

Yue Yang

📍 220 South 33rd Street, Towne 299, Philadelphia, PA, 19104

✉ yueyang1@seas.upenn.edu

☎ (+1)-215-452-1517

🌐 yueyang1996.github.io

EDUCATION BACKGROUND

University of Pennsylvania

Department of Computer and Information Science (Penn NLP)

Ph.D. in Computer and Information Science, advised by Dr. Chris Callison-Burch and Dr. Mark Yatskar.

Concentrate on the intersection area of Natural Language Processing (NLP) and Computer Vision (CV).

Philadelphia, PA

Aug. 2020 - May. 2025

University of Pennsylvania

Department of Computer and Information Science (GRASP)

Master of Science in Engineering in Robotics, GPA: 3.88/4.00.

Philadelphia, PA

Aug. 2018 - May. 2020

Zhejiang University (ZJU)

School of Energy Engineering

Bachelor of Engineering in Mechanical Design and Manufacture and Automatization, GPA: 3.83/4.00.

Hangzhou, China

Sept. 2014 - July. 2018

RESEARCH INTERESTS

My current research focuses on two directions to marry Vision and Language:

- Vision helps Language: Some commonsense knowledge is not explicitly written in text, e.g., “an orange is orange” and thus it is difficult for language models to learn such knowledge. However, it is possible to leverage visual signals to alleviate the reporting bias of language models. [5, 6]
- Language helps Vision: End-to-end models have achieved great success in vision tasks but they lack interpretability. However, since language is human-readable, it is feasible to improve interpretability on vision tasks by leveraging language as the intermediate representation. [4, I]

In addition, I am also interested in procedural knowledge [2, 3, II], especially learning it from multimodal data, e.g., videos [4] and images [1], and its applications in different downstream tasks.

PUBLICATIONS

7. **Yue Yang**, Artemis Panagopoulou, Shenghao Zhou, Daniel Jin, Chris Callison-Burch, Mark Yatskar. **Language in a Bottle: Language Model Guided Concept Bottlenecks for Interpretable Image Classification**. *Preprint*, 2022.
6. **Yue Yang**, Wenlin Yao, Hongming Zhang, Xiaoyang Wang, Dong Yu, Jianshu Chen. **Z-LaVI: Zero-Shot Language Solver Fueled by Visual Imagination**. *EMNLP 2022*.
5. **Yue Yang***, Artemis Panagopoulou*, Marianna Apidianaki, Mark Yatskar and Chris Callison-Burch. **Visualizing the Obvious: A Concreteness-based Ensemble Model for Noun Property Prediction**. *Findings of EMNLP 2022*.
4. **Yue Yang**, Joongwon Kim, Artemis Panagopoulou, Mark Yatskar and Chris Callison-Burch. **Induce, Edit, Retrieve: Language Grounded Multimodal Schema for Instructional Video Retrieval**. *Open-Domain Retrieval Under Multi-Modal Settings Workshop, CVPR 2022*.
3. Artemis Panagopoulou, Manni Arora, Li Zhang, Dimitri Cugini, Weiqiu You, **Yue Yang**, Liyang Zhou, Yuxuan Wang Zhaoyi Hou, Alyssa Hwang, Lara Martin, Sherry Shi, Chris Callison-Burch, Mark Yatskar. **QuakerBot: A Household Dialog System Powered by Large Language Models**. *1st Proceedings of Alexa Prize TaskBot (Alexa Prize 2021)*.
2. Shuyan Zhou*, Li Zhang*, **Yue Yang**, Qing Lyu, Graham Neubig and Chris Callison-Burch. **Show Me More Details: Discovering Event Hierarchies from WikiHow**. *ACL 2022*.
1. **Yue Yang**, Artemis Panagopoulou, Qing Lyu, Li Zhang, Mark Yatskar and Chris Callison-Burch. **Visual Goal-Step Inference using wikiHow**. *EMNLP 2021*.

CURRENT RESEARCH PROJECTS

I. Language-Model-Guided Concept Bottleneck Models

Philadelphia, PA

PhD Independent Study | [website]

Mar. 2022 - Present

- Concept Bottleneck Model (CBM) aims to solve the black-box problem of end-to-end models by mapping the inputs onto a set of interpretable concepts and using the concepts to make predictions.
- Previous work relies on human-annotated concepts, which are costly and time-consuming to collect. In addition, CBMs are often less effective than end-to-end models in terms of accuracy.
- To fill these two gaps, we propose a novel approach to automatically construct the bottlenecks for any image classification tasks by leveraging the knowledge of large language models like GPT-3.
- Our LM-Guided CBM outperforms the end-to-end model in few-shot scenarios and achieves competitive performance under full supervision across 9 image classification datasets.

II. Implicit Entity State Changes in Procedures

Philadelphia, PA

Collaborate with CMU

May. 2022 - Present

- Reasoning about entity state changes in procedures is incredibly beneficial for building embodied agents and smart household systems. For example, after the step “turn on the stove”, the agent should know the *stove* is becoming *hot* and thus cannot be touched by hand.
- We create a dataset of more than 200 procedures, in which the models need to answer 300 questions regarding the environment that require multihop knowledge of entity changes.
- We showed that existing language models struggle at the task, while our proposed model based on Codex achieves high performance by converting procedures into python scripts.

PROFESSIONAL EXPERIENCE

Tencent AI Lab

Seattle, WA

Research Scientist

May. 2022 - Aug. 2022

- Improve the zero-shot ability of language models through visual imagination.

University of Pennsylvania/Coursera

Philadelphia, PA

Associate Instructor for Online Courses

Sept. 2020 - Present

- Develop an Artificial Intelligence course on Coursera.
- Over 200 students enrolled in the 2021 fall semester.
- Design the R2D2 projects such as obstacle avoidance, AR path planning, mask detection, etc.

TEACHING EXPERIENCE

- CIS-521 Artificial Intelligence (19fall, 20fall, 21summer, 21fall, 22spring, 22fall), Head TA.
- CIS-530 Computational Linguistics (21spring), TA.

AWARDS & ACHIEVEMENTS

- Alexa Taskbot Challenge Finals, Amazon May. 2022
- SEAS Outstanding Teaching Award, UPenn May. 2020
- Elite Liu Yongling Scholarship, offered by Hong Kong Elite Co., Ltd. (1/224) Aug. 2017
- Official Delegate of 32nd Student Congress, ZJU Jun.2017

SKILLS

- Programming languages: Python, Matlab, C, Bash, Coq.
- Machine learning libraries: PyTorch, TensorFlow, Keras, scikit-learn, NumPy.