

ABSTRACTS

Josh Bongard

The Effects of Cartesian Gravity on Robots

Although great strides have been made recently in robotics, “Cartesian gravity”---the unceasing and hidden assumption that body and brain are separate---continues to distort the field’s research agenda. I will show various examples of this, especially the fact that roboticists tend to assume that the bodies of their robots do not change, only their neural controllers do. I will then describe how my group has attempted to combat Cartesian gravity by assuming instead that robot body plans, as well as their neural controllers, can adapt over different time scales. I will show how this expanded research program has enabled us to address new questions in robotics and formulate new solutions to old problems. I will touch on how this may pave the way toward safe and general AI, and its implications for biology, cognitive science, and the philosophy of mind.

Emma Borg

Understanding Agency in Others and Ourselves

The idea that we are intentional agents, for the most part acting rationally in light of our mental states, possesses strong intuitive appeal. Following on from this, it seems plausible to think that we make good sense of one another’s behaviour by reasoning about what we take them to believe, desire, etc. However, this model is philosophically tendentious and recently some new challenges to our prima facie status as practical agents have emerged. This talk explores two such challenges. First, work in comparative human/animal research which suggests that predicting and explaining another subject’s actions may standardly be a matter of smart behaviour reading alone, rather than involving any attribution of mental states. Second, work in social psychology which seems to show that our own actions are regularly subject to significant influence from sub-conscious, irrational factors. I seek to reject both these new challenges to the intuitive view, arguing that an appropriately undemanding account of reason-responsive action still provides the most attractive account of how we gain access to, and act within, our social world.

Lisa Bortolotti

When Irrationality Supports Agency: the Case of Confabulation and Optimism

In my talk I introduce the case of confabulatory explanations of our choices and overly optimistic beliefs about our talents and future. Confabulations and optimistic biases result in epistemically irrational beliefs, that is, beliefs about ourselves and the world surrounding us that are not supported by the evidence available to us and that are not responsive to counterevidence. Notwithstanding their irrationality, and maybe in virtue of it, such beliefs have been found to enhance wellbeing and even performance in some circumstances. I will argue that confabulatory explanations and overly optimistic beliefs support our sense of ourselves as competent, efficacious, and largely coherent agents. Although our sense of competence, efficacy and coherence is largely illusory, it does make a positive contribution to our capacity to persevere in the pursuit of our goals, especially after setbacks. My conclusion will be that, when examining what makes typical humans successful agents, careful consideration should be given to the role of epistemically irrational beliefs. This may have interesting implications for growing agency in artificially intelligent systems.

Daniel Dennett and Keith Frankish

Building Up to Consciousness and Autonomy

It is tempting to suppose that there is a “headquarters” in the brain, a Cartesian Theatre, where conscious experience is generated and autonomy is exercised. After all, how could there be consciousness unless there is a place and a time at which it occurs? But who is the boss, the controller, the theatre goer, who witnesses the stream of consciousness and decides how to react? The task for cognitive science is precisely to explain how there can be consciousness and autonomy without a headquarters, while at the same time explaining why the headquarters

image is such a tempting one for conscious, autonomous beings like us.

In this project, we address this problem from a new angle. We consider a military drone, remotely operated by a human. Here there really is a headquarters, an autonomous, conscious pilot. We then imagine how we might remove the human pilot and build all of their relevant competencies into the drone itself, making it autonomous and conscious. One approach would be to build subsystems that would implement each of the pilot's competences artificially. This top-down approach, adopted by traditional AI, is still employed by many philosophers thinking about the mind. But it is very hard to analyse, let alone implement, human competences in this way.

Here, we take a different approach, inspired by the work of the cyberneticist and roboticist Valentino Braitenberg, who showed how a simple wheeled robot can produce remarkably lifelike behaviours, indicative of emotions, knowledge, and decision making. Having ejected the pilot from his drone cockpit, we imagine building increasingly sophisticated competences into the drone itself, in a bottom-up fashion, driven by operational needs. In this way, with engineering, tinkering, and tricks, we aim to build up to consciousness and autonomy.

Cecilia Heyes

Growing Gadgets

Cognitive gadgets are distinctively human 'thinking tools' – such as mindreading, imitation and language – that have been fashioned by cultural evolution. They are built in infancy and childhood by and from old, genetically inherited parts, but social interaction is the driver of gadget development, and cultural selection is what has made gadgets good at what they do. Using gaze tracking, empathy and imitation as examples, I will suggest that cognitive gadgets, although products of cultural evolution, can inspire intelligent design.

Murray Shanahan

Is General Intelligence Possible Without Consciousness?

A feature of human intelligence is its generality, the capacity to adapt to a wide range of tasks and problems. By contrast, even the most sophisticated of today's artificial intelligence systems are highly specialised. If the field of AI is to achieve generality, and to approach human-level intelligence, will it necessarily end up building machines with consciousness? In this talk, I argue that consciousness is a multi-faceted concept. Some psychological attributes we associate with consciousness in humans and other animals are indeed likely to be required for general intelligence, while others, including the capacity for suffering, are not.