

An Exploration of Composite Language Modeling for Speech Recognition

Xiaolin Xie

Dr. Yunxin Zhao, Thesis Supervisor

Abstract

Language models are one of the most critical knowledge sources of automatic speech recognition (ASR) systems. In the past decades, many language models have been developed, and some have been proved useful and successful in speech recognition systems. However, almost all language models only capture one or two aspects of natural language. This study aims to investigate the effects of a syntactic, semantic, and lexical language model on speech recognition. In this study, we refer this language model as the composite language model (CLM). The parameters of the CLM in our study are distributed among hundreds of computer nodes in a supercomputer because they are too large to be stored in just one computer node. A distributed application has been developed to implement two speech rescoring techniques by using the CLM: lattice rescoring and confusion network rescoring. Experiments on a Wall Street Journal task have shown that using CLM to rescore word lattices and confusion networks have led to improvements in word accuracy over the commonly used trigram language model, with the latter offering a larger performance gain.