The Human Mind Project

Computers and Minds

21st November 2014

University of Edinburgh

As part of the *Being Human festival of the humanities*, the fourth public event of *The Human Mind Project* brought together leading academics to discuss the current state of the art of research at the interface of philosophy, computational neuroscience, robotics and artificial intelligence. "Computers and Minds" featured presentations on computational theories of mind, the cognitive neuroscience of self-consciousness and body awareness, and the implications of the latest advances in artificial intelligence for understanding the human mind. The workshop was followed in the early evening by a public event, with a presentation by Sethu Vijayakumar on "Robots that Learn: The Future of Man or the 'Man of the Future'?" including a demonstration on robotics, and a lecture by Andy Clark entitled "Being and Computing: Are You Your Brain, and Is Your Brain a Computer?"

@BeingHumanFest #BeingHuman14

beinghumanfestival.org

Event Programme

The Institute for Advanced Studies in the Humanities (IASH)

9.30	*Registration and Coffee*
9.45	Welcome Addresses

Barry Smith, Director of the Institute of Philosophy, London; and Director of *Being Human* festival of the Humanities

Colin Blakemore, Professor of Neuroscience and Philosophy, London; and Project Leader of *The Human Mind Project*

Morning Session

Chair: **Sanjay Modgil**, King's College London

10.00	**Murray Shanahan,** Professor of Cognitive Robotics, Imperial College London
	ARE WE ON THE ROAD TO HUMAN-LEVEL ARTIFICIAL INTELLIGENCE?
10.45	Discussion
11.00	*Coffee Break*
11.30	**Frances Egan,** Professor of Philosophy, Rutgers University
	UNDERSTANDING COMPUTATIONAL MODELS OF MIND
12.15	Discussion
12.30	*Lunch Break*

Afternoon Session

Chair: **Mattia Gallotti** University of London

13.30	**Aaron Sloman**, Honorary Professor, University of Birmingham
	DISEMBODIED MOTIONLESS INTELLIGENCE
14.15	Discussion
14.30	*Coffee Break*
15.00	**Anil Seth**, Professor of Cognitive and Computational Neuroscience, University of Sussex
	MIND THE BODY

15.45	Discussion
16.00	General Discussion
16.15	*End of Academic Workshop*

David Hume Tower, Faculty Room North

Chair: **Shamil Chandaria**, Institute of Philosophy, London

17.00	**Sethu Vijayakumar**, Professor of Robotics, University of Edinburgh
	ROBOTS THAT LEARN: THE FUTURE OF MAN OR THE 'MAN OF THE FUTURE'?
18.00	Discussion
18.15	**Andy Clark**, Professor of Philosophy, University of Edinburgh
	BEING AND COMPUTING: ARE YOU YOUR BRAIN, AND IS YOUR BRAIN A COMPUTER?
19.00	Discussion
19.30	*Wine Reception*

Abstracts

Andy Clark, Professor of Philosophy, University of Edinburgh

BEING AND COMPUTING: ARE YOU YOUR BRAIN, AND IS YOUR BRAIN A COMPUTER?

There's more to minds than brains alone. Biologically evolved intelligence makes the most of brain, body, and world. This talk looks at the resulting complexity, and highlights some of the unexpected advantages of solutions that span multiple levels of organization (neural, bodily, worldly) and that take shape at multiple time-scales (evolution, development, learning). I end by asking how we should conceive of the role of the biological brain in these complex ecologies.

Frances Egan, Professor of Philosophy, Rutgers University

UNDERSTANDING COMPUTATIONAL MODELS OF MIND

The so-called computational theory of mind claims that mental processes are computational processes. After explaining how this claim should be understood, I consider the popular view that it is just a metaphor to speak of the mind as a computer. I argue that this view is mistaken: computationalism, if true, implies that the human mind is a computer in the same sense that iPads and hand calculators are computers. I then consider how computational explanations of cognitive processes are constrained by our commonsense conception of ourselves, and how computationalism puts pressure on the very conception of mind.

Anil Seth, Professor of Cognitive and Computational Neuroscience, University of Sussex

MIND THE BODY

The brain is not just a machine to think with, and the body is not just a machine to move with. We are "beast machines" with our intelligence, and our conscious experiences, being deeply rooted in our specific physiology. In particular, the experience of embodied selfhood rests critically on how the brain perceives its body, both from the outside (exteroception), and from the inside (interoception). In this talk I will develop a view of embodied selfhood through the lens of 'predictive processing', by which perception is construed as a process of (active) inference on the causes of sensory signals. I will introduce a model of interoceptive inference which says that subjective feeling states (emotions) arise from actively inferred predictive models of the causes of interoceptive signals. The model also predicts that embodied selfhood is grounded in active inference of those signals "most likely to be me" across interoceptive and exteroceptive domains. I will present some recent evidence illustrating this view, including a novel version of the 'rubber hand illusion' incorporating visual feedback of heartbeat signals via augmented reality. These ideas and findings guard against naïve computational theories of mind, brain, and consciousness, which push the body into the background. At the same time they also bring new relevance to some old ideas in AI that are typically associated with cybernetics and predictive control.

Murray Shanahan, Professor of Cognitive Robotics, Imperial College London

ARE WE ON THE ROAD TO HUMAN-LEVEL ARTIFICIAL INTELLIGENCE?

Recent progress in fields such as machine learning and computer vision has revived interest in the prospect of so-called strong artificial intelligence, or human-level artificial general intelligence. But the field of AI has experienced several such periods of excitement in the past. Although each of those periods produced useful technology, none of them seemed to bring us much closer to the goal of human-level AI. Is there any reason to think things are different this time? In this talk I will overview some of the current technological trends in AI, and ask what else might be needed to take us all the way to human-level intelligence. I will conclude with some speculations about the implications for human society if we were to succeed.

Aaron Sloman, Honorary Professor, University of Birmingham

DISEMBODIED MOTIONLESS INTELLIGENCE

This is about AI as science (explanation and modelling), not engineering. In the early days of AI several prominent researchers grossly underestimated the difficulties of the tasks, and their failed predictions about timescales were erroneously interpreted as failures of AI, and computational theories of mind. That provoked a stream of rival fashions claiming to have the answer, e.g. based on neural nets, evolutionary computation, "behaviour-based" design, subsumption architectures, embodiment, enactivism, situated cognition, morphological computation, dynamical systems, autopoesis, Bayesian learning, deep learning, use of "Big Data", (not all in that order), and probably others. Some of the ideas had already been thought of long before suitable computers were available and were discussed by Marvin Minsky in "Steps Toward Artificial Intelligence" in 1962 -- half a century ago. All those fashions ignore complexities in what needs to be explained, so that each tries to explain too little -- focusing only on problems for which the fashion is suited. My presentation will identify a class of sub-problems involving "offline intelligence" uses of intelligence without performing any visible actions, though the long term consequences for future actions may be profound. Examples include making mathematical discoveries, such as the discoveries leading to

Euclid's Elements, perhaps the most important book ever written. Others include wondering whether, wondering why, trying to remember, composing a poem, enjoying a poem, designing a building, debugging a design, trying to understand why someone is late, and many more. There are also unconscious examples, such as deriving grammar based language understanding from pattern based language understanding, then extending the grammar-based version to handle exceptions: a difficult software engineering feat performed unconsciously by children. But I think there are many more, including toddler theorems discovered and used, without anyone noticing (though Piaget noticed some). A key methodological step is adopting virtual-machine functionalism as a key part of the research framework. This is part of the very long term Turing-inspired Meta-Morphogenesis project, which needs much help.

Sethu Vijayakumar, Professor of Robotics, University of Edinburgh

ROBOTS THAT LEARN: THE FUTURE OF MAN OR THE 'MAN OF THE FUTURE'?

What is your science fiction fantasy: A personal robot butler doing your household chores autonomously or going to the surgeon to buy a new bionic part to augment your body's capabilities? Today, robots are increasingly making the journey from industry floors to our homes — examples include self-driving vehicles (on road and underwater), prosthetic devices, surgical assistants and service robots for drilling, mining and cleaning. Professor Sethu Vijayakumar will explore the scientific challenges in the exciting domain of 'interactive, autonomous robotics' and show some of the cutting edge research that is aimed at making robots as versatile, safe, reactive and adaptive as us humans. He will illustrate the spills and thrills of working with some of the world's most sophisticated anthropomorphic robots like the SARCOS DB, HONDA ASIMO, the iLIMB Hand, the KUKA robot arm and the NAO footballers through interactive demonstrations and videos. In the field of robotics, the science fiction of truly embodied artificial intelligence has never been this close to science fact and he will illustrate how the tremendous progress in data driven machine learning is fuelling the progress.

Scientific Organization

Mattia Gallotti (University of London) and Sanjay Modgil (King's College London)

Many thanks to the *Institute for Advanced Studies in the Humanities* (IASH) and to the School of Philosophy, Psychology and Language Sciences at the University of Edinburgh for their generosity in hosting the academic workshop and for organizational support.