Arkil **Patel**

Grad Student, Mila

📱 +1 514-549-6593　✉ arkil.patel@gmail.com　🌐 arkilpatel.github.io　⌨ arkilpatel　in arkil-patel　🎓 Google Scholar

Education

Present **Aug 2022**	**McGill University** MSc. (Thesis), Computer Science Advisors: Prof. Dzmitry Bahdanau and Prof. Siva Reddy	**Montreal, Canada** GPA: 4.0/4.0
July 2020 **Aug 2016**	**Birla Institute of Technology and Science Pilani** B.E. (Hons.), Computer Science Graduated with *Distinction*	**Goa, India** CGPA: 9.03/10

Experience

Present **Aug 2022**	**Mila - Quebec AI Institute** *Graduate Research Assistant | Advisors: Prof. Dzmitry Bahdanau and Prof. Siva Reddy* Working on multiple analysis projects evaluating the abilities and limitations of Large Language Models. Also working on understanding the generalization abilities of Transformers as well as emeregent phenomenon such as In-context learning.	**Montreal, Canada**
Aug 2023 **May 2023**	**Allen Institute for Artificial Intelligence - AllenNLP team** *Research Intern | Advisor: Dr. Pradeep Dasigi* Worked on evaluation of Large Language Models pre-trained on code.	**Seattle, USA**
Jul 2022 **Aug 2020**	**Microsoft Research** *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.	**Bangalore, India**
Dec 2019 **Jun 2019**	**Microsoft Research** *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.	**Bangalore, India**

Publications

MAGNIFICo: Evaluating the In-Context Learning Ability of Large Language Models to Generalize to Novel Interpretations [pdf] [code]
Arkil Patel, Satwik Bhattamishra, Siva Reddy, Dzmitry Bahdanau
2023 Conference on Empirical Methods in Natural Language Processing [EMNLP'23]

Understanding In-Context Learning in Transformers and LLMs by Learning to Learn Discrete Functions [pdf]
Satwik Bhattamishra, Arkil Patel, Varun Kanade, Phil Blunsom
Under Review at ICLR 2024

Simplicity Bias in Transformers and their Ability to Learn Sparse Boolean Functions [pdf] [code]
Satwik Bhattamishra, Arkil Patel, Varun Kanade, Phil Blunsom
61st Annual Meeting of the Association for Computational Linguistics [ACL'23]

When Can Transformers Ground and Compose: Insights from Compositional Generalization Benchmarks [pdf] [code]
Ankur Sikarwar, Arkil Patel, Navin Goyal
2022 Conference on Empirical Methods in Natural Language Processing **[Oral]** [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, <u>Arkil Patel</u>, Navin Goyal
2020 Conference on Computational Natural Language Learning   [CoNLL'20]

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

Selected Projects

Evaluating Large Language Models: Acquisition, Creation, and Usage of Code Modules   May'23 - Present
Project Advisors: *Dr. Pradeep Dasigi, Dr. Dzmitry Bahdanau, and Prof. Siva Reddy*

> Developing a benchmark to assess the ability of LLMs to acquire, create, and use code functions from information presented in-context.

Evaluating Large Language Models: Learning Novel Interpretations   Jan'23 - Jun'23
Project Advisors: *Dr. Dzmitry Bahdanau and Prof. Siva Reddy*

> Developed a benchmark based on text-to-SQL semantic parsing to assess the ability of LLMs to generalize to novel interpretations using in-context learning.
> Experimented with contemporary open-source LLMs such as LLaMA and StarCoder as well as state-of-the-art models such as GPT-3.5-turbo and GPT-4. Our findings highlighted various capabilities and limitations associated with in-context learning in LLMs. Our paper for this work is under submission at EMNLP'23.

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

2023 - 2024	**Canada Graduate Scholarship (Master's)**, awarded funding of $17,500	NSERC, Canada
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   From Natural Language to Data Science, McGill University   Jan'23 - May'23

Responsible for creating and grading assignments.

Teaching Assistant   Neural Networks and Fuzzy Logic, BITS Pilani   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.

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 talks.

Reviewer   ACL Rolling Review, ACL 2023, EMNLP 2023, 2022, 2021, AAAI 2022, NAACL 2021