

Assignment 3: Group Reimplementation Project

Check Absalon for deadline.

Introduction

You will work as a group to reimplement Generalization without Systematicity: On the Compositional Skills of Sequence-to-Sequence Recurrent Networks by Brendan Lake and Marco Baroni¹, and reproduce the results reported in Experiments 1, 2, and 3.

Tasks

- A. Implement a data loader for the SCAN tasks dataset.
- B. Re-implement² the **overall-best** architecture, and the **best** models described in Section 3. Note that the architecture of the best model³ is different for each experiment.
- C. Use your re-implementation to reproduce the results of Experiments 1, 2, and 3 in the paper.

Check Absalon for your group assignment. **Report any problems with the assignment as soon as possible.**

Grading (20%)

Your group will deliver a 10 minute presentation during the final lab / coursework session in December about the progress that has been made with the reimplementation project.

Your presentation will be evaluated on four criteria:

1. Description of how your group implemented the encoder-decoder models
2. Presentation and analysis of Experiments
3. Handling of questions
4. Clarity of the presentation

Submission

In addition to preparing and delivering the presentation, your group **must** upload the slides of your presentation before 0900 on the day of your presentation.

Failure to complete this step will result in a failure to complete the assignment.

Techincal Notes

It is **your responsibility** to make sure your computer is compatible with the projector in the lecture room.

¹ <http://proceedings.mlr.press/v80/lake18a.html>

² You are encouraged to using existing implementations of the Simple RNN, LSTM, and GRU. If you base your implementation on the PyTorch Seq2Seq Tutorial, be aware of the differences between its implementation of the Attention-based Decoder compared to the implementation described in the Supplementary Material of Lake & Baroni.

³ Described as “top-performing network” (Experiment 1), “best result” (Experiment 2), and “best performance” (Experiment 3)