Ankur **Sikarwar**

Research Engineer, A*STAR Singapore

sikarwarank.com @ ankursikarwardc@gmail.com github.com/ankursikarwar Google Scholar

N09C46 Connexis North, 1 Fusionopolis Way, Singapore 138632

Education

June 2021 | **Birla Institute of Technology, Mesra** — **India**
July 2017 | Bachelor of Engineering | Electronics & Communication
First Class with Distinction

Summer of 2023 | **MIT Center for Brains, Minds and Machines** — **Cambridge, USA**
Brains, Minds, and Machines Summer Course (Non-Degree)
Awarded Fujitsu Laboratories Fellowship

Publications

R=In Review, C=Conference, P=Preprint

[C.1] **When Can Transformers Ground and Compose: Insights from Compositional Generalization Benchmarks.**
Ankur Sikarwar, Arkil Patel, Navin Goyal
Conference on Empirical Methods in Natural Language Processing. **[Oral]** [EMNLP'22]

[C.2] **Learning to Learn: How to Continuously Teach Humans and Machines.**
Parantak Singh, You Li, Ankur Sikarwar, Weixian Lei, Daniel Gao, Morgan Bruce Talbot, Ying Sun, Mike Zheng Shou, Gabriel Kreiman, Mengmi Zhang
International Conference on Computer Vision. [ICCV'23]

[C.3] **Decoding the Enigma: Benchmarking Humans and AIs on the Many Facets of Working Memory.**
Ankur Sikarwar, Mengmi Zhang
Conference on Neural Information Processing Systems. (Dataset & Benchmark Track) [NeurIPS'23]

[R.1] **Reason from Context with Self-supervised Learning.**
Xiao Liu, Ankur Sikarwar, Joo Hwee Lim, Gabriel Kreiman, Zenglin Shi, Mengmi Zhang
[In Review]

[R.2] **Human or Machine? Turing Tests for Vision and Language.**
Mengmi Zhang, Giorgia Dellaferrera, Ankur Sikarwar, Marcelo Armendariz, Noga Mudrik, Prachi Agrawal, Spandan Madan, Mranmay Shetty, Andrei Barbu, Haochen Yang, Tanishq Kumar, Shui'Er Han, Aman Raj Singh, Meghna Sadwani, Stella Dellaferrera, Michele Pizzochero, Brandon Tang, Hanspeter Pfister, Gabriel Kreiman
[In Review]

[P.1] **On the Efficacy of Co-Attention Transformer Layers in Visual Question Answering.**
Ankur Sikarwar, Gabriel Kreiman
Preprint.

Research Experience

Present | **Agency for Science, Technology and Research | Institute for Infocomm Research** — **Singapore**
Oct 2022 | *Research Engineer | Advisor: Dr. Mengmi Zhang*
Modeling human learning by training self-supervised methods on egocentric infant visual experiences (SAY-Cam dataset). Also, working on self-supervised methods for contextual reasoning.

Aug 2022 | **Microsoft Research** — **Bangalore, India**
Feb 2022 | *Research Intern | Advisor: Dr. Navin Goyal*
Developed models capable of generalizing compositionally in grounded language understanding tasks. Also, worked on the mechanistic interpretability of grounding and composition in multimodal transformers.

July 2021 | Worked on modular neural networks and on obtaining faithful interpretations of individual reasoning modules. Investigated compositional generalization benchmarks and exposed key design flaws in out-of-distribution testing.

July 2021 | **Harvard University | Kreiman Lab** — **Cambridge, USA**
Jan 2021 | *Research Assistant | Advisor: Dr. Gabriel Kreiman*
Investigated the efficacy of cross-modal attention in tasks like Visual Question Answering. Conducted interpretability studies on vision-language transformers using human attention maps.

July 2019 – May 2019 | **IIIT, Hyderabad | Center for Visual Information Technology** | **Hyderabad, India**
Research Intern | Advisor: Dr. Avinash Sharma
Worked on an end-to-end network for reconstructing 3D models of humans from monocular video. Developed tools for pre-processing & generating 3D mesh data of humans from a vertex-based template model.

Selected Research Projects

Self-supervised Learning for Contextual Reasoning Oct'22 - Present
Advisor: Dr. Mengmi Zhang, Prof. Gabriel Kreiman

- Working on a self-supervised learning method that captures associations between objects and their contexts.
- Proposed a new task, *Object Priming*, to evaluate contextual reasoning capabilities of models.
- Designed and conducted large-scale human psychophysics experiments to curate object priming maps from human subjects.

Memory-augmented Networks for Better Generalization Jan'23 - Present
Advisor: Dr. Mengmi Zhang

- Working on novel memory-augmented architectures with information bottlenecks for out-of-distribution generalization.
- Benchmarked contemporary memory architectures like RNNs, GRUs, LSTMs, and Transformers on challenging working memory tasks and showed that recurrent networks exhibit better alignment with human behavior compared to transformers.

Compositional Generalization in Grounded Language Understanding July'21 - Aug'22
Advisor: Dr. Navin Goyal

- Developed a transformer-based approach that achieves state-of-the-art performance on grounded compositional generalization benchmarks like gSCAN and ReaSCAN.
- Investigated bottlenecks for compositional generalization in current models and exposed key design flaws in previous benchmarks. Also showed that transformers generalize to higher depths of reasoning even when trained for shallower depths.
- Derived an explicit construction to mechanistically explain grounding and composition in transformers.

Analysis of Co-Attention in Multimodal Transformers Jan'21 - July'21
Advisor: Dr. Gabriel Kreiman

- Demonstrated that attention in co-attention transformer layers correlates more with human attention when compared with traditional CNN/LSTM networks.
- Evaluated the influence of question semantics in driving visual attention of vision-language transformers. Demonstrated that words, particularly nouns drive visual attention rather than grammar or semantics.

3D Reconstruction of Human Bodies from Monocular Video May'19 - July'19
Advisor: Dr. Avinash Sharma

- Worked on a 3D Human Reconstruction model capable of predicting 3D mesh from a few frames of a monocular RGB video.
- Integrated OpenPose in the pipeline for predicting joint locations of humans. Also, worked on texture stitching and mapping for the reconstructed 3D models.

Talks

"Decoding the Enigma: Benchmarking Humans and AIs on the Many Facets of Working Memory"

- Libedinsky Lab, National University of Singapore — July 2023

"When Can Transformers Ground and Compose: Insights from Compositional Generalization Benchmarks"

- EMNLP 2022 — Dec 2022
- Deep NeuroCognition Lab, A*STAR Singapore — Nov 2022
- Lab Sabha, Microsoft Research India — July 2022

"On the Efficacy of Co-Attention Transformer Layers in Visual Question Answering"

- Kreiman Lab, Harvard University — June 2021

Academic Service and Leadership Roles

Reviewer NeurIPS'23, EMNLP'23, ACL'23, EMNLP'22
Organizer Reading Group, Deep NeuroCognition Lab, A*STAR Singapore
Volunteer National Service Scheme | *Participated in STEM outreach programs for underprivileged kids.*

Skills and Relevant Coursework

Languages	Python, C, C++, MATLAB
Frameworks	PyTorch, Tensorflow, Keras
Other Skills	Amazon Mechanical Turk, jsPsych, psiTurk, Blender, Unity
Relevant Coursework	Linear Algebra, Probability Models & Stochastic Processes, Convex Optimization, Neural Networks & Fuzzy System, Machine Learning, Convolutional Neural Networks for Visual Recognition, Natural Language Processing with Deep Learning, Multivariable Calculus, Real Analysis, Data Structures, Information Theory & Coding

Honours and Awards

Fujitsu Laboratories Fellowship, 2023 For attending MIT Center for Brains, Minds and Machines Summer Course. Only undergrad to be selected from a pool of 300+ graduate students.

Graduated in First Class with Distinction, 2021 Birla Institute of Technology

iHack Alpha: AI-Enabled Solutions, 2021 Among Top 8 Finalists globally.

Bengaluru Tech Summit Global Hackathon, 2019 | Top 20 Finalists For developing "FOCUS: A Wearable Device for People with Speech and Motor Impairments."

NASA International Space Apps Challenge, 2019 | Global Nominee For developing "Prophet: A distributed system for identifying and mitigating lunar dust for future moon missions."

Microsoft Codefundo++, 2019 Runner's Up, Birla Institute of Technology.

Siemens MakeIT Real Hackathon, 2018 | Winner For developing the winning prototype "TetraChrome Lenses: Smart Glasses for Visually Impaired People" within 24 hours.

References

- Prof. Gabriel Kreiman ..... *Professor, Harvard University, Center for Brains, Minds and Machines, USA*
- Dr. Navin Goyal ..... *Principal Researcher, Microsoft Research, India*
- Prof. Mengmi Zhang ..... *Principal Investigator and Assistant Professor, A*STAR and NTU, Singapore*