ESANN'2009 proceedings, European Symposium on Artificial Neural Networks - Advances in Computational Intelligence and Learning. Bruges (Belgium), 22-24 April 2009, d-side publi., ISBN 2-930307-09-9.

Phenomenal Weightless Machines

Igor Aleksander[1] and Helen B. Morton[2]

[1]Imperial College
London, UK

[2]Brunel University
Uxbridge, Middlesex, UK

Abstract: This paper describes how early designs of dynamic weightless neural systems were developed to enable some of the states of a state structure to have a phenomenal character. Such states reflect the features of a sensory reality and allow the storage of aspects of sensory experience and access to it. The 'machine consciousness' paradigm is summarised in this paper. The paper concludes with a description of the current state-of-the-art of a phenomenal approach to a model of consciousness which is based on the first of a set of introspective axioms.

1. Introduction.

This paper is a review of the role that weightless systems (see Aleksander and Morton, 1990, for an early definition) have played in the development of the 'machine consciousness' paradigm. The aims of those who contribute to 'machine consciousness' are first, to clarify what it is for an organism, whether it be human, animal or artefact to be conscious. Second is the aim to examine the potential for informational machines to *be* conscious and what benefit this might bring to the general area of cognitive computation.

A phenomenal state in a recursive system is one that is responsible for the behaviour of the system by reflecting implicitly the properties of the real world. After considering the origins of the 'machine consciousness' paradigm, the nature of computationally phenomenal states is introduced and a distinction between functional and phenomenal virtual machines is drawn as phenomenology is thought to be essential for being conscious. Phenomenal designs will be seen to evoke the neural computational domain. We describe in particular, how weightless neurons are suited not only to create neural state machines that internalise experience in their state structures, but also how such states may be made subjective (i.e reflect reality as sensed from the point of view of the organism). It is seen that this property comes from two specific modes of operation of the weightless neuron: *iconic transfer* and *muscular indexing*. This is based on the 'axiomatic/introspective method', which decomposes the concept of being conscious into elements which have reasonably clear transitions into neural architectures.

2. A brief introduction to machine consciousness[1]

In 2001, the Swartz foundation, which usually organises meetings on brain science, organised a 3-discipline (philosophy, computation, neuroscience) workshop on the question of ‘could a machine be conscious?’. While there were many disagreements, one area of agreement (as summarised by one of the organisers, Christof Koch)[2] was:
“*...we know of no fundamental law or principle operating in this universe that forbids the existence of subjective feelings in artefacts designed or evolved by humans.*”
This statement carries a streak of optimism as well as a challenge to find ways in which machines with subjective states could be designed. This gave rise to several projects, briefly listed below, that attempt to do this.

This declaration, when published, led to several expressions of scepticism (which, indeed, may occur to the current reader) that need to be made explicit before proceeding. These mainly assume the unassailability of Chalmers ‘Hard Problem’ (Chalmers, 1996)[3]. In this paper it is argued that the ‘hard problem’ is addressed by relating it to what is known of the relationship of physical structure to state structure in automata theory. In addition to this it is necessary to show how the inner states of a state machine can become subjective, here done through iconic transfer and muscular indexing.

One of the oldest models is known as ‘Global Worksapace Theory’, developed by Bernard Baars (1988). This assumes that there are several *unconscious* processes such as various forms of memory, volitional and emotional activities that compete for entry into an architectural element known as the ‘global workspace’. The competition is won by the process that is most salient for the sensory input present at the time. Then takes place a key step: the content of the global workspace are broadcast to the competing processes changing their state. This is the moment of consciousness and it is a sequence of such moments that constitutes the system’s ‘stream of consciousness’.

While this system has no pretence of phenomenal consciousness, a move towards phenomenology was executed by Murray Shanahan (2005) using the weightless neural methods that are discussed in this paper. Does ‘global workspace theory’ have a meaning in neurophysiology? A positive answer was given by Dehaene and Naccache (2001) who showed that areas of the brain that include the prefrontal cortex, the anterior cingulate and related areas, form a global workspace and stand in appropriate relation to distant areas that carry unconscious memory processes.

Another noteworthy contributor to machine consciousness is Pentti Haikonen who published two major books on the subject (2003, 2007). He believes that most characteristics of being conscious can be represented in a repetitive architecture of conventional, weighted neural networks.

---

[1] Those wishing to read a summary of this paradigm may do so on: http://www.scholarpedia.org/article/Machine_consciousness
[2] http://www.theswartzfoundation.org/abstracts/2001_summary.asp
[3] This suggests that science can only be done on the physical (body) and only correlations can be found to the subjective (mind) . Chalmers has argued that the ‘hard problem’ for science is that it cannot prove that the physical implies the subjective..

ESANN'2009 proceedings, European Symposium on Artificial Neural Networks - Advances in Computational Intelligence and Learning. Bruges (Belgium), 22-24 April 2009, d-side publi., ISBN 2-930307-09-9.

3. Virtual machine functionalism

We recall that a virtual machine is one that runs on a host machine and the properties of which can be studied independently, without reference to the operation of the host machine. Functionalism is a philosophical notion that addresses the *behaviour* of an organism in a real world as a result of its perception of that real world. In the general case, philosopher Ned Block (1996) has characterised a functional system as a state machine, where the state (seen as a 'mental' state) change is necessary to keep track of a developing environmental reality without any particular restriction on the coding of such states. He illustrates this by suggesting that if a mental state moves from tranquillity into a state of pain, all this instigates is a propensity to 'say ouch' or have other 'thought states' that are contingent on the pain state. Calling this an 'atomic' view of functionalism, Sloman and Chrisley (2003) pointed out that a lack of clarity sets in if a functional system where many conscious processes may be ongoing, is represented as a single state. This led them to define virtual machine functionalism (VFM) by stating that a functional mental state as one in which many conscious processes are present simultaneously each with its own state structure. For example a headache state might be accompanied by thoughts of phoning a doctor, the effect on writing a paper, needing to cook a meal and paying one's bills. That is, it is important to recognise that several automata may be acting simultaneously each providing an element of an overall mental state. Such automata are highly variable, and their essence is 'virtual' in the brain.

Taking one step back it is noted that philosopher Daniel Dennett (1993), evoked a virtual machine approach:

> Human consciousness … can best be understood as the operation of a "*Von Neumannesque"* virtual machine *implemented* in the parallel architecture of the brain that was not designed for any such activities.

The key phrase here is that it may be wrong to look for a design that specially supports the states of a functionally conscious system, but that such a system almost accidentally does this task. The real meaning is that it is the virtual machine that defines the presence of mental states and this can have many if not an infinity of supports. The reference to a 'Von Neumannesque' machine also appears unnecessary. The key issues for VFM then are that whatever it is for the machine to be conscious might be expressed as a Virtual Machine that reflects the complexity of multiple interacting state machines. As bounded infinity of physical structures can support such a VM, the trick is to find some bounding constraints. Sloman and Chrisley have done this by identifying interacting layered schemes: *horizontal* going from the reactive to the deliberative to the managerial and *vertical* going from sensory input to its interpretation ending in a vetoed system of action.

We leave aside work in machine consciousness concerned with robots in order to now show how the VMF approach leads to a phenomenal approach that can be satisfied by using weightless systems.

4. Virtual Machine Phenomenology

Phenomenology is a study of consciousness said to have been funded by German philosopher Edmund Husserl who defined it as (1901): *"The reflective study of the essence of consciousness as experienced from the first-person point of view"*. A phenomenal system therefore is one in which the internal state(s) have a capacity for representing reality directly in a way that is a decent approximation of some external reality. Such states must be parts of state structures (i.e. virtual machine) that represent the behavioural experience of the organism. In order to achieve an unrestricted reflection of reality, a fine-grain representation is implied where the grain is determined by the minimal changes in an external world of which the system is to become conscious.

4.1 A definition of a weightless neuron use in phenomenal systems

We recall that one type of weightless neuron maps an n-input binary vector X into a binary variable z which can have value 0, 1 and u , where u represents a random choice between 0 and 1. Learning takes place during a training period when a special binary 'teaching' input line d (desired) of the neuron determines whether X is associated with z=0 or z=1 which is stored in the neuron's lookup table which is normally in state u before training takes place. If during a training sequence, the stored value of 0 or 1 is contradicted, the stored lookup state for the contradicted X reverts to the u state.

As generally defined, the weightless neuron also generalises to the extent that if an unknown input vector X_u is compared to the X_j of (X_j, d_j) pairs on which the neuron was trained, and there is a distinct X_j which is closer to X_u (in Hamming distance, say) than any other, then the neuron will respond with the corresponding d_j.

4.2 Iconic Transfer and Phenomenal states

Say that a network consists of k neurons, each with n inputs, which is 'connected' to a pattern P that consists of a bits. The connection is made at random. Then there exists a set of teaching lines $D=\{d_1,d_2 \ \dots \ d_k\}$ which, after a training step, defines the k-bit output pattern Q. Now, if D is connected to pattern P as well, Q learns to be a sampling of P.

Transferring this now to a recursive network in which the n inputs of each neuron not only sample P, but also Q (possible with a defined ratio), Q becomes the state of a neural automaton. We submit that this is a *phenomenal* state as it is determined by P alone which is the interface at which the reality of the automaton's environment is represented. Figure 1 shows the development of a phenomenal state in a 144x144 (the dimension of Q) neuron network with a 144x144 input (the dimension of P). Each neuron has 288 binary inputs, 144 randomly drawn from the input P and 144 randomly drawn from state This is a model of the tool-making ability of 'Betty', a crow studied in the zoology department at Oxford University[4]

[4] http://users.ox.ac.uk/~kgroup/index.html

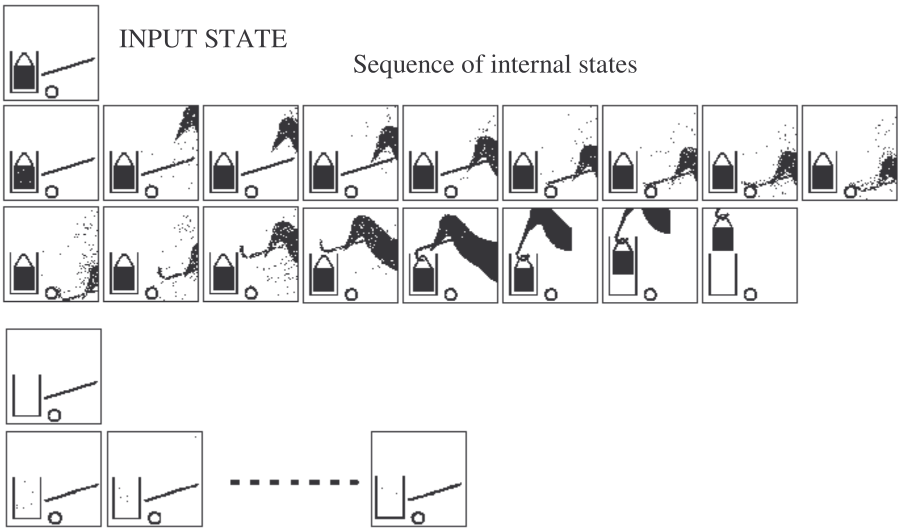

Figure 1: Quasi-phenomenal behaviour of an 144x144 weightless neural network for making the tool that can extract a food canister from a jar.

The above weightless neural state machine was trained by being exposed to the shown sequence, illustrating that *iconic transfer* may create a state sequence that represents past sensory experience. This may be triggered by an initial input state, and the internal sequence then becomes an imaginational representation of future action. When executed, the action leads to the new input state in the lower group which leads to a different internal sequence – one for taking no action.

The reason for referring to this as a quasi-phenomenal representation lies in the fact that it is a 'third person' view and not a 'first person' experience as required in phenomenology. To go beyond the third person we briefly look mainly at the first ('presence') of some previously published introspective axioms (Aleksander, 2005) and comment on the role that weightless systems play in their consideration. These five axioms are a breakdown of what important elements of consciousness feel like and how they may be translated into neural mechanisms: presence, imagination, attention, volition and emotion.

5. Axioms and weightless systems.

The presence axiom states that a primary element of consciousness is the feeling that one is an entity in the 'out-there' world. To achieve this, the 'out-there-world' needs to be phenomenally represented as being unaffected by the actions (e.g. eye movement, head movement, body movement …) of the organism. That is, it makes it possible to represent the independence of the 'self' in the world. To achieve this it is required that whatever sensory input is being represented, is must be compensated for the acquisition actions of the organism. Say that the eye is foveally fixated on the nose of a face. Say we give the position of the nose the spatial origin *x,y* coordinates 0,0, and allow that an internal phenomenal representation of the nose in a neural area indexed 0,0. Now say that the gaze shifts slightly to see the ear at coordinates 1,0 (in non-defined units). This means that a new neural area centred on 1,0 has to be iconically activated. The implication is that the neural network training of weightless neurons is *indexed* on muscular activity. That is, eye movements create a phenomenal inner state larger than the foveal area. However other more major movements extend the phenomenal representation into state trajectories structured by the more major movements. That is, a head movement from coordinates *x,y* to *x',y'* will cause a related state change without changing the set of neural state variables.

6. Conclusion

This paper briefly traced the early definition of a weightless neuron into its application as the building brick of informational models of neuronal systems that support phenomenal consciousness. The machine consciousness paradigm is addressed and two necessary properties for the weightlessneural modelling have been seen to be necessary to ensure phenomenal representations: iconic learning and muscular indexing.

References:

Aleksander, I, and Morton, H.B.: *An Introduction to Neural Computing,* London: Chapman and Hall, (1990).

Chalmers, D.: *The Conscious Mind:In Search of a Fundamental Theory,* Oxford (1996).

Baars, B.: *A cognitive theory of consciousness*, NY: Cambridge University Press (1988).

Shanahan, M.: Cognition, Action Selection and Inner Rehearsal, *Proceedings IJCAI 2005 Workshop on Modelling Natural Action Selection*, pages 92-99 (2005).

Dehaene, S., Naccache, L.: Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1-37 (2001)

Haikonen, P.: *The Cognitive Approach to Conscious Machines*, Exeter: Imprint Academic, (2003)

Haikonen, P.: *Robot Brains: Circuits and Systems for Conscious Machines,* Chichester: John Wiley and Sons, (2007)

Block, N.: 'What is functionalism?'. The Encyclopedia of Philosophy Supplement, Macmillan (1996)

Sloman, A. and Chrisley, R.: Virtual Machines and Consciousness, *Journal of Consciousness Studies*, 10, No. 4-5, (2003)

Husserl, E.: *Logical investigations.* Eng. Trans by J. N. Findlay, London Routledge,1973. Original in German (1901)

Aleksander, I.: *The World in my Mind, My Mind in the World: Key Mechanisms of Consciousness in Humans, Animals and Machines",* Exeter: Imprint Academic (2005)