Arkil **Patel**

Grad Student, Mila

+1 514-549-6593 | arkil.patel@gmail.com | Website | arkilpatel | arkil-patel | Google Scholar

Education

Present – Aug 2022 | **McGill University** | **Montreal, Canada**
MSc. (Thesis), Computer Science
Advisors: Prof. Dzmitry Bahdanau and Prof. Siva Reddy

July 2020 – Aug 2016 | **Birla Institute of Technology and Science Pilani** | **Goa, India**
B.E. (Hons.), Computer Science | CGPA: 9.03/10
Graduated with *Distinction*

Experience

Present – Aug 2022 | **Mila - Quebec AI Institute** | **Montreal, Canada**
Graduate Research Assistant | Advisors: Prof. Dzmitry Bahdanau and Prof. Siva Reddy
Working on compute-efficient Natural Language Processing and self-critiquing generations from Large Language Models.

Jul 2022 – Aug 2020 | **Microsoft Research** | **Bangalore, India**
Pre-doctoral Research Fellow | Advisor: Dr. Navin Goyal
Developed models capable of generalizing compositionally on semantic parsing and grounded language understanding tasks. Also analyzed (theoretically as well as empirically) the capabilities of neural models and exposed the deficiencies in existing datasets.

Dec 2019 – Jun 2019 | **Microsoft Research** | **Bangalore, India**
Research Intern | Advisor: Dr. Navin Goyal
Worked on theoretically understanding the abilities of Transformers. Also worked on developing robust and interpretable models for semantic parsing.

Publications

When Can Transformers Ground and Compose: Insights from Compositional Generalization Benchmarks [pdf]
Ankur Sikarwar, Arkil Patel, Navin Goyal
Accepted as a long paper at the 2022 Conference on Empirical Methods in Natural Language Processing [EMNLP'22]

Revisiting the Compositional Generalization Abilities of Neural Sequence Models [pdf] [code]
Arkil Patel, Satwik Bhattamishra, Phil Blunsom, Navin Goyal
60th Annual Meeting of the Association for Computational Linguistics [ACL'22]

Are NLP Models really able to Solve Simple Math Word Problems? [pdf] [code]
Arkil Patel, Satwik Bhattamishra, Navin Goyal
2021 Conference of North American Chapter of the Association for Computational Linguistics [NAACL'21]

On the Computational Power of Transformers and Its Implications in Sequence Modeling [pdf] [code]
Satwik Bhattamishra, Arkil Patel, Navin Goyal
2020 Conference on Computational Natural Language Learning [CoNLL'20]

VehicleChain: Blockchain-based Vehicular Data Transmission Scheme for Smart City [pdf]
Arkil Patel, Naigam Shah, Trupil Limbasiya, Debasis Das
2019 IEEE International Conference on Systems, Man and Cybernetics **[Oral]** [SMC'19]

Selected Projects

Compute-efficient Fine-tuning of Large Language Models Sep'22 - Present
Project Advisors: *Dr. Dzmitry Bahdanau and Prof. Siva Reddy*

- Working on designing transformer-based models for learning to optimize Large Language Models when finetuning.
- The objective is to eliminate compute-heavy backpropagation by automatically obtaining parameter update values.

Self-critiquing Large Language Models Sep'22 - Present
Project Advisors: *Dr. Dzmitry Bahdanau and Prof. Siva Reddy*

- Designing prompts with chains-of-thought and critiques that can enable LLMs to correct themselves when wrong.
- Developing approaches to automatically generate critiques for LLM generations.

Grounded Language Understanding Jul'21 - Jul'22

Project Advisor: *Dr. Navin Goyal*

- Designed a transformer-based approach that achieves state-of-the-art performance on grounded systematic generalization challenges such as gSCAN and ReaSCAN.
- Analysed the task to understand difficulty bottlenecks and exposed issues with the training set.
- Derived an explicit and interpretable construction that captures the model's behavior and completely describes the detailed computations corresponding to grounding and composition. Work accepted at **EMNLP'22**.

Analysing the Compositional Generalization Capabilities of Neural Sequence Models May'21 - Nov'21

Project Advisors: Dr. Navin Goyal and Prof. Phil Blunsom

- Showed that neural sequence models such as LSTMs and Transformers do have some inductive biases that enable them to generalize compositionally in the setting defined by SCAN, Colors and COGS datasets.
- Examined the learned embeddings of models to understand how they are able to generalize.
- Investigated the extent to which the bias exists by experimenting with different training distributions, model capacities and analysing transferability. Work published at **ACL'22**.

Semantic Parsing: Automatically Solving Math Word Problems Jun'19 - May'21

Project Advisor: *Dr. Navin Goyal*

- Worked on building robust and interpretable models to automatically solve math word problems.
- Conducted various experiments to show that existing models rely on shallow heuristics to solve the problem. Also created a challenge set to enable better evaluation of models. Work published at **NAACL'21**.

Analysing the Computational Power of Transformers Jun'19 - Dec'19

Project Advisor: *Dr. Navin Goyal*

- Theoretically analysed the computational power of transformers, as measured by Turing-completeness.
- Analysed the necessity (for Turing-completeness) of various components in the transformer architecture such as residual connections, attention blocks and FFNs.
- Empirically verified the relevance of our theoretical results. Work published at **CoNLL'20**.

Honours and Awards

2016 - 2020	**Institute Merit Scholarship,** awarded to top 10% students in the batch	BITS Goa, India
2012 - 2020	**National Talent Search Scholarship,** awarded to top 1000 students in the country	New Delhi, India

Skills

Languages	Python, C++, C, Java, SQL, MATLAB, Verilog
Frameworks	PyTorch, TensorFlow, Keras
Relevant Coursework	Machine Learning, Natural Language Understanding with Deep Learning, Neural Networks, Data Mining, Data Structures and Algorithms, Object Oriented Programming, Linear Algebra, Probability and Statistics, Multivariate Calculus, Discrete Mathematics

Services, Teaching and Leadership Roles

Teaching Assistant **Neural Networks and Fuzzy Logic** Jan'20 - May'20

Responsible for conducting tutorials for teaching the theory behind Deep Learning models. Also responsible for teaching implementation of DL models in PyTorch and designing the programming assessments.

Teaching Assistant **Data Mining** Jan'19 - May'19

Responsibilities included conducting the programming tutorials to teach implementation of ML algorithms and designing the programming assessments.

Teaching Assistant **Database Systems** Jan'19 - May'19

Assisted the faculty in conducting labs for a class of 200+ students and designed the questions for regular lab assessments as well as the final lab evaluation.

Organiser **Speech & NLP Reading Group, Microsoft Research India** May'21 - Jul'22

Organised the weekly Speech and NLP reading group at MSR India. Co-ordinated the scheduling and format of the meetings. Invited external speakers for giving expert talks.

Reviewer **ACL Rolling Review, EMNLP 2022, 2021, AAAI 2022, NAACL 2021**