NICHOLAS BARNFIELD, Harvard University

Phase Transitions in Linear Associative Memory

Associative memory, the ability to store and retrieve relationships between objects, has recently attracted attention as a model of factual recall in large language models. A central question is how large a memory must be to retrieve a given number of associations under a prescribed criterion. In this talk, we study it in the simple but already nontrivial setting of linear associative memory with Gaussian embeddings. For top-1 retrieval, where each key must rank its target first, we prove a sharp threshold in memory size separating success from impossibility. Where top-1 retrieval fails, we introduce the Tail-Average Margin, a convex criterion that compares the target's score with the average of its strongest competitors and certifies top-k retrieval. We characterize the learned memory exactly in the high-dimensional limit, revealing a sharp phase transition. Our analysis also contributes a coupled leave-one-out method for matrix-valued empirical risk problems in which each sample enters many dependent comparisons, a tool of independent interest.

Joint work with Juno Kim, Eshaan Nichani, Jason D. Lee and Yue M. Lu.