University of Illinois System in HOO Text Correction Shared Task

Alla Rozovskaya, Mark Sammons, Joshua Gioja and Dan Roth
{rozovska, mssammon,gioja,danr}@illinois.edu

Text Correction Task

The **Text Correction task** addresses the problem of detecting and correcting mistakes in text. This task is challenging, since many errors are not easy to detect.

HOO Text Correction Shared Task

- Writing mistakes made by **non-native speakers of English.**
- Focuses on papers in the **Natural Language Processing** community.

Our Contributions

- We target several common types of errors: **articles, prepositions, word choice, punctuation**.
- We implement **adaptation techniques** for article and preposition error correction and demonstrate their success.

Adaptation Techniques

- Mistakes made by non-native speakers are systematic.
- Injecting knowledge about typical errors into the system improves its performance significantly.
- In our previous work, we proposed **methods to adapt a model to the typical errors using error statistics**.
- The preposition and article systems use these methods with additional improvements.

System Components

Component	Relative Freq.	Targeted Errors	Examples
Article	18%	Missing/Unnecessary/ Replacement	Section 5:1 describes the details of None*/the evaluation metrics. The main advantage of the*/None phonetical ignment is that it requires no training data.
Preposition	9%	Replacement	Pseudo-word searching problem is the same to*/as decomposition of a given sentence into pseudo-words.
Word choice	-	Various lexical and grammatical errors	
Punctuation	18%	Missing/Unnecessary	In the thesaurus we incorporate LCSbased*/LCS-based semantic description for each verb class.

The column "Relative frequency" shows the proportion of a given error type in the pilot data. The category "Article" is based on the statistics for determiner errors, the majority of which involve articles.

Type-Based Performance

Team	Run	Detection	Recognition	Correction
JU	0	0.029	0.029	0.029
LI	3	0.048	0.048	0.033
NU	0	0.372	0.368	0.276
UD	-	-	-	
UI	8	0.505	0.505	0.449
UT	1	0.040	0.025	0.025

Articles. For each team, the F-scores for the best run are shown.

Team	Run	Detection	Recognition	Correction
JU	0	0.035	0.035	0.035
LI	8	0.039	0.039	0.039
NU	0	0.266	0.266	0.168
UD	5	0.079	0.079	0.000
UI	8	0.488	0.488	0.363
UT	4	0.202	0.202	0.117

Prepositions. For each team, the F-scores for the best run are shown.

- In the overall performance, **our system ranked first** in all three evaluation metrics (Detection, Recognition, and Correction).
- Dale and Kilgarriff (2011) give only Recall scores for type-based performance because it is not possible to compute Precision for open-class errors. Since it is easy to obtain high recall by proposing many edits and, similarly, easy to obtain high precision by just proposing no edits, we have done a slightly different evaluation for closed-class errors, articles and prepositions, and present results sorted by F-score.

Article/Preposition Errors

- Trained on the ACL Anthology corpus
- Features: words and part-of-speech tags in a 4-word window.
- A discriminative learning framework (Averaged Perceptron) in Learning Based Java (Rizzolo and Roth, 2007).
- **Adaptation to the typical errors** based on the methods proposed in Rozovskaya and Roth (2010, 2011).

Word Choice Errors

- Various context-sensitive confusions: spelling errors, grammatical, and word choice errors.
- A Naïve Bayes classifier trained on the ACL Anthology and the North American News NY Times Text.

Punctuation Errors

- **Missing commas**: corrected with a set of rules.
- **Misuse of hyphens**: discover automatic rules by extracting mappings between hyphenated and non-hyphenated sequences using n-gram counts computed from the ACL Anthology Corpus.

  LCSbased → LCS-based
  para-linguistics → paralinguistics

COGNITIVE COMPUTATION GROUP UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

This research is supported by the U.S. Department of Education and DARPA.