CBS Forum

Entropy Based AI: Estimation in Natural Sequences and Information Topology

Dr. Andrew D. Back

School of Information technology and Electrical Engineering, University of Queensland

June 24 (Monday), 2019

13:30-14:30

1F Seminar Room, CBS East Bldg.

Abstract

A significant problem in AI which is largely unsolved, is how to determine the intent and potentially the meaning behind events. Machine learning algorithms are now capable of driving cars, yet cannot solve simple tasks which humans find trivial, such as knowing the difference between what is said is not really what is meant. For example, today's machine learning systems can't tell the difference between a true 'yes' said with enthusiasm or a hesitant 'yes' indicating perhaps just the opposite. We present new results in the field of information theory, particularly entropy based methods applied to natural sequences which are aimed at solving these problems.

While entropy is widely used, a previously unanswered question is how many samples are needed to compute entropy? We present new results which give a method for estimating how many samples are required using a modified Zipf-Mandelbrot law and the Dvoretzky-Kiefer-Wolfowitz inequality with a given confidence level. Secondly, we present an efficient new entropy estimation algorithm for natural sequences which is shown to yield accurate results with a very small number of samples. We then show how degenerative entropic solutions can occur when using current entropy measures are applied to the field of dynamic sequences such as language or dynamical behaviour due to. To overcome this, we introduce Transitive Entropy which can be understood in terms of Ricci Curvature Tensors. This provides a natural pathway to a new entropy-based approach to AI, which we term Information Topology, extending Amari's information geometry with topological data analysis applied to Riemannian manifolds. We give some examples, demonstrating the proposed entropy measure applied to classifying patient dementia by analysing their speech.

Host: **Hiro. Nakahara** Lab for Integrated Theoretical Neuroscience