

## L2X: Interpolated n-gram models

Marco Kuhlmann

While the neural models that you have seen in lab L2 define the state of the art in language modelling, they require substantial computational resources. Where these are not available, the older generation of probabilistic language models can make a strong baseline. Your task in this lab is to evaluate one of these models on the WikiText-2 dataset.

### Problem description

Read the section on Interpolation in Eisenstein's book (section 6.2.3).

1. Implement an interpolated trigram model with a uniform distribution over the model weights ( $\lambda$ ); you can do this in pure Python. Report the perplexity of the model on the validation data.
2. Implement the Expectation–Maximization algorithm for finding the best values of the model weights. Tune the weights on the *test* section of the data and report the perplexity of your best model on the validation section.
3. Summarize your results and observations in a short report (PDF, ca. 1 page). Submit a single zip-file with your code, instructions for how to run it in order to reproduce your results, and your report.