 Edition), Edward N. Zalta ed., `http://plato.stanford.edu/archives/win2009/entries/chinese-room¿`
11. Cosmelli D, and Thompson E (2010) Embodiment or Envatment? Reflections on the Bodily Basis of Consciousness. In *Enaction: Towards a New Paradigm for Cognitive Science*, eds. John Stewart, Olivier Gapenne, and Ezequiel di Paolo, MIT Press
12. Daly JJ, and Wolpaw JR (2008) Brain-computer interfaces in neurological rehabilitation. *Lancet Neurol.* 7:1032–1043
13. Deisseroth K (2011) Optogenetics. *Nat Methods.* 8(1):26–9
14. De Marse TB, Wagenaar DA, Blau AW and Potter SM (2001) The neurally controlled animat: biological brains acting with simulated bodies. *Auton Robot* 11:305–310
15. Downes JH, Hammond MW, Xydas D, Spencer M Becerra VM Warwick K, Whalley BJ and Nasuto SJ (2011) Emergence of a Small-World Functional Network in Cultured Neurons. submitted
16. Fins JJ (2004) Deep Brain Stimulation. In, *Encyclopedia of Bioethics*, 3rd Edition. S.G. Post , Editor-in-Chief. New York: MacMillan Reference. Volume 2, 629–634

17. Fornos A, Sommerhalder J and Pelizzone M (2011) Reading with a simulated 60-channel implant. *Frontiers in Neuroscience* 5:57
18. Froese T, and Di Paolo EA (2011) The Enactive Approach: Theoretical Sketches From Cell to Society. *Pragmatics & Cognition* 19(1): 1–36
19. Froese T and Ziemke T (2009) Enactive Artificial Intelligence: Investigating the systemic organization of life and mind. *Journal of Artificial Intelligence.* 173(3-4):466–500
20. Gradinaru V, Thompson KR, Zhang F, Mogri, M, Kay K, Schneider MB and Deisseroth K (2007) Targeting and readout strategies for fast optical neural control in vitro and in vivo. *J Neurosci.* 26;27(52):14231–8
21. Hammond MW, Xydas D, Downes JH, Becerra VM, Warwick K, Nasuto SJ and Whalley BJ (2012) Endogenous cholinergic tone modulates spontaneous network level neuronal activity in primary cortical cultures. submitted
22. Hasselmo ME and Cekic M (1996) Suppression of synaptic transmission may allow combination of associative feedback and self-organizing feedforward connections in the neocortex. *Behavioural Brain Research* 79(1-2): 153–161
23. Hasselmo ME, Schnell E et al. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *J. Neurosci.* 15(7): 5249–5262
24. Hatsopoulos NG, Donoghue JP (2009) The science of neural interface systems. *Annual Review of Neuroscience* 32:249–266
25. Kringelbach ML, Jenkinson N, Owen SLF and Aziz TZ (2007) Translational principles of deep brain stimulation. *Nature Reviews Neuroscience.* 8:623–635
26. Krusienski DJ, Grosse-Wentrup M, Galan F, Coyle D, Miller KJ, Forney E and Anderson ChW (2011) Critical issues in state-of-the-art brain–computer interface signal processing. *J. Neural Eng.* 8: 025002
27. Lebedev MA, and Nicolelis MA (2006) Brain-machine interfaces: past, present and future. *Trends Neurosci.* 29:536–546
28. Marom S, Shahaf G (2002) Development, learning and memory in large random networks of cortical neurons: lessons beyond anatomy. *Quarterly Reviews of Biophysics.* 35:63–87
29. Maye A and Engel AK (2011) A discrete computational model of sensorimotor contingencies for object perception and control of behavior. 2011 IEEE International Conference on Robotics and Automation (ICRA)
30. McCarthy J, Minsky ML, Rochester N and Shannon CE (1955), A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, Stanford, USA
31. Mintz M (2011) A biomimetic model aimed at recovering learning in a brain damaged animal: Converging neuroscience with technology. Strategies for Engineered Negligible Senescence Meeting, Cambridge, UK
32. Morse AF, Herrera C, Clowes R, Montebelli A and Ziemke T (2011) The role of robotic modelling in cognitive science. *New Ideas in Psychology* 29(3):312–324
33. Mussa-Ivaldi FA, and Miller LE (2003) Brain-machine interfaces: computational demands and clinical needs meet basic neuroscience. *Trends Neurosci* 26:329–334
34. Muzumdar A (2004) Powered Upper Limb Prostheses: Control, Implementation and Clinical Application. Springer
35. Newell A and Simon HA (1976) Computer Science as Empirical Inquiry: Symbols and Search *Communications of the ACM* 19 (3): 113–126
36. Nicolelis MAL and Chapin JK (2002) Controlling robots with the mind. *Scientific American* 287:46–53
37. Nöe A (2004) Action in perception. Cambridge, Mass: MIT Press
38. Nöe A (2010) Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness
39. Novellino A, Chiappalone M, Tessadori J, D'Angelo P, Defranchi E, and Martinoia S (2011) Bioartificial Brains and Mobile Robots. In *Mobile Robots - Control Architectures, Bio-Interfacing, Navigation, Multi Robot Motion Planning and Operator Training*, ed. Janusz Brzedkowski, InTech, 2011

40. O'Regan JK and Nöe A (2001) A sensorimotor account of visual perception and consciousness. *Behavioral and Brain Sciences*, 24, 939–1011
41. O'Regan JK (2010) How to make a robot that feels. 4th International Conference on Cognitive Systems, CogSys 2010, Zurich, Switzerland
42. O'Regan JK (2011) Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness New York: Oxford University Press
43. Pfeifer R and Scheier C (1999) Understanding intelligence. Cambridge MA: MIT Press
44. Pfeifer R and Bongard J (2007) How the body shapes the way we think: a new view of intelligence,T̃ MIT Press
45. Preston J and Bishop M eds. (2002) Views into the Chinese Room: New Essays on Searle and Artificial Intelligence, New York: Oxford University Press
46. Searle J (1981) Minds, Brains, and Programs, *Behavioral and Brain Sciences* 3: 417–457, Cambridge, UK: Cambridge University Press
47. Shahaf G, Marom S (2001) Learning in networks of cortical neurons. *J Neurosci* 21:8782–8788
48. Spencer M, Downes JH, Xydas D, Hammond MW, Becerra VM, Warwick K, Whalley BJ and Nasuto SJ (2011) Multi Scale Evolving Complex Network Model of Functional Connectivity in Neuronal Cultures, IEEE Transactions on Biomedical Engineering, (epub ahead of print)
49. Sporns O (2011) Networks of the Brain. The MIT Press
50. Stuber GD, Sparta DR, Stamatakis AM, van Leeuwen WA, Hardjoprajitno JE, Cho S, Tye KM, Kempadoo KA, Zhang F, Deisseroth K and Bonci A. (2011) Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature* 475:377–380
51. Talwar SK, Xu S, Hawley E, Weiss S, Moxon K and Chapin J (2002) Rat navigation guided by remote control. *Nature* 417, 37–38
52. Tan DS and Nijholt A eds. (2010) Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction. Springer
53. Varela F, Thompson E and Rosch E (1991) The embodied mind: Cognitive science and human experience. Cambridge, MA: MIT Press
54. Warwick K, Gasson M, Hutt B, Goodhew I, Kyberd P, Andrews B, Teddy P and Shad A (2003) The Application of Implant Technology for Cybernetic Systems. *Archives of Neurology* 60 (10): 1369–73
55. Warwick K, Xydas D, Nasuto SJ, Becerra, VM, Hammond MW, Downes JH, Marshall S and Whalley BJ (2010) Controlling a mobile robot with a biological brain. *Defence Science Journal* 60 (1): 5–14
56. Warwick K, Nasuto SJ, Becerra VM and Whalley BJ (2010) Experiments with an In-Vitro Robot Brain. Chapter in *Instinctive Computing*, ed. Yang Cai, Lecture Notes in Artificial Intelligence, Springer, Vol. 5987
57. Zrenner E (2002) Will retinal implants restore vision? *Science* 295: 1022–5
58. Xydas D, Downes JH, Spencer MC, Hammond MW, Nasuto SJ, Whalley BJ, Becerra VM and Warwick K (2011) Revealing ensemble state transition patterns in multi-electrode neuronal recordings using hidden Markov models. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* , 19 (4): 345–355