SYMPOZJUM SPECJALNE

MACHINE CONSCIOUSNESS

ABSTRAKTY:

Piotr Bołtuć (University of Illinois Springfield)

Non-Reductive Consciousness may be more like Hardware than Software

Progress of cognitive architectures towards functional consciousness is impres-sive (e.g. Franklin's LIDA, Thaler's imagitrons). How to attain what Block calls phenomenal consciousness is less clear, largely because the notion is ambigu-ous. It covers two different areas: 1. Functional phenomenal facts (e.g. a drone's responce to color or other impulses that help produce qualia in humans). 2. Phe-nomenal facts unique to first-person awareness. Most authors either deny that the latter consciousness can even be discussed (Ryle, Dennett, the Churchlands) or deny that such consciousness could be attributed to machines (Searle); Chalmers who is a non-reductive functionalist is an exception. Paradigmatic case of con-sciousness is exemplified when a health professional says that a patient regained consciousness (he or she may have no understanding of their surroundings, very low intelligence, they may even suffer the locked in syndrome). I argue that the unique aspect of first-person consciousness comes from the background of the stream of awareness that is present even in patients where content of conscious-ness is minimal. Yet I argue (contra Searle) that once neuroscience learns how such stream is generated in animals we should be able to use the same blueprint to build it in artificial cognitive architectures. This is the engineering thesis in machine consciousness [Boltuc 2007, 2009, 2012]. Time permitting I shall tackle the main objections to this argument, including the problem of epiphenomenalism.

David Gamez (University of Sussex, Brighton)

From Human to Machine Consciousness

In philosophy and popular culture there has been a great deal of speculation about the consciousness of machines. In these discussions people often use their intuition and imagination to decide whether a machine is conscious. If a robot looks like a human and behaves in a similar way to a conscious human, then we are inclined to attribute consciousness to it. One problem with this approach is that there are an extremely large number of ways in which a robot's external behaviour can be generated – for example, a giant lookup table, the popula-tion of China communicating with radios and satellites or a standard computer program. It is hard to believe that all of these systems are conscious, and it has been argued that any system can be interpreted as implementing a given func-tion, which would lead to an

untenable panpsychism. A second problem is that a system can be conscious without any external behaviour. For example, we are often conscious when we are dreaming or using our imagination. In humans it is generally believed that consciousness is linked to spatiotemporal patterns in the brain that might or might not be connected to external behaviour. A human brain is judged to be in a particular conscious state if it contains spatiotemporal patterns that are correlated with the state. This talk argues that we should take a similar approach to machine consciousness: we should base our judgements about a machine's consciousness on the spatiotemporal patterns in its internal states. First, we need to study the human brain to develop mathematical theories that link spatiotemporal patterns in the physical world to conscious states. We can then use these theories to make reliable inferences about the consciousness of artificial systems.

Włodzisław Duch (Nicolaus Copernicus University)
Signs of consciousness in humans and machines

What could convince us that artificial system is conscious? Although the class of conscious machines could be broader I'll focus on artifacts that have brain-like structure. The von Neumann universal computer metaphor does not capture dynamical, self-modifying, data-flow reactive information processing nature characteristic of the brain processes. Neurodynamical framework seems to be sufficient to analyze spatiotemporal neural activity and relate it to subjective, first person experiences. Processes at molecular level may pose important constraints on neurodynamics and are critical to understand influence of psychoactive drugs, but such effects may be captured by parameters defining neurodynamics. Dynamic dataflow machines that use content-addressable memory may serve as higher-order abstraction for brain processes but we do not have yet examples of real-time machines of this sort.

Neurodynamical theories of consciousness have been developed by many scientists, including Freeman, Varela, Thomson, Scott Kelso, Bressler, Llinas, Singer, Crick, Koch, Dehaen, Changeux, and their collaborators. Edelman and Tononi focused on integration and differentiation of brain processes, providing first quantitative measures of the consciousness level, applicable to brains under anesthesia, or to the disorders of consciousness. In their approach a group of neurons with highly correlated activity forms a functional cluster, a dynamic core process that integrates information and is constantly changing. This can be seen in neuroimaging and also evaluated in simulated neural networks. The brain-computer interfaces learn to identify and assign meaning to spatiotemporal patterns of activity in the brain, translating such patterns into decisions and actions. Progress in this direction will eventually show how to transform brain activity into areas of psychological spaces defined by dimensions that characterize subjective experiences.

Communication between cognitive systems, including brains and machines, may be characterized in terms of resonance patterns, invoking functionally equivalent states in both cognitive systems. Adding the ability to comment on a sequence of transitions between internal states will create a verbal stream of consciousness. Arguments against consciousness in such machines will be difficult to find.

Recently Cruse and Schilling have created an artificial hexapod walking insect controlled by small artificial neural network. The insect is moving in an unpredictable environment. Some attributes of consciousness may already be identified in such simple embodied system. Prospects for building conscious machines capable of creating theory of mind will be discussed.