

You can be Shakespeare!

A Case Study in Paraphrase Targeting Writing Styles

Author^{1,2} *Author*^{2,3}

(1) INSTITUTE_1, address 1

(2) INSTITUTE_2, address 2

(3) INSTITUTE_3, address 3

author1@institute1, author@institute2

ABSTRACT

We present initial investigation into the task of paraphrasing language while targeting a particular writing style. The plays of William Shakespeare and their modern translations are used as a testbed for evaluating paraphrase systems targeting a specific style of writing. We empirically show that even with a relatively small amount of parallel training data available, it is possible to learn paraphrase models which capture stylistic phenomenon, and these models outperform baselines based on dictionaries and out-of-domain parallel text. In addition we present an initial investigation into automatic evaluation metrics for paraphrasing writing style. To the best of our knowledge this is the first work to investigate the task of paraphrasing text with the goal of targeting a specific style of writing.

KEYWORDS: Paraphrase, Writing Style.

1 Introduction

Identical meaning can be expressed or *paraphrased* in many different ways; automatically detecting or generating different expressions with the same meaning is fundamental to many natural language understanding tasks(), so much previous work has investigated methods for automatic paraphrasing. Two different utterances may be semantically equivalent, however they can also be stylistically quite different. For example, the same information about a new product is likely to be conveyed using very different lexical and grammatical patterns in advertising materials v.s. technical manuals, or in Shakespearean plays v.s. Hollywood movies.

In this paper, we investigate the task of automatic paraphrasing when targeting a specific writing style, focusing specifically on the style of Early Modern English employed by William Shakespeare. We exploit modern translations of 17 plays written to help literature students to better understand Shakespeare. These modern translations are used to generate a parallel corpus of Shakespeare's style and modern English, which is then used to train phrase-based translation models which are capable of automatically paraphrasing ordinary sentences into Shakespearean English. In addition we develop several baseline systems which don't make use of this source of parallel text and instead rely on dictionaries of expressions commonly found in Shakespearean English, or parallel monolingual text gathered through Amazon's Mechanical Turk ().

We evaluate these models both through human judgements and standard evaluation metrics from the Machine Translation and paraphrase literature, however no previous work has investigated the ability of automatic metrics to capture the notion of writing style. We propose

several new metrics for evaluating the task of paraphrasing while targeting a specific style, and measure correlation with human judgement showing promising results for this particular style.

Systems which are capable of automatically paraphrasing literary writing styles could be directly beneficial for educational applications, for example helping students to experiment with writing literature in the style of authors they are studying. Additionally note that out of the 37 surviving plays written by William Shakespeare, only 17 currently have modern translations available; although we have not yet formally evaluated paraphrasing in the other direction, this work has the potential to additionally make the other 20 plays more accessible to students of Shakespeare.

2 Data

We propose to use Shakespeare's plays and their modern English translations as a testbed the task of paraphrasing targeting a specific writing style. Having access to parallel text in the target style allows us to train statistical models for generating paraphrases, and also perform automatic evaluation using BLEU which requires access to a set of reference translations. For this purpose we scraped modern translations of 17 Shakespeare plays from <http://nfs.sparknotes.com>.

- Motivate the need for a benchmark dataset for evaluating the writing style paraphrase task
- Present Shakespeare / Modern translation data as a situation where we have parallel data available (useful for building models & evaluating automatic evaluation metrics)

3 Evaluation Metrics

- Describe the need for automatic evaluation metrics.
- Describe previously used evaluation metrics for paraphrase.
- Highlight problems with previous metrics when targeting a specific writing style.
- Propose new metrics.

4 Experiments

- Experimental setup.
- Present results from human evaluation comparing various systems.
- Analyze correlation between evaluation metrics and human judgments.

5 Related Work

- Kevin Knight's work on poetry generation
- Any work on writing style (e.g. classification)? Possibly cite work on author attribution...
- work on paraphrase evaluation metrics (David Chen, CCB, etc...)

6 Conclusions