Information-Theoretic Bounded Rationality*

Pedro Ortega

Abstract: In this talk I provide an overview of information-theoretic bounded-rationality for planning in sequential decision problems. I show how to ground the theory on a stochastic computation model for large-scale choice spaces and then derive the free energy functional as the associated variational principle for characterising bounded-rational decisions. These decision processes have three important properties: they trade off utility and decision complexity; they give rise to an equivalence class of behaviourally indistinguishable decision problems; and they possess natural stochastic choice algorithms. I will discuss a general class of bounded-rational sequential planning problems that encompasses some well-known classical planning algorithms as limit cases (such as Expectimax and Minimax), as well as trust- and risk-sensitive planning. Finally, I will point out formal connections to Bayesian inference and to regret theory. [This is joint work with Daniel A. Braun, Kee-Eung Kim, Daniel D. Lee, and Naftali Tishby (in alphabetical order).]

Bio: Pedro A. Ortega is postdoctoral researcher at the GRASP Robotics Lab, University of Pennsylvania, working with Daniel D. Lee. His research focuses on the mathematical foundations of intelligent behaviour. Most of his past work applies information-theoretic and statistical mechanical ideas to sequential decision-making, leading to contributions in novel bounded rationality models and recasting adaptive control as a causal inference problem. He obtained his PhD in Engineering from the University of Cambridge (Zoubin Ghahramani), and he has been a post-doctoral fellow at the Department of Engineering in Cambridge (Simon Godsill), at the Max Planck Institute for Biological Cybernetics/Intelligent Systems (Daniel Braun) and at the Hebrew University in Jerusalem (Naftali Tishby).

*Internal Talk, Gatsby Unit, September 8, 2015.