Abhay Zala

aszala@cs.unc.edu github.com/aszala aszala.com cs.unc.edu/~aszala/ Google Scholar

Education

Chapel Hill, NC
2019 - Exp. 2022

University of North Carolina at Chapel Hill

B.S in Computer Science (Minor: Data Science)

- Advisor: Prof. Mohit Bansal (MURGe-Lab, UNC-NLP Group)
- GPA: 3.52
- Relevant Coursework: Algorithms and Analysis - Data Structures - Modern Web Programming - Programming Language Concepts - Java Programming - System Fundamentals - Computer Organization - Models of Langauge and Computation - Special Topics in Computer Science - Foundations of Statistics and Data Science - Introduction to Probability - Discrete Mathematics - Calculus I, II, III

Publications

ArraMon: A Joint Navigation-Assembly Instruction Interpretation Task in Dynamic Environments

Hyounghun Kim, Abhay Zala, Graham Burri, Hao Tan, Mohit Bansal

Findings of EMNLP 2020

- We combine Vision-and-Language Navigation, assembling of collected objects, and object referring expression comprehension, to create a novel joint navigation-and-assembly task.
- During this task, the agent is asked to collect different target objects one-by-one by navigating based on natural language instructions in a complex, realistic outdoor environment, but then also arrange the collected objects part-by-part in an egocentric grid-layout environment.
- We implement a 3D dynamic environment simulator and collect a dataset (in English; and also extended to Hindi) with human-written navigation and assembling instructions, and the corresponding ground truth trajectories.

FixMyPose: Pose Correctional Captioning and Retrieval

Hyounghun Kim, Abhay Zala*, Graham Burri, Mohit Bansal*

Proceedings of AAAI 2021

- We introduce a new captioning dataset named FixMyPose.
- We collect descriptions of correcting poses (in both English and Hindi). The collected descriptions have interesting linguistic properties requiring an understanding of spatial relations and commonsense knowledge about postures.
- We introduce two new tasks. During the correctional-captioning task, models must generate descriptions of how to move from the current to target pose image, whereas in the retrieval task, models should select the correct target pose given the initial pose and correctional description.
- We present strong cross-attention baseline models (uni/multimodal, RL, multilingual) and also show that our baselines are competitive with other models when evaluated on other image-difference datasets.

Multimodal Counterfactual Commonsense Reasoning

Hyounghun Kim, Abhay Zala*, Mohit Bansal*

In Submission to AAAI 2022

- We introduce a new dataset and task for Counterfactual Commonsense.

** indicates equal contribution*

Experience

Undergraduate Research Assistant, MURGe-Lab

2019 (September) - Present

- Drove new **NLP** and **ML** ideas at the nationally recognized MURGe-Lab at the University of North Carolina at Chapel Hill (UNC).
- Working on integrating NLP with other fields such as **Computer Vision** and **Embodiment**.
- Created new tasks for multi-modal models and several baselines models to prove their validity.
- Analyzed models on existing metrics as well as several new metrics.
- Created two large datasets of several thousands of NLP instructions and images, along with simulation environments for multi-modal models.
- Presented research work at the EMNLP and AAAI conferences.

AI Researcher Summer Intern, MURGe-Lab

2020 (May - August)

- Developed AI model with **PyTorch** to interpret language instructions and perform robotic actions.
- Setup simulation environment in **Unity** with **C#** and hosted on **Amazon Web Services (AWS)**.
- Created a new task and dataset (7.7k task instances and 30.8k NLP instructions) for multi-modal models.
- Presented several baseline models (integrated and biased) and their performance on metrics (nDTW, CTC, rPOD, PTC).

Skills

- Python, PyTorch, NLP / Machine Learning, Java, C#, C/C++, Git, Github, Docker, Kubernetes, Linux, AWS, GCP, GKE, IBM Cloud, NodeJS, PHP, SQL, HTML5, CSS3, JavaScript/jQuery, Google Firebase, Unity Game Engine, OpenGL Microsoft and Adobe Software Suites, Web/Mobile/Desktop App Development

Other Projects and Awards

- UNC-NLP Group Website improvement

- Currently working to improve the UNC-NLP Group website.

- VR training system for healthcare workers

- **Won 1st place** (Best use of AR/VR for healthcare) at the HackReality 2021 Hackathon. Allows people that require lab equipment training to learn from home and without risk. Uses **Unity/C#** to create interactions with virtual patients to assist in worker training.

- Social media platform enabling networking and message sharing

- **Won 1st place** (Best use of GCP), **2nd place** (Best Hack for Accessibility / Inclusivity) and the **Wolfram Award** at the HackNC 2020 Hackathon. Containerized and deployed web app on a Kubernetes cluster running on GKE. Technologies used: **NodeJS**, **jQuery** and **Google Firebase**.

- AI based presentation control system

- **Won 1st place** at the NC State University PackHacks 2019 Hackathon. System listens to the speaker and automatically progresses the slides. Allows speakers to focus on the delivery while the AI handles the visuals.

- Online peer / AI tutoring service

- **Won 2nd place** (College Track) at the PackHacks 2021 Hackathon. Users can interact live with tutors or an AI powered by WolframAlpha. Developed using **NodeJS**, **Python**, **PyTorch**, **jQuery**, and **Google Firebase**.

- Website for showcasing movies and shows using the Red Ventures API

- **Won 3rd place** (Tech and Overall) at the Red Ventures Case Competition 2019. Leveraged REST APIs from the Red Venture digital platform.

- ACL Mentorship Program Logo

- Designed logo for the newly created Association for Computational Linguistics (ACL) Mentorship Program.

- Website development

- Personal portfolio and awarding winning site for competitions (aszala.com/tsa - 1st and 3rd place on state level for webdesign).