



THE UNIVERSITY OF
CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

YI-HUNG LIU

Department of Statistics
The University of Chicago

Causal Relations and Feature Extraction in Contextualized Sentence
Representation

WEDNESDAY, April 29, 2020, at 1:00 PM

ABSTRACT

Pre-trained contextualized sentence representations have achieved phenomenal success in natural language understanding tasks. On the other hand, causality is an unsettled topic discussed in different fields last more than 2000 years. Causal relation extraction in natural language is, therefore, a field that requires scientific examination. As people are having progress on it using pre-trained embeddings, it is important to understand what information did they embedded and how do they generate predictions. In this paper, we first confirm the pre-trained embeddings themselves can detect causal relations and the importance of context in implicit causal relations. Thereafter, we try to detect various components of causality derived from theories in psychology, philosophy, and computer science. The result shows causal power is more important in implicit sentences while probability differences are enough for explicit relations. Additionally, the predicted causes by BERT are correlated with causal power and CPMI that could potentially be interpreted as an internal generation process of BERT.