Statement of Purpose

Since high school, where I thrived in algorithm competitions, my passion for technology has evolved through hands-on experiences in programming, machine learning, and natural language processing (NLP). My undergraduate research, particularly in NLP and causal inference, revealed the potential of language models in creating context-aware, robust and efficient computing systems**.** For my PhD, I am committed to developing *systems that merge data, symbolic reasoning, and NLP*, aiming to revolutionize how technology adapts to specific needs while minimizing costs. To achieve this goal, I have prepared myself through 4 research projects in building integrated language and logic systems, coupled with my deep understanding of language structures and pragmatic problem-solving approach.

Structural text planning and generation

Since Spring 2023, I have been collaborating with Kevin Yang and Prof. Dan Klein at **Berkeley AI Research**, focusing on integrating language and logic for coherent long-form storytelling. Our aim is to address **factual contradictions in narrative structures**, a challenge I was initially drawn to due to the potential of structural text generation. I quickly recognized the existing limitations – the independent generation of different narrative parts often led to incoherence and factual conflicts. To tackle this, my approach has been to **manage an external world state**, **breaking down complex narratives into atomic facts** and **designing a data structure for dynamic world states**. This involves ensuring **time-bound fact validity** and **detecting contradictions within the story**.

Through this project, I've learned that LLMs need additional support, be it external modules or internal training, to align with human logic and the world effectively. This realization has not only bolstered my determination to pursue hybrid logic-language systems but also equipped me with essential technical and methodological skills. As I prepare for our submission to **ARR 2023**, I am excited to further both my personal growth and contribute to the broader field of AI research.

In-depth understanding about language and causality

Under the mentorship of Zhijing Jin, Prof. Mrinmaya Sachan, and Prof. Rada Mihalcea, I have delved into the complex interplay between causality and NLP. One of our focus has been on unraveling the structures within natural language data, particularly examining causal and anticausal properties in sentiment analysis. We explored how causal directions in prompts influence the behavior of Large Language Models (LLMs), using these insights to enhance causal discovery and performance. Our work culminated in **my first-author paper at UAI** 2022 workshop. Additionally, we investigated language's hidden causal mechanisms through discourse and psychological analysis.

One of my other projects with **Kernel-based Conditional Independence Test statistics** as a loss penalty, I grappled with biases and robustness in causality data training, pushing the boundaries of my understanding of high-dimensional data. Additionally, my involvement in various projects, from transforming symbolic language in our correlation-to-causation inference data construction (ICLR 2024 under review) to examining translationese for causal and anti-causal learning (NAACL 2022 oral)**,** and contributing to data structuring in **AIScholar** and **LogicalFallacy** (EMNLP 2022), further deepened my comprehension of natural language data structures.

Recognizing the vital role of data in mixed systems through those projects, my future research will also focus on understanding data characteristics more profoundly. I aim to **mine and construct better data representations** and develop corresponding model systems for complex tasks. My approach will involve three key methods: **1)** mining **additional supervisory signals** like causality and discourse structure from natural texts, **2)** generating **synthetic data** using symbolic or mixed ways, **3)** collecting data from **simulation** among multiple LLM agents.

Other Technology and Leadership Experience

In my career, I've also cultivated a blend of technical expertise and leadership skills crucial for my PhD in AI. On the technical front, my work at **Megvii** on AutoML systems and collaboration with **Zhejiang University** on Large Language Models directly feeds into my research focus on language and logic integration. Contributing to the **causal-learn** project and developing the **causalnlp-toolbox** have deepened my practical AI knowledge, showcasing my ability to impact the field. In terms of leadership, leading my team to **gold medals** and **entering the World Final** in the ICPC demonstrated my problem-solving and team coordination skills, vital for research collaboration. Furthermore, guiding HKU Robotics+AI and founding the AI4Good Community underscored my capacity for mentorship and community building, essential qualities for a future in academic or research leadership.

Research & Career Plans

I aspire to explore on a new computing paradigm about intelligence that **integrates language and logic** for my PhD research, focusing on two main areas:

1) Innovating in **structural text planning and generation**, such as long-form storytelling, exploring algorithms that merge language and logic and enhancing human interaction with text.
2) I plan to advance intelligence emergence **through a data-driven approach**, focusing on extracting richer supervisory signals from texts and experimenting with synthetic data and multi-agent interactions.

My long-term career goal is to become an established **independent researcher** or an **academic professor**, with a focus on developing algorithms that contribute to the achievement of general intelligence. My aim is to apply these advancements to tackle more complex, real-world problems, making such technology more accessible and beneficial for humanity. In addition to my research endeavors, I take great pleasure in collaboration and mentorship, working alongside others in the field to expand our collective impact and influence.