GSAS Workshops
Language Universals

Speaker: Dr. Yoon Kim (MIT)

Date: April 15, 2022 - 12.00 1.30

Location: Sever Hall 214

Title: Corpus-driven discovery of linguistic structures with neural networks

Abstract: Natural language has inherent structure. Words compose with one another to form hierarchical structures to convey meaning. While these compositional structures (such as parse trees) are crucial for mediating human language understanding, they are unobserved during human language acquisition. Yet, human learners have little trouble acquiring the syntax of their native language without explicit supervision. This has motivated the classic task of grammar induction, (i.e., data-driven discovery of syntactic structure from raw text), which has proven to be empirically difficult for artificial language learners. In this talk, I show how recent advances in model parameterization and inference can lead to improved computational tools for discovering syntactic structure from raw text. I also show how such techniques can be applied to induce bilingual synchronous grammars for tasks such as machine translation.