Quantum Natural Language Processing and Machine Learning

Natural Language Processing-Lab

Indiana University at Bloomington

Natural Language Processing

Challenges:

- Meaning and Understanding: SOTA NLP and AI technologies lack understanding, fail to process semantic and meaning properties of language and other modi of communication
- Machine Learning: Language Model Optimization computation are costly and time consuming

Quantum Expectations:

- Optimization algorithms for ML algorithms
  - Traditional NLP technologies (e.g., symbolic, probabilistic, neural)
  - Large Language Models and Generative AI
- Machine Learning: Language Model Optimization computation are costly and time consuming

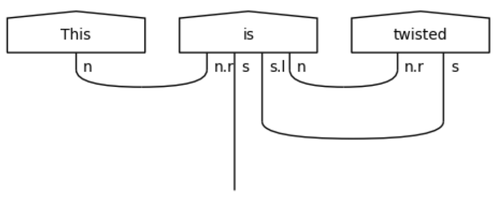
Our Goals

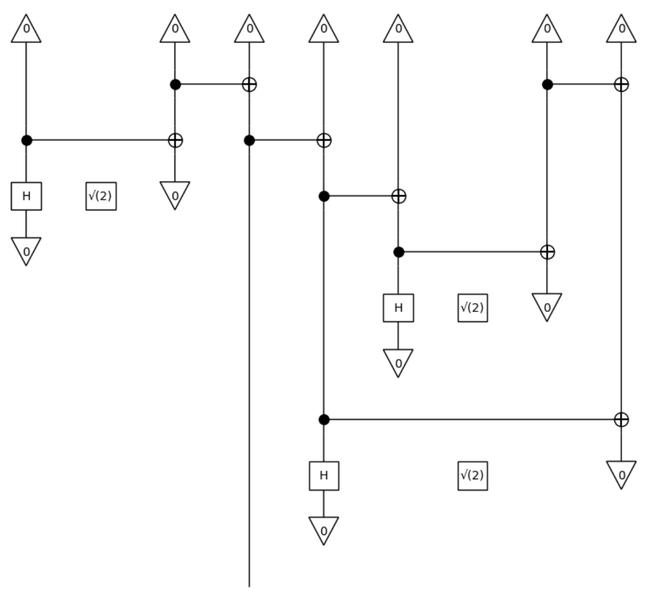
We are aimed to establish a local infrastructure and gather a collaborative team of colleagues and graduate students dedicated to exploring QNLP and ML.

- **Algorithms** We will innovate and evaluate NLP and ML algorithms utilizing current quantum computing platforms.
- **Optimization** We aim to enhance the efficiency and effectiveness of both training and testing cycles within optimization–focused ML frameworks.
- **Improvements** We plan to utilize the probabilistic models from quantum mechanics to overcome the shortcomings in classical NLP, especially at semantic and common–sense levels.

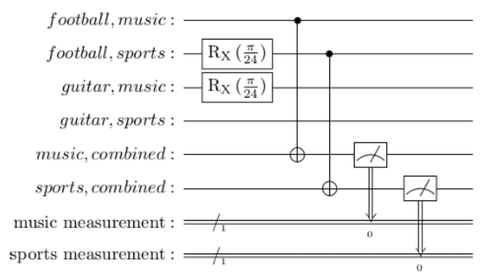
Technologies and Platforms

- **NLP Platform:** `lambeq` - Quantum NLP and ML environment.
  - NLP environment: Linguistic Representations → Quantum Circuits





- **IBM Quantum:** `qiskit` - General computing platform.
  Text Topic classification circuit:



Quantum Natural Language Processing

- **Powerful** The largest quantum computer in the world, created by Atom Computing, has surpassed the 1000–qubit mark. The potential of existing quantum computers makes them capable for data–hungry fields, such as NLP.
- **Refined Theoretical Foundation** After nearly a century of development, quantum mechanics and category theory have evolved into a unified language of science. Quantum mechanics and category theory are naturally suited for handling large data and non-local correlations, making them quite suitable for the field of natural language processing
- **Accessible Tools** There are a lot of open–source tools for quantum computing and QNLP, like `qiskit` from IBM and `lambeq` from Quantinuum. All these tools contribute to making quantum computing more accessible to a broader audience, from beginners to advanced learners, and supports the growing community.

Applications

- **Medical** QNLP can be potentially used to better understand and analyze patient data, optimizing communications between medical care providers. This could improve decision - making and diagnosis accuracy.
- **Linguistics and Syntax** Quantum solutions offer new tools for reasoning with probabilities or creating hybrid neuro-symbolic solutions for serious NLP problems.
- **Intelligence and National Security** Quantum computers can make analysis of language–based information more precise, especially in communications using spoken language and text highly relevant to Intelligence and National Security.
- **Translation** By employing the mathematical foundation of quantum theory, QNLP can represent linguistic aspects in a manner that's potentially more nuanced and sophisticated than traditional NLP.

Connections and Collaborations

Local Quantum Centers:

- QSEc – Quantum Science and Engineering Center (IU)
- Center for Quantum Technologies (CQT) – IU, Purdue, Notre Dame
  - Center for Quantum Technologies (CQT) – NSF Site

The NLP Lab has worked on Quantum NLP and advanced NLP and ML approaches for many years, bringing together scientists and students from various disciplines.

- Prof. Jerome Busemeyer (IU)

Besides colleagues on the Indiana University Bloomington campus, we are in touch with:

- Prof. Mehrnoosh Sadrzadeh (University College London)

Team

Prof. Damir Cavar
Prof. Larry Moss
Chi Zhang
Shane Sparks
Jose Benavides

Natural Language Processing Lab

The NLP-Lab (`https://nlp-lab.org/`) projects and results: