

Invitation

Models of Consciousness Workshop

The Search for a Unified Theory

Organized by

Ricardo Sanz, Universidad Politécnica de Madrid, Spain

Aaron Sloman, University of Birmingham, U.K.

Ron Chrisley, University of Birmingham, U.K.

With the support of

The European Science Foundation

Manor House, University of Birmingham

September 1-3, 2003

V 1.0 2003/05/29

Invitation

*The organisers
of the
Models of Consciousness Exploratory Workshop
would like to invite you
to participate in this multidisciplinary event.*

Presentation

The **Models of Consciousness** workshop tries to advance the elaboration of a unified scientific theory of consciousness. A central research topic of the workshop is *unified theories of natural and artificial consciousness* and this workshop focuses on the particular aspects of *models of consciousness* that are also suitable for implementation, i.e. theories of consciousness that can support the construction of conscious machines and also serve as explanation of the experimental data about the natural kind of consciousness.

The objective of the workshop is the presentation, evaluation and discussion of available models of consciousness both from a *scientific point of view* (providing explanation of, e.g., observed behavior) and *technological point of view* (serving as base design for building conscious machines). The objective of the evaluation is the definition of potential candidate architectures for various forms of consciousness that will serve as focal points for future collaborative research projects.

The workshop will gather a multidisciplinary collection of European researchers into the fields of artificial intelligence, neuroscience, robotics, psychology, physics, automatic control, computer science, philosophy, etc.

This workshop will be very valuable to initiate the exchange of knowledge and experiences between researchers from across Europe in an emerging area of research; due to its necessary multidisciplinaryity it will help establish new collaborative links between different disciplines; and as a major activity it will serve to test innovative ideas and develop potential collaborations.

The Context for the Workshop

The objective of building *conscious machines* was already a research topic in the early years of artificial intelligence, but the extreme difficulties encountered at that time in developing implementable models of even the simplest features of human intelligence halted the research and put *machine consciousness* into the bin of *Utopian research topics* (more or less like *time-travel, immortality or hair-restoring*).

But the case for *consciousness* is a little bit different because consciousness *does exist now*. Consequently, we know *a priori* that the construction of a conscious entity is possible. Research in artificial consciousness is not any longer Utopian research for several reasons:

Recent advances in *neuroscience* and *physiological psychology* have provided a deeper knowledge of how human and animal consciousness arises.

Modern *robotic mental architectures* have refocused from abstract problem solving systems onto systems that continuously interact with the world, which some take to be an important step toward the explanation and creation of (a particular kind of) awareness (i.e. flea-type awareness).

Developments in the field of *complex control systems* delved into theories of how model-based control-loop integration can lead to an integrated perception of the self of the controller and the plant.

Computer technology now provides computational power that is many orders of magnitude beyond what was available in the past. It has been estimated that in ten to twenty years, computers will reach the computational power of the human brain (from a certain perspective).

A recent explosion of interest in consciousness on the part of philosophers has led to the most sophisticated conceptual understanding to date of the possibilities for a scientific theory of consciousness, and the potential obstacles to such.

We understand that the main problem for having a good scientific theory of consciousness lies in the wide scattering of knowledge among a wide collection of disciplines. It is necessary to employ coherent, interdisciplinary approaches to the problem to get a glimpse of a good solution. This is why this workshop is essential.

There have been previous attempts to advance the development of a theory of consciousness but, to our knowledge, this is the first attempt to filter theories on the basis of technological applicability (engineering filtering).

Non-implementable theories (*i.e.* theories that claim that consciousness is *indefinable or unknowable*, theories that say that consciousness is *epiphenomenal* and hence has no causal powers, theories that say that consciousness is *just a myth invented* by philosophers, theories that say that *no machine could have it* or theories that say that *machines that are indistinguishable from us could lack consciousness*) are not useful for the engineering work. It is necessary to re-consider their suitability as scientific explanations.

Focus on Interdisciplinarity

This work is interdisciplinary by nature and by need because there is no single discipline that can provide all the relevant knowledge, nor the necessary broadness, nor the essential tools to build such a machine.

The people who have the data about natural consciousness can be found in the disciplines of **psychology, medicine, neurobiology, ethology, linguistics, anthropology, and philosophy.**

The available scientific models of natural consciousness can be found in **neuroscience, cognitive science and biophysics.**

Some **artificial intelligence** and **robotics** researchers now regard artificial consciousness as one of the basic research objectives of the field.

If we are going to “engineer” conscious machines it is obvious that it is necessary to have engineering disciplines represented in the workshop. We have selected people from **computer engineering, software engineering and automatic control engineering.**

Expected outcome of the workshop

The expected outcome of the workshop is divided into two parts. The first one is a collection of scientific-technical results:

- A sound characterization of the field
- An engineering definition of consciousness
- A seed for a sound theory of consciousness
- Draft requirements for architectures for conscious machines of various sorts

The second part is a planning of future research activities in the field:

- Research roadmap
- Launch a network of excellence
- Elaborate concrete collaborative research projects

Workshop Draft Programme

The workshop will be three days long and each day the structure will be similar around a central topic: the data, the models, the implementations. The structure of the discussion will be as follows:

Day One: **The Relevant Data**

1. Introductory speech to the topic of the day: *Data to be explained by a Theory of Consciousness*
2. Morning session M1 with workshop participants' contributions
3. Afternoon session A1 with workshop participants' contributions
4. Closing speech: *A summary of the data that a consciousness theory must explain. A summary of the data that any proposal of a conscious cognitive architecture should be capable of explaining.*
5. Day Discussion: *The relevant data and the targets of potential explanations of consciousness*

Day Two: **Theories of Consciousness**

1. Introductory speech to the topic of the day: *Meanings, Theories and Models of Consciousness*
2. Morning session M2 with workshop participants' contributions
3. Afternoon session A2 with workshop participants' contributions
4. Closing speech: *Candidate architectures for natural and artificial consciousness.*
5. Day Discussion: *The (multiple) nature of consciousness and the feasibility of a ultimate theory*

Day Three: **Engineering Consciousness**

1. Introductory speech to the topic of the day: *Engineering Conscious Machines*
2. Morning session M3 with workshop participants' contributions
3. Afternoon session A3 with workshop participants' contributions
4. Closing panel: *Architectures for Consciousness. Candidate architectures for engineering conscious machines. Summary of the workshop and proposals for cognitive architectures that can guide an engineering process of building conscious machines.*
5. Closing Discussion: *Constructibility of conscious entities*

People Invited

*The participation in the workshop is only possible by invitation.
We will be very glad of having your ideas at this workshop.*

This is the list of people invited to participate in the event:

Antti Revonsuo	(philosophy, University of Turku, Finland)
Axel Cleermans	(psychology, Université Libre de Bruxelles, Belgium)
François Anceau	(computer science, CNAM, France)
Geraint Rees	(neuroscience, University College London, UK)
Germund Hesslow	(neurobiology, Lund University, Sweden)
Giorgio Butazzo	(computer science, University of Pavia, Italy)
Rodeney Cotterill	(biophysics, Technical University of Denmark)
Giulio Sandini	(robotics, Genova University, Italy)
Igor Aleksander	(electrical engineering, Imperial College, UK)
Jackie Andrade	(psychology, University of Sheffield, UK)
Jim Doran	(artificial intelligence, University of Essex, UK)
John Taylor	(mathematics, King's College, UK)
Pentti Haikonen	(cognitive science, Nokia, Finland)
Kevin O'Regan	(psychology, CNRS, France)
Manfred Spitzer	(psychiatry, University of Ulm, Germany)
Miguel Salichs	(robotics, Carlos III University, Spain)
Mike Denham	(control systems, University of Plymouth, UK)
Owen Holland	(robotics, University of Essex, UK)
Riccardo Manzotti	(robotics, Genova University, Italy)
Susan Greenfield	(neuroscience, Oxford University, UK)
Thomas Christaller	(robotics, Fraunhofer, Germany)
Tom Ziemke	(computer Science, University of Skovde, Sweden)
Jacques Lacombe	(neurosurgery/engineering, European Commission)
Ralph Dum	(physics, European Commission)
Pekka Karp	(robotics, European Commission)
Colette Maloney	(computer science, European Commission)
Stan Franklin	(artificial intelligence, University of Memphis, USA)

Organisers

Aaron Sloman	(artificial Intelligence, University of Birmingham, UK)
Ron Chrisley	(philosophy, University of Birmingham, UK)
Ricardo Sanz	(automatic control, Technical Univ. of Madrid, Spain)

Costs

No registration cost.

Accommodation and meals will be paid by the organisers to people invited to present their ideas (if coming from an ESF country).

Travel grants (100€ UK, 200€ Europe) can be provided if necessary.

Do you want to participate ?

If you have any question or want to participate, send an e-mail to any of the proposers before June 15:

Ricardo.Sanz@etsii.upm.es
A.Sloman@cs.bham.ac.uk
R.L.Chrisley@cs.bham.ac.uk

Please give full address, proposed speech title and abstract and desired session/focus (relevant data, theories, engineering models).



See you in Birmingham !!!