

# Length Generalization on Multi-Digit Integer Addition with Transformers

Imran Ibrahimli  
Universität Hamburg  
Dept. Informatik – Knowledge Technology, WTM



<https://www.informatik.uni-hamburg.de/WTM/>

# Outline

- 1 Introduction
- 2 Motivation
- 3 Research Questions
- 4 Background
- 5 Approach
- 6 Conclusion

# Outline

- 1 Introduction
- 2 Motivation
- 3 Research Questions
- 4 Background
- 5 Approach
- 6 Conclusion

# Introduction

- Overview of length generalization challenges in sequence tasks
- Importance of transformers for AI applications

# Outline

- 1 Introduction
- 2 Motivation**
- 3 Research Questions
- 4 Background
- 5 Approach
- 6 Conclusion

# Motivation

- Challenges with transformers in multi-digit addition
- Real-world implications of length generalization issues

# Problem Statement

- Transformers struggle with sequences longer than seen in training
- Importance of positional encoding for accurate digit alignment

# Outline

- 1 Introduction
- 2 Motivation
- 3 Research Questions**
- 4 Background
- 5 Approach
- 6 Conclusion



# Research Questions

- **RQ1:** Why do transformers with standard positional encodings fail in length generalization?
- **RQ2:** How do sub-task data influence length generalization?
- **RQ3:** How can interpretability techniques reveal transformer mechanisms?

# Outline

- 1 Introduction
- 2 Motivation
- 3 Research Questions
- 4 Background**
- 5 Approach
- 6 Conclusion

# Background - Transformers

- Key transformer components: self-attention, feed-forward networks
- Encoder-decoder vs. decoder-only models

# Background - Positional Encoding

- Absolute, relative, and random encoding methods
- Importance of positional encoding in handling sequence tasks

# Related Work

- Overview of research on length generalization and transformers in arithmetic tasks

# Outline

- 1 Introduction
- 2 Motivation
- 3 Research Questions
- 4 Background
- 5 Approach**
- 6 Conclusion

# Approach - Overview

- Focus on minimal architectural changes and positional encoding

# Experimental Setup

- Model configurations and evaluation metrics



# Data Formatting

- Techniques: zero padding, reversing answer, etc.

# Limitations of Absolute Positional Encoding

- Issues with absolute encoding and digit alignment

# Random Spaces Technique

- Description and impact on model generalization

# Results - Baseline Model Performance

- Baseline results with standard positional encoding

# Attention Map Analysis

- Visual analysis of attention patterns

# Impact of Data Formatting

- Comparing results across different formatting methods

# Sub-task Learning

- Role of sub-tasks in improving length generalization

# Discussion - Length Generalization

- Key findings on generalization capabilities



# Challenges and Limitations

- Limitations of random spaces and Abacus encoding

# Outline

- 1 Introduction
- 2 Motivation
- 3 Research Questions
- 4 Background
- 5 Approach
- 6 Conclusion**

# Conclusion

- Summary of findings and contributions

# Future Work

- Suggestions for further research in encoding methods and model interpretability

# The End

Thank you for your attention.  
Any questions?