- **Independent Research & Versatility:** I have demonstrated the ability to **conduct NLP research independently and enthusiastically**, in **3** years, through **4** diverse projects: event extraction (DARPA BETTER, member), schema induction (DARPA KAIROS, leader), entity-state tracking (AI2, member & leader), and natural to symbolic reasoning (ongoing). My growth has evidently shown in **comprehensive research skills**: identify problems, navigate challenges, achieve goals, and present work. Now, I have **1** ACL publication, **2** in submissions, all first-authored; **1** ongoing research on using Symbolic Languages to imitate and assist human planning (See Figure.1).
- **Creativity:** My creativity is derived from integrating my **human cognition expertise** (**6** years) into the **NLP** research. I envision the concrete-abstract architecture unexplored in current LM designs. (See more in future work) I have developed a Natural to Symbolic Language distillation approach, aligned with human information encoding. (See more in past & ongoing work)
- **Collaboration & Leadership:** I have been extensively cooperating with Prof. Dan Roth, Prof. Chris Callison-Burch, Prof. Lyle Ungar, Prof. Heng Ji, and Prof. Martha Palmer, all of whom have provided positive feedback.

  Leading a team of 6 members on the KAIROS project, we successfully published an **ACL** demo paper within 4 months. I fostered a collaborative environment within and across the groups through weekly meetings and more. As a **TA** for the Deep Learning course, I **guided 13** students, receiving **high ratings** (≈1st out of 15 TAs). My goal was to provide all possible support to my students. I applied human learning theories to progressive materials, engaged students through insights and interactions, and offered extra personal discussions to cater to distinct needs.
- **Expectation for PhD:** I aspire to advance reasoning research through intensive study, and extensive cooperation, in the long term, to contribute to society. A professor who is dedicated to improving the world deeply attracts me.

I am passionate about building anthropomorphic robots to enhance human life, especially for the underprivileged. Specifically, I endeavor to devise **intelligent agents that emulate human understanding and reasoning** of world events. In contrast to human learning, which assimilates and accommodates information into brain schemas (Piaget, J., 1952),[3] a significant challenge with current Language Models (LMs), including the SOTA GPT-4, is their inability to automatically acquire and anchor structured knowledge in the network. This deficiency leads to unreliable reasoning and hallucinations. To alleviate it, my research **directs LMs to construct and reason with structured and symbolic representations**. I will present these efforts through event extraction, schema induction, and entity-state dynamics in Symbolic Language, in chronological order.

- **Understanding the Basics: Event Extraction**

  Event Extraction derives structured triples from unstructured text,
  e.g., "I go to a restaurant." elicits a 'Travel' event: (trigger: go, person-arg: I, place-arg: restaurant).

- **Problem-Solving:** I implemented a Roberta pipeline where triggers underwent conventional BIO tagging, while arguments leveraged a unique Question-Answering (QA) approach. This strategy promoted increased information sharing of arguments and facilitated more effective transfer learning from abundant QA datasets. However, the lack of in-domain annotation data hinders the performance. To resolve the bottleneck, **inspired by a free-form QA data augmentation work,**[6] **I introduced synthetic QA to restrictive event contexts.**[1] I implemented an answer extraction to question generation (AE-QG w/ Bert-T5) approach in reverse. Not only did it circumvent the no-answer questions making up 60% of the inquiries, but it also proved notably more effective, with our 8k synthetic data surpassing the effectiveness of 80k SQuAD data as tested on the ACE. In the next phase, I mapped the isolated events into graphs, akin to human intellect.
- **Work Impact:** my work was incorporated in BETTER and adapted in another work.[7]

- **The Bigger Picture: Event Schema Induction**

  Event schema is a graphical depiction of event knowledge, e.g., first "I go to a restaurant", followed by "I eat food".

- **Problem-Solving:** I joined this project when other groups had worked on it for years. I quickly caught up on the progress by discussing with professors and reading the previous documents. To manage the task effectively, **I began by summarizing the human reasoning patterns** from 11 annotated schemas**, then progressed to auto-generation**. This involved using GPTs to infer the details and merge the graphs iteratively. Despite their impressiveness in common sense, GPTs succumbed to intricate relations. While human annotation is precise but prohibitively onerous (≈ 1 hour/schema).[8] To strike a balance, I led the team to develop an interface supporting human-GPT interactive schema generations (≈ 15 mins/schema) (See Figure.2).[2] As a leader, my objectives were both to ensure every member feels valuable and to achieve project requirements as a sub-team in KAIROS. I navigated through accommodating **diverse interests, opinions, and levels of engagement**. In the next phase, my work on reasoning shifted from an action-centered perspective to a fine-grained entity-centered approach. It modeled events as revolving around main subjects via goal-oriented actions.
- **Work Impact:** Our demo was used by the UIUC group to collect data efficiently.

- **From Action to Entity: Causal Reasoning in Symbolic Language**
  Causality can be understood as entity-state dynamics shaped by actions,
  e.g., init: (I, at home) (I, hungry) → Action go: (I, at a restaurant) → Action eat: not (I, hungry) → goal: not (I, hungry).

- **Problem-Solving:** Previous works treated causality either grossly as temporal relations or tediously as one-condition inferences designed for rigorous logic deduction. I creatively envisioned causality as the entity-state transitions articulated in Symbolic Language (SL), like PDDL. In other words, a causal unit was viewed as entity.state before → action → entity.state after. **I designed a 3-step approach, based on human information encoding, to elicit a causal unit in SL from NL descriptions** (See Figure.3).[3] It allowed GPTs to capture often-ignored implicit entity states, e.g., shape changed, in a schema. In the next phase, I will introduce probability into each $Predicate_i$ in a state, where $P_{Predicate_i} \leq 1$ (cf. PPDDL[9]: for all $State_i$, $\sum State_i = 1$).
- **Work Impact:** the fine-grained view in the initial work advances our ongoing reasoning work (See Figure.1).

- **Gazing Ahead: A Future Toward Anthropomorphic Intelligence (AI)**
  Humans have consistently crafted more sophisticated tools. For me, the pursuit of Anthropomorphic Intelligence offers numerous encouraging paths to investigate:

- First, continually enhance event reasoning with Reinforcement Learning (RL). Similar to game scenarios, human action series are policy-based. Consequently, it is feasible to **learn a policy** from documents on latent states and observed actions, $\pi(s, a)$, **to model event reasoning**. To my knowledge, while RL prevails in games and robotics, it has not been incorporated in event causality work.
- Second, build efficient knowledge-transferring approaches. Current artificial boundaries limit the utilization of heterogeneous texts, lagging behind the seamless transitions humans make between NL and structured knowledge. I posit that **divergent texts can be jointly trained through concrete-abstract architectures**, such as Transformer encoder → GNN → Transformer decoder. To my knowledge, there are some preliminary attempts combining Transformers and GNN [10, 11, 12] but none of them give a systematic view on the design of architecture.
- Third, explore multimodal learning. Studies in neural and cognitive science suggest that distinct senses, e.g., vision, language, etc., are linked to central concepts through star networks. For instance, when we say 'horse', its image, definition, and other sensory associations immediately come to our mind. Accordingly, **cross-learning multimodal embeddings through grounded concepts** is promising. To my knowledge, there are many works on multimodality, [13, 14] but they fail to theorize their approaches and overlook the essence of grounding in concepts.

Every challenge I've faced has only deepened my love for this field, fueling my determination to make meaningful research progress. I would be honored to join your esteemed lab and to make contributions with my passion, expertise, and vision.

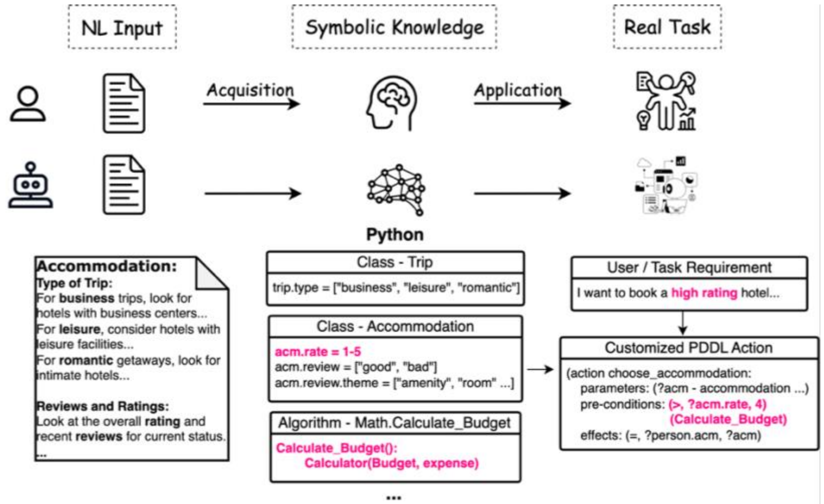

Figure.1. We apply human reasoning process (1st row) to agent reasoning (2nd & 3rd rows). Problem solving contains 2 steps: Acquisition, where NL input is encoded as symbolic knowledge, and Application, where task specific knowledge is employed. Python functions as a cerebral symbolic network (mid column), compiling structured relations and logics succinctly. We dynamically implement the Planning Task, e.g., PDDL Actions, (right column) facing various constraints (pink part).

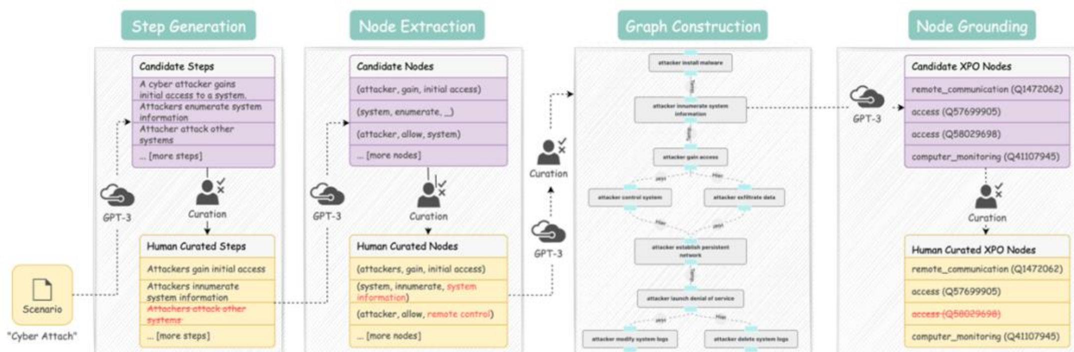

Figure.2. Our schema curation system includes 4 stages: Step Generation, Node Extraction, Graph Construction and Node Grounding; Outputs from the model are highlighted with a purple background; Outputs that have been curated by humans are highlighted with a yellow background; Modifications made by humans are marked in **red**.

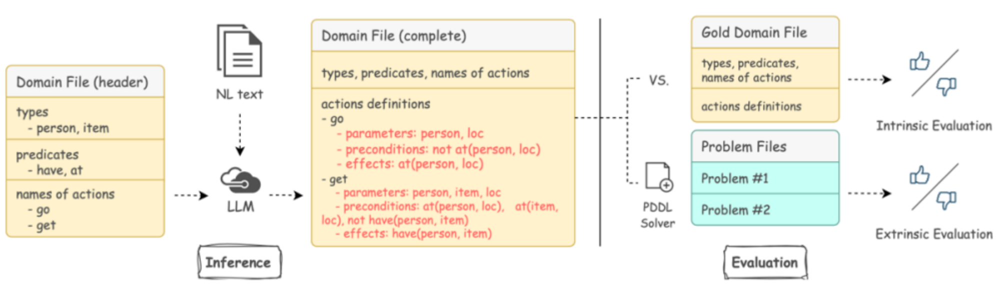

Figure.3. Our CoT approach predicts Actions in the DF, including parameters, preconditions, and effects based on textual descriptions. The CoT guides LM to extract symbolic representation from natural language with 3 steps: content extraction, entity-state inference, and translation.

Reference:

[1] **Zhang, T.**, Sulem, E., Roth, D. Question-Answering Data Augmentation for Argument Role Labeling. In submission

[2] **Zhang, T.** *, Tham, I. *, Hou, Z. *, Ren, J., Zhou, L., Xu, H., Zhang, L., Martin, L., Dror, R., Li, S., Ji, H., Palmer, M., Brown, S., Suchocki, R., and Callison-Burch, C. (2023). Human-in-the-Loop Schema Induction. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations), Toronto, Canada. Association for Computational Linguistics.

[3] **Zhang, T.** *, Zhang, L. *, Hou, Z., Wang, Z., Gu, Y., Clark, P., Callison-Burch, C., and Tandon, N. PROC2PDDL: Towards Open-Domain Symbolic Planning. In submission

[4] Piaget, J., & Cook, M. (1952). The origins of intelligence in children (Vol. 8, No. 5, pp. 18-1952). New York: International Universities Press.

[5] Vygotsky, L. S., & Cole, M. (1978). Mind in society: Development of higher psychological processes. Harvard university press.

[6] Riabi, A., Scialom, T., Keraron, R., Sagot, B., Seddah, D., & Staiano, J. (2020). Synthetic data augmentation for zero-shot cross-lingual question answering. arXiv preprint arXiv:2010.12643.

[7] Deutsch, D., & Roth, D. (2021). Incorporating Question Answering-Based Signals into Abstractive Summarization via Salient Span Selection. arXiv preprint arXiv:2111.07935.

[8] Ciosici, M. R., Cummings, J., DeHaven, M., Hedges, A., Kankanampati, Y., Lee, D. H., ... & Freedman, M. (2021). Machine-assisted script curation. arXiv preprint arXiv:2101.05400.

[9] Younes, H. L., & Littman, M. L. (2004). PPDDL1. 0: The language for the probabilistic part of IPC-4. In Proc. International Planning Competition.

[10] Yu, D., Zhu, C., Yang, Y., & Zeng, M. (2022, June). Jaket: Joint pre-training of knowledge graph and language understanding. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 10, pp. 11630-11638).

[11] Yasunaga, M., Bosselut, A., Ren, H., Zhang, X., Manning, C. D., Liang, P. S., & Leskovec, J. (2022). Deep bidirectional language-knowledge graph pretraining. Advances in Neural Information Processing Systems, 35, 37309-37323. (architecture)

[12] Hua, Y., Deng, Z., & Mckeown, K. (2023, July). Improving Long Dialogue Summarization with Semantic Graph Representation. In Findings of the Association for Computational Linguistics: ACL 2023 (pp. 13851-13883). (architecture)

[13] Ahn, M., Brohan, A., Brown, N., Chebotar, Y., Cortes, O., David, B., ... & Zeng, A. (2022). Do as i can, not as i say: Grounding language in robotic affordances. arXiv preprint arXiv:2204.01691. (multimodal)

[14] Sharma, P., Torralba, A., & Andreas, J. (2021). Skill induction and planning with latent language. arXiv preprint arXiv:2110.01517. (multimodal)