THO-MAS: Seq2seq Morphological Analyzer
Byung-Doh Oh, Pranav Maneriker, Nanjiang Jiang

Shared baseline encoder

Binary Relevance

Seq2seq decoder

Results

Analysis

- Binary relevance helped a lot by narrowing down output space
- Seq2seq helped with accuracy but not F1 (probably due to error propagation)
- On average over 107 datasets, seq2seq outperforms baseline by 14.25 accuracy, 4.6 F1

Seq2seq ... than Bin. Rel.:
- Reproduces longer sequences
- Predicts less invalid tag sequences, better at copying
  - On average over the 6 treebanks: 220 fewer unseen tag pairs in tests predicted
- Performs better on tense/aspect/mood tokens, capture tag dependencies
  - On Turkish: 10.6+ accuracy 12+ F1