Embedding Spaces for Decision Making

A variety of domains, including images, text, and brain measures, can be captured in embedding spaces. For example, word embedding models place each word at some point in a high-dimensional space with the relative positions of words conveying similarity. In this talk, I will consider what comparing embedding spaces to one another can tell us about cognition and its brain basis. First, I will cover model-based neuroscience research that compares model and brain representations. I will discuss limitations of existing model-based approaches, including deep learning accounts of the ventral visual stream, and suggest an alternative approach to linking models and brain measures that assesses causal efficacy within the overall computation rather than simply shared variance between embedding spaces, which can be misleading. Second, I will discuss how embedding spaces derived from large-scale studies of human behaviour can help us evaluate models. One conclusion is that better performing models are not necessarily better models of humans. In the final part of the talk, I will consider how people rely on multiple embedding spaces (akin to memory systems) when making open-ended decisions, such as deciding what to add next to their online shopping cart. Overall, these results indicate the value of embedding spaces for developing and evaluating models of mind and brain at scale.