Research Statement

Tianqing Fang

Knowledge, including but not limited to skills, factual and commonsense knowledge, is fundamental to human cognition and of vital importance in developing human-level language models. Despite being knowledgeable on many aspects by scaling up to billions of parameters, Large Language Models (LLMs) are still criticized for *hallucination* and a limited capacity of *reasoning* with knowledge. To this end, there is an urgent need for developing both knowledge-enhanced LLMs and robust reasoning paradigms. My research vision is thus centered on building **robust knowledge-enhanced Natural Language Processing (NLP) systems beyond scaling up**.

My current research direction can be roughly divided into two scopes:

1. **Complex knowledge acquisition**: Acquisition of commonsense knowledge, knowledge conflicts, and complex logical queries with information extraction, crowdsourcing, and LLMs.

2. **Knowledge injection and reasoning on LLMs**: (Lightweight) injection of knowledge, including constrained decoding, retrieval-augmented, and information-theoretic injections. Elicit the complex reasoning ability of LLMs using internal and external knowledge, particularly on complex structured, counterfactual, long-tail, and long-context knowledge.

1 Complex Knowledge Acquisition

I have studied knowledge acquisition of commonsense knowledge with complicated structures, such as through graph reasoning, abstraction, and complex logical queries. I also study knowledge conflicts, where the actual context contradicts the parametric knowledge in the language models.

I mainly studied **Complex Commonsense Knowledge** mining for inferential commonsense knowledge of daily events and entities. Based on the commonsense paradigm of ATOMIC (e.g., *PersonX repels PersonY's attack*, then PersonX is seen as *brave*), I used graph-enhanced BERT-based models to automatically convert information-extracted discourse relations to commonsense knowledge [7]. I built several follow-up works, including benchmarking such a commonsense knowledge base population process [5] and using semi-supervised learning for knowledge acquisition [3]. In addition, I'm interested in abstraction [9], indicating a higher level of knowledge (e.g., conceptualizing *watching football games* to *relaxing activity* for further inferences). I further built a semi-supervised abstraction/instantiation acquisition system [11] and applied it for downstream zero-shot commonsense question answering tasks [10], showing state-of-the-art zero-shot reasoning performance, even better than ChatGPT. However, despite understanding those one-hop inferences, LLMs still struggle to reason about complicated structures, such as logical queries on knowledge graphs [2]. I sampled complex first-order logic queries from ATOMIC and verbalized them to narratives to derive both a harder commonsense evaluation set and better reasoning supervision signals for LLMs. Experiments show that complex queries equip language models with better reasoning ability on both complex reasoning and original one-hop reasoning tasks.

I also worked on **Knowledge Conflicts**, which refer to the cases where the parametric knowledge from the language model contradicts with the actual context. I used the idea of *reporting bias* to calculate knowledge conflict statistics to mine temporal knowledge conflicts of various types [6]. I developed counterfactual data augmentation that can be used for both fine-tuning and in-context learning to mitigate such knowledge conflicts.

2 Knowledge Injection and Reasoning

I study injecting knowledge into LLMs without scaling them up from three perspectives. First, lightweight **Knowledge Constrained Decoding**. In the framework KCTS [1], a Monte-Carlo Tree Search module is applied to estimate the future groundness to the intended knowledge, and a novel token-level hallucination detection method is used by constructing synthetic supervision signals by setting a hallucination inflection point at a token level. KCTS is plug-and-play to LLMs and only requires fine-tuning on 0.21% of additional training weights while significantly improving factuality. Second, I study knowledge fusion in a **Data Augmentation** perspective to denoise the knowledge [8] with training dynamics, i.e., a clean distiller and a self-regularization module, and to use counterfactual data augmentation for both fine-tuning and in-context learning to mitigate knowledge conflicts [6]. Last but not least, I used **Graph Reasoning Networks** to fuse knowledge subgraphs to language models. I studied integrating supporting sub-graphs of knowledge to reasoning modules, including using GraphSAGE to aggregate the knowledge in ATOMIC to perform commonsense knowledge base population [5] and fusing embeddings of knowledge graphs to each layer of transformers to perform better dialogue generation [4].

3 Future Works

In the future, I will pursue my lifelong research goal to enable machines with the ability of human cognition and reasoning that leads to AGI, especially in the following directions:

Robust Complex Reasoning. First, though current LLMs possess knowledge about numerous one-hop scenarios, they are not robust in handling complex graph-structured reasoning tasks, even with the chain-of-thought. In my future research, I aim to identify and integrate complex knowledge and reasoning capabilities into language models, enabling them to effectively address complex reasoning challenges prevalent in real-world applications, such as planning. Second, the Zipf's Law indicates the existence of substantial amount of long-tail knowledge, which is crucial for achieving human-level AI systems while intrinsically infeasible for LLMs to learn by only scaling up. I'm interested in acquiring long-tail knowledge based on reporting bias [6] and abstraction-instantiation relations [9]. This knowledge can be utilized for evaluating or enhancing long-tail reasoning ability through approaches such as data augmentation [10] and counterfactual in-context learning [6].

Effective and Efficient Knowledge Injection. I aim to focus on two main research problems regarding knowledge injection and reasoning. First, inject knowledge into LLMs without catastrophic forgetting and high-cost fine-tuning. To accomplish this, I will focus on inference-time algorithms, including guided decoding [1], retrieval-augmentation, and information-theoretic integration of in-context examples. Second, reason with entailment knowledge inspired human cognition (*k-line* theory by Marvin Minsky). Certain long-tail scenarios, particularly in commonsense reasoning, can be abstracted/entailed to a higher level for reasoning. For instance, if an LLM lacks knowledge about *ghijak*, conceptualizing it as an *instrument* enables effective resolution of reasoning scenarios related to instruments without requiring scaling up or fine-tuning the LLM.

Real-world Application of Knowledge Reasoning. Most existing (commonsense) reasoning benchmarks are limited to toy datasets designed to challenge language models rather than addressing practical tasks. To bridge this gap, I aim to integrate various forms of knowledge, including scientific knowledge, moral and cultural knowledge, and long-context knowledge from textbooks, into language models as curriculums, guiding their utilization in real-world applications involving interactions with human agents, procedural planning, and science discovery. My goal is to create a unified pipeline that encompasses salient information extraction or direct knowledge generation from (long) contexts, leading to the formation of a logical reasoning graph. This graph will be processed using the graph-aware reasoning algorithms I have developed on top of LLMs.

References

[1] Choi, S., Fang, T., Wang, Z., and Song, Y. KCTS: knowledge-constrained tree search decoding with token-level hallucination detection. *EMNLP* (2023).

[2] Fang, T., and Bosselut, A. Complex commonsense reasoning on commonsense knowledge graphs. *ongoing work (to be released in Dec 2023)* (2023).

[3] Fang, T., Do, Q. V., Zhang, H., Song, Y., Wong, G. Y., and See, S. Pseudoreasoner: Leveraging pseudo labels for commonsense knowledge base population. *Findings of EMNLP* (2022).

[4] Fang, T., Pan, H., Zhang, H., Song, Y., Xu, K., and Yu, D. Do boat and ocean suggest beach? dialogue summarization with external knowledge. In *AKBC* (2021).

[5] Fang, T., Wang, W., Choi, S., Hao, S., Zhang, H., Song, Y., and He, B. Benchmarking commonsense knowledge base population with an effective evaluation dataset. In *EMNLP* (2021).

[6] Fang, T., Wang, Z., Zhang, H., Song, Y., and Chen, M. Getting sick after seeing a doctor? diagnosing and mitigating knowledge conflicts in event temporal reasoning. *arxiv* (2023).

[7] Fang, T., Zhang, H., Wang, W., Song, Y., and He, B. DISCOS: bridging the gap between discourse knowledge and commonsense knowledge. In *WWW '21: The Web Conference* (2021).

[8] Fang, T., Zhou, W., Zhang, H., Liu, F., Chen, M., and Song, Y. On-the-fly denosing for data augmentation in natural language understanding. *arxiv* (2022).

[9] He, M., Fang, T., Wang, W., and Song, Y. Acquiring and modelling abstract commonsense knowledge via conceptualization, 2022.

[10] Wang*, W., Fang*, T., ..., Song, Y., and Bosselut, A. CAR: conceptualization-augmented reasoner for zero-shot commonsense question answering. *Findings of EMNLP* (2023).

[11] Wang*, W., Fang*, T., Xu, B., Bo, C. Y. L., Song, Y., and Chen, L. CAT: A contextualized conceptualization and instantiation framework for commonsense reasoning. *ACL* (2023).

* indicates equal contribution.