WANJUN ZHONG

zhongwj25@mail2.sysu.edu.cn || https://zhongwanjun.github.io || (+86) 13609749192

ABOUT ME

SYSU-MSRA joint PhD in Natural Language Process field at present

- Advised by Dr. Ming Zhou, Prof. Jian Yin and Prof. Jiahai Wang
- Strong research experience on Natural Language Processing.
- Solid knowledge on ML and DL algorithms.
- Research Interests:
  (1) **machine reasoning**: logical reasoning, commonsense reasoning, neural-symbolic reasoning, multi-hop reasoning, etc.
  (2) **natural language understanding**: question answering, fact checking, etc.
  (3) **knowledge-aware pre-training:** pre-trained models with linguistic and semantic knowledge.
  (4) **Multi-modal Pre-training and Understanding**

EDUCATION

Joint PhD program between Sun Yat-Sen University (SYSU) and Microsoft Research Asia (MSRA) **Sep 2018 – June 2023**
PhD in Computer Science and Technology (Expected June 2023)
Advisors: Dr. Ming Zhou, Prof. Jian Yin and Prof. Jiahai Wang

Sun Yat-Sen University (SYSU) **Sep 2014 – June 2018**
Bachelor of Science in Software Engineering

RESEARCH EXPERIENCE

Machine Reasoning for Multi-modal Knowledge Understanding **MSRA Intern**
This research aims to build machine reasoning models that can better understand multi-modal knowledge (video and text), and better model the interaction between knowledge reside in different modalities.

Unified Model for QA and General NLP tasks **Visitor on Tsinghua University**
This research aims to build unified models for question answering and general NLP tasks with the bridge of an extensible and flexible structural prompt and prompt-based pre-training, which provides the benefits of better sharing common knowledge among tasks, and the ability to distinguish different tasks.

Machine Reasoning for General NLP tasks **MSRA Intern**
This research aims to build machine reasoning models that can integrate different types of knowledge (i.e., unstructured knowledge, structured knowledge and logical knowledge) and reasoning techniques for general NLP tasks that are more applicable in real user scenario, including dialog and open-domain question answering.

Complex Machine Reasoning with Logical-Level Reasoning Model **MSRA Intern**
This research aims to build machine reasoning models that can leverage logical knowledge for solving reasoning-required downstream tasks.

- We propose a method called LReasoner, which uses logical rules to generate expanded facts and use them as evidence in inference. LReasoner is the first AI model that achieves human parity on the ReCLor leaderboard for logical reasoning questions.
- We propose a method called ARM, which uses logical rules as guidance to make multi-step deduction to find the solution. A benchmark dataset called AR-LSAT is also built for the analytical reasoning task.
- We propose a method to use logical rules as constraints to learn better representations for downstream tasks.
- We apply these methods to Logical Reasoning (under review), Analytical Reasoning (under review) and Propaganda Detection (EMNLP 2020).

Machine Reasoning over Unstructured Knowledge with Graph-Based Reasoning Model **MSRA Intern**
This research aims to build graph-based machine reasoning models that can leverage unstructured texts as knowledge for solving reasoning-required downstream tasks.

- We build a semi-structured graph based on unstructured knowledge, which can include not only each task input, but also the related passages retrieved from unstructured texts as evidence. A graph-based reasoning module initialized by the pre-trained language model is used to reason over the semi-structured graph and predict the task output.
- We apply this framework to Fake News Detection (ACL 2021), Fact Checking (ACL 2020) and Deepfake Detection (EMNLP 2020).

Machine Reasoning over Structured Knowledge **MSRA Intern**
This research aims to build a machine reasoning framework that can leverage structured knowledge (i.e., commonsense knowledge, knowledge graph and web table) for solving reasoning-required downstream tasks.

- We propose a pre-training method to learn entity representations based on a given structured knowledge base.
- We propose to leverage logical operations for table-based fact checking with program-guided neural modular network.
- We apply these methods to Table-based Fact Checking (ACL 2020) and Commonsense Question Answering (NLPCC 2019).

TEACHING EXPERIENCE

- Teaching Assistant of two artificial intelligence courses, 2019
- Teaching Assistant of courses: linear algebra and discrete mathematics, 2015

HONORS & AWARDS

Academic Competition

- ECCV-Ego4D Challenge for Episodic Memory Natural Language Queries — 3rd — 2022
- Global Artificial Intelligence Application Competition — Merit Awards — 2018
- China College Students Data Modeling Contest — National Second Prize — 2016
- FASHIONAI GLOBAL CHALLENGE — 3rd, 7th in two channels of quarter-finals — 2018

Scholarship

- Microsoft Research Fellowship Award (11 Outstanding Ph.D. Students in Computer Science in the Asia-Pacific Area), 2021
- Baidu Scholarship (Global Top 40), 2021
- National Scholarship of PhD (top 0.2%), by Ministry of Education of China, 2020
- The First Prize Scholarship, 2016
- The Second Prize Scholarship, 2017, 2018, 2019, 2020
- Individual Scholarship, 2015

SKILLS

- **Languages:** C++, Python, Java, English CET-6
- **Deep Learning Frameworks:** TensorFlow, PyTorch

Publications

[1] Wanjun Zhong, Junjie Huang, Qian Liu, Ming Zhou, Jiahai Wang, Jian Yin and Nan Duan. *Reasoning over Hybrid Chain for Table-and-Text Open Domain Question Answering*, **IJCAI Oral (2022)**

[2] Wanjun Zhong, Yifan Gao, Ning Ding, Yujia Qin, Zhiyuan Liu, Ming Zhou, Jiahai Wang, Jian Yin, Nan Duan. *ProQA: Structural Prompt-based Pre-training for Unified Question Answering*. **NAACL (2022)**

[3] Wanjun Zhong, Siyuan Wang, Duyu Tang, Zenan Xu, Daya Guo, Jiahai Wang, Jian Yin, Ming Zhou and Nan Duan. *AR-LSAT: Investigating Analytical Reasoning of Text.* **NAACL (2022)**

[4] Junjie Huang*, Wanjun Zhong*, Qian Liu, Ke Xu, Daxin Jiang, Nan Duan. Mixed-modality Representation Learning and Pre-training for Joint Table-and-Text Retrieval in OpenQA (* indicates equal contribution) **EMNLP (2022)**

[5] Xinyu Pi*, Wanjun Zhong*, Yan Gao, Jian-guang Lou and Nan Duan. *LogiGAN: Learning Logical Reasoning via Adversarial Pre-training* (* indicates equal contribution) **NeurIPS (2022)**

[6] Wanjun Zhong, Duyu Tang, Jiahai Wang, Jian Yin and Nan Duan. *UserAdapter: Few-Shot User Learning in Sentiment Analysis.* **ACL (2021)**

[7] Wanjun Zhong, Duyu Tang, Zhangyin Feng, Nan Duan, Ming Zhou, Ming Gong, Linjun Shou, Daxin Jiang, Jiahai Wang and Jian Yin. *LogicalFactChecker: Leveraging Logical Operations for Fact Checking with Graph Module Network.* **ACL (2020)**

[8] Wanjun Zhong, Jingjing Xu, Duyu Tang, Zenan Xu, Nan Duan, Ming Zhou, Jiahai Wang and Jian Yin. Reasoning Over Semantic-Level Graph for Fact Checking. **ACL (2020)**

[9] Wanjun Zhong, Duyu Tang, Zenan Xu, Ruize Wang, Nan Duan, Ming Zhou, Jiahai Wang and Jian Yin. *Neural Deepfake Detection with Factual Structure of Text.* **EMNLP (2020)**

[10] Wanjun Zhong, Duyu Tang, Nan Duan, Ming Zhou, Jiahai Wang, Jian Yin. Improving Question Answering by Commonsense-Based Pre-Training. **NLPCC (2019)**

[11] Wanjun Zhong, Duyu Tang, Nan Duan, Ming Zhou, Jiahai Wang and Jian Yin. A Heterogeneous Graph with Factual, Temporal and Logical Knowledge for Question Answering Over Dynamic Contexts. arXiv, 2020

[12] Siyuan Wang, Wanjun Zhong, Duyu Tang, Zhongyu Wei, Zhihao Fan, Daxin Jiang, Ming Zhou and Nan Duan. *Logic-Driven Context Extension and Data Augmentation for Logical Reasoning of Text.* **ACL (2022)**

[13] Siyuan Wang, Zhongkun Liu, Wanjun Zhong, Ming Zhou, Zhongyu Wei, Zhumin Chen, Nan Duan. *From LSAT: The Progress and Challenges of Complex Reasoning.* **TASLP（2022)**

[14] Zenan Xu, Daya Guo, Duyu Tang, Qinliang Su, Linjun Shou, Ming Gong, Wanjun Zhong, Xiaojun Quan, Daxin Jiang and Nan Duan. *Syntax-Enhanced Pre-trained Model.* **ACL (2021)**

[15] Linmei Hu, Tianchi Yang, Luhao Zhang, Wanjun Zhong, Duyu Tang, Chuan Shi, Nan Duan and Ming Zhou. *Compare to The Knowledge: Graph Neural Fake News Detection with External Knowledge.* **ACL (2021)**

[16] Junjie Huang, Duyu Tang, Wanjun Zhong, Shuai Lu, Linjun Shou, Ming Gong, Daxin Jiang and Nan Duan. *WhiteningBERT: An Easy Unsupervised Sentence Embedding Approach.* **EMNLP (2021)**

[17] Ruize Wang, Duyu Tang, Nan Duan, Wanjun Zhong, Zhongyu Wei, Xuanjing Huang, Daxin Jiang and Ming Zhou. *Leveraging declarative knowledge in text and first-order logic for fine-grained propaganda detection.* **EMNLP (2020)**
[18] Zhijian Hou*, Wanjun Zhong*, Leiji, Kun Yan, Difei Gao, Wing-Kwong Chan, Chong-Wah Ngo, Zheng Shou, Nan Duan. *An Efficient COarse-to-fiNE Alignment Framework @ Ego4D Natural Language Queries Challenge 2022* (* indicates equal contribution) **ECCV Workshop (2022)**

Under Submission:
[19] Zenan Xu*, Wanjun Zhong*, Qinliang Su, Zijing Ou, Fuwei Zhang. *Semantic Composition and Alignment with Cross-Modality-Aware Syntactic Hypergraph Convolutional Network for Video Question Answering.* (* indicates equal contribution)
[20] Zenan Xu, Zexuan Qiu, Qinliang Su, Wanjun Zhong. *ReadE: Learning Relation-Dependent Entity Representation for Knowledge Graph Completion.*
[21] Wanjun Zhong, Yifan Gao, Ning Ding, Zhiyuan Liu, Ming Zhou, Jiahai Wang, Jian Yin and Nan Duan. *Improving Task Generalization via Unified Schema Prompt*
[22] Qiqi Gao*, Wanjun Zhong, Xiao Li, Jinglu Wang, Jie Li, Yan Lv. *CLUE: Contrastive Language-guided Learning for Referring Video Object Detection*
[23] Zhijian Hou*, Wanjun Zhong*, Leiji, Kun Yan, Difei Gao, Wing-Kwong Chan, Chong-Wah Ngo, Zheng Shou, Nan Duan. *CONE: An Efficient COarse-to-fiNE Alignment Framework for Long Video Temporal Grounding* (* indicates equal contribution)
[24] Wanjun Zhong*, Tingting Ma*, Jiahai Wang, Jian Yin, Tiejun Zhao, Chin-Yew Lin and Nan Duan. *Disentangling Reasoning Capabilities from Language Models with Compositional Reasoning Transformers* (* indicates equal contribution)