Speaker: Hugo Chao CUI

Title: **A Phase Transition between Positional and Semantic Learning in a Solvable Model of Dot-Product Attention**

Abstract: We investigate how a dot-product attention layer learns a positional attention matrix (with tokens attending to each other based on their respective positions) and a semantic attention matrix (with tokens attending to each other based on their meaning). For an algorithmic task, we experimentally show how the same simple architecture can learn to implement a solution using either the positional or semantic mechanism. On the theoretical side, we study the learning of a non-linear self-attention layer with trainable tied and low-rank query and key matrices. In the asymptotic limit of high-dimensional data and a comparably large number of training samples, we provide a closed-form characterization of the global minimum of the non-convex empirical loss landscape. We show that this minimum corresponds to either a positional or a semantic mechanism and evidence an emergent phase transition from the former to the latter with increasing sample complexity. Finally, we compare the dot-product attention layer to linear positional baseline, and show that it outperforms the latter using the semantic mechanism provided it has access to sufficient data.