

# Neural and Symbolic Arabic Paraphrasing with Automatic Evaluation

## First Author

Affiliation / Address line 1  
Affiliation / Address line 2  
Affiliation / Address line 3  
email@domain

## Second Author

Affiliation / Address line 1  
Affiliation / Address line 2  
Affiliation / Address line 3  
email@domain

## Abstract

We present a tool for Arabic paraphrasing that yields good paraphrasing accuracy. We present and compare several methods for paraphrasing and obtaining monolingual parallel data. We also present first results on Arabic paraphrasing using neural methods. Additionally, we propose a new evaluation metric for paraphrasing that is shown to correlate highly with human judgement.

agencies covering the same event. We expect headlines describing the same event to have some degree of semantic similarity yet different surface realizations.

The sequence-to-sequence models required a relatively long time to run which limited the testing of other architectures and model options. We plan to conduct more experiments on different architectures and compare results.

## Acknowledgments

- 1 Introduction
- 2 Related Work
- 3 Obtaining Parallel Monolingual Data
- 4 Extracting Paraphrases
- 5 Generating Paraphrased Sentences
  - 5.1 Phrase Substitution Method
  - 5.2 Neural Seq-to-Seq Method
- 6 An Automatic Evaluation Metric
  - 6.1 Semantic Similarity
  - 6.2 Surface Variation
- 7 Analysis
- 8 Evaluation
- 9 Future Work

We plan to explore other options for obtaining monolingual parallel data. One possible approach is to retrieve headlines of news articles from different