

# Arkil Patel

Grad Student, Mila

+1 514-549-6593   [arkil.patel@gmail.com](mailto:arkil.patel@gmail.com)   [arkilpatel.github.io](https://github.com/arkilpatel)   [arkilpatel](https://www.linkedin.com/in/arkil-patel)   [in arkil-patel](https://scholar.google.com/citations?user=arkil-patel)   [Google Scholar](https://scholar.google.com/citations?user=arkil-patel)

## Education

<b>Present</b> <b>Aug 2022</b>	<b>McGill University</b> MSc. (Thesis), Computer Science Advisors: <a href="#">Prof. Dzmitry Bahdanau</a> and <a href="#">Prof. Siva Reddy</a>	<b>Montreal, Canada</b> GPA: 4.0/4.0
<b>May 2020</b> <b>Aug 2016</b>	<b>Birla Institute of Technology and Science Pilani</b> B.E. (Hons.), Computer Science Graduated with <i>Distinction</i>	<b>Goa, India</b> CGPA: 9.03/10

## Experience

<b>Present</b> <b>Aug 2022</b>	<b>Mila - Quebec AI Institute</b> <i>Graduate Research Assistant</i> / Advisors: <a href="#">Prof. Dzmitry Bahdanau</a> and <a href="#">Prof. Siva Reddy</a> Working on multiple analysis projects evaluating the abilities and limitations of Large Language Models. Also working on understanding the factors behind the effective generalization abilities of Transformers as well as emergent phenomenon such as in-context learning.	<b>Montreal, Canada</b>
<b>Aug 2023</b> <b>May 2023</b>	<b>Allen Institute for Artificial Intelligence - AllenNLP team</b> <i>Research Intern</i> / Advisor: <a href="#">Dr. Pradeep Dasigi</a> Worked on evaluating how well can Large Language Models in-context learn novel programming libraries for code generation.	<b>Seattle, USA</b>
<b>Jul 2022</b> <b>Aug 2020</b>	<b>Microsoft Research</b> <i>Pre-doctoral Research Fellow</i> / Advisor: <a href="#">Dr. Navin Goyal</a> 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.	<b>Bangalore, India</b>
<b>Dec 2019</b> <b>Jun 2019</b>	<b>Microsoft Research</b> <i>Research Intern</i> / Advisor: <a href="#">Dr. Navin Goyal</a> Worked on theoretically understanding the abilities of Transformers. Also worked on developing robust and interpretable models for semantic parsing.	<b>Bangalore, India</b>

## Publications

### Evaluating In-Context Learning of Libraries for Code Generation [\[pdf\]](#)

[Arkil Patel](#), Siva Reddy, Dzmitry Bahdanau, Pradeep Dasigi  
*Preprint, under review*

### 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 [\[Oral\]](#)

[\[EMNLP '23\]](#)

### Understanding In-Context Learning in Transformers and LLMs by Learning to Learn Discrete Functions [\[pdf\]](#)

Satwik Bhattamishra, [Arkil Patel](#), Phil Blunsom, Varun Kanade

*Preprint, under review*

### 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, 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

---

### Evaluating Large Language Models: Acquisition and Usage of Code Modules

May'23 - Nov'23

Project Advisors: Dr. Pradeep Dasigi, Dr. Dzmitry Bahdanau, and Prof. Siva Reddy

- > Developed a framework to assess the ability of LLMs to learn novel libraries and code functions from information presented in-context.
- > Showed that even smaller, open-source models such as StarCoder can learn code libraries in-context and that models such as GPT-4 can learn as efficiently from natural language descriptions as they can from demonstrations.

### 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 was accepted at EMNLP'23.

### Understanding In-Context Learning in Large Language Models

Jan'23 - Oct'23

Project Advisors: Prof. Varun Kanade and Prof. Phil Blunsom

- > Showed limitations of Transformers in in-context learning certain function classes known to be efficiently learnable.
- > On a testbed of boolean functions, showed that attention-free architectures can perform in-context learning but they lag behind Transformers on certain tasks.
- > Showed that LLMs pretrained on text data encode mechanisms that enable them to learn new functions in-context.

### 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.

## Honours and Awards

---

2023 - 2024	<b>Canada Graduate Scholarship (Master's)</b> , awarded funding of \$17,500	NSERC, Canada
2016 - 2020	<b>Institute Merit Scholarship</b> , awarded to top 10% students in the batch	BITS Goa, India
2012 - 2020	<b>National Talent Search Scholarship</b> , awarded to top 1000 students in the country	New Delhi, India

## Skills

---

**Languages** Python, C++, C, Java, SQL, MATLAB, Verilog  
**Frameworks** PyTorch, TensorFlow, Keras

## Teaching and Service

---

**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.

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