



Comparing Natural Language Embeddings for Libc Functions as Rich Labels

Bachelor defense

Ruben Triwari

Ludwig Maximilian University Munich

19, February 2025

Outline

Motivation & Research Objective

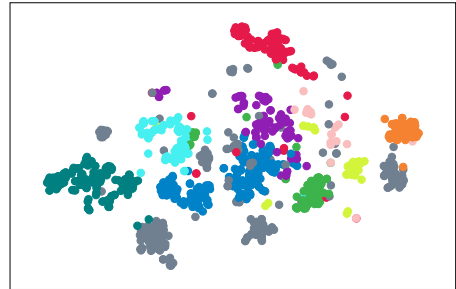
State Of The Art

Methodology

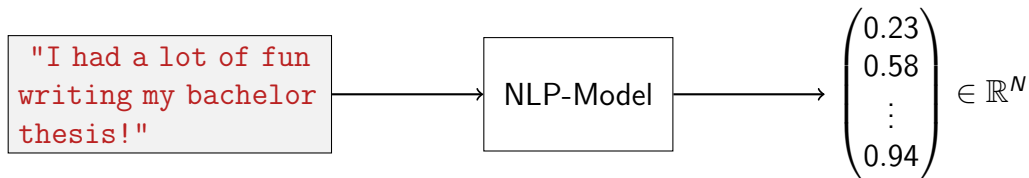
Results

Limitations

Conclusion & Future Work

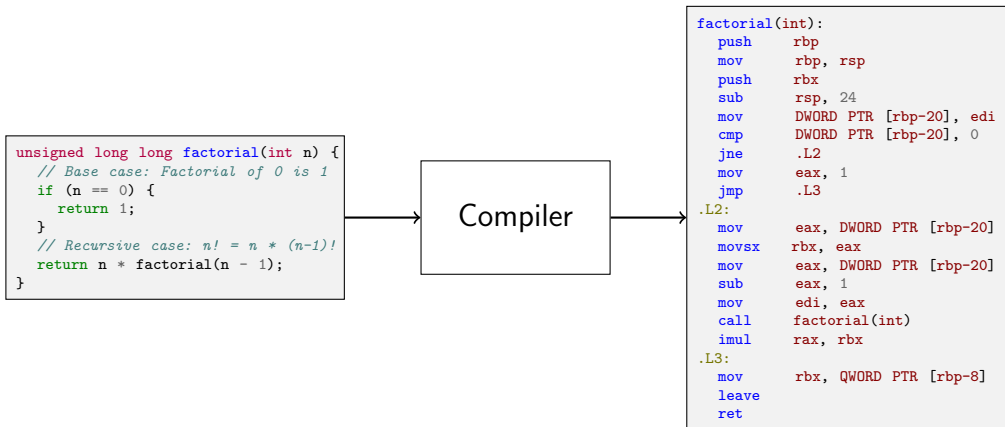


Motivation



- ↪ Encoding natural language was a huge factor in recent nlp advancements
- ↪ Information described as a vector can be used in many downstream task
- ↪ That motivates encoding binary code and describing them as a vector
- ↪ That motivates using NLP tools to encode binary code

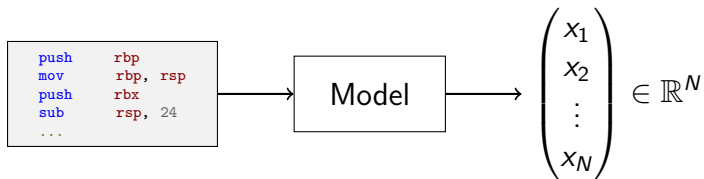
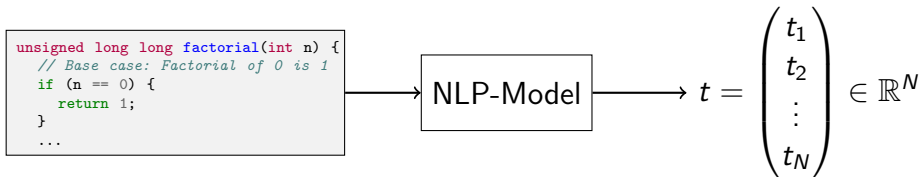
Motivation



⇒ Compiler removes important information in natural language

Motivation

Idea:



Research Objective

Simply Typed Lambda Calculus – Motivation

Curry-Howard Correspondance

Denotational Semantics

HALLO

HALLO

Discussion