Task-Oriented Paraphrase Analytics

LREC-COLING 2024

Marcel Gohsen
Matthias Hagen
Martin Potthast
Benno Stein
Paraphrase Definitions

“convey the same meaning but use different words” [Bhagat and Hovy, 2013]

“restatements with approximately the same meaning” [Wang et al., 2019]

“talking about the same situation in a different way” [Hirst, 2003]
Paraphrase Definitions

“convey the same meaning but use different words” [Bhagat and Hovy, 2013]

“restatements with approximately the same meaning” [Wang et al., 2019]

“talking about the same situation in a different way” [Hirst, 2003]
Paraphrase Definitions

“convey the same meaning but use different words” [Bhagat and Hovy, 2013]

Original: The future of the nation is in your hands.
Paraphrase: The nation’s future is in your hands.

“restatements with approximately the same meaning” [Wang et al., 2019]

Original: This was a bland dish.
Paraphrase: I’ve made dishes that are much tastier than this one.

“talking about the same situation in a different way” [Hirst, 2003]
Paraphrase Definitions

“convey the **same meaning** but use different words”  [Bhagat and Hovy, 2013]

**Sentence Compression**
“generate a shorter paraphrase of a sentence”  
[Filippova et al., 2013]

“restatements with **approximately the same meaning**”  [Wang et al., 2019]

**Reframing**
“shift in perspective without distorting the underlying meaning”  
[Ziems et al., 2022]

“talking about the same situation **in a different way**”  [Hirst, 2003]
Creating a Paraphrasing Task Taxonomy

Extensive literature research to find tasks that

1) have been explicitly defined as paraphrasing tasks or
2) align with common paraphrase definitions

Classifying found tasks as

- Semantically *equivalent* paraphrasing
- Semantically *similar* paraphrasing

Overall, we found 25 paraphrasing tasks and arranged them in a task taxonomy.
Paraphrasing Task Taxonomy

Taxonomy of semantically equivalent paraphrasing tasks:

- Paraphrase generation task
  - Semantically equivalent paraphrasing
  - Semantically similar paraphrasing

- Copy editing
  - Improvement of coherence
  - Text simplification
  - Sentence compression
  - Sentence expansion

- Data augmentation
  - Adversarial example generation

- Linguistic steganography
  - Acrostification
  - Natural language watermarking

- Style adjustment
  - Author obfuscation
  - Plagiarizing
  - Style transfer

Taxonomy of semantically similar paraphrasing tasks:

- Paraphrase generation task
  - Semantically equivalent paraphrasing
  - Semantically similar paraphrasing

- Context change
  - Image recaptioning
  - Positive reframing
  - Text localization

- Conversational interaction
  - Argument repetition
  - Question dodging
  - Rogerian rhetoric application
  - Utterance clarification

- Textual entailment generation

- Information disguise

- Query suggestion

© marcel_gohsen 2024
Paraphrasing Task Taxonomy
Semantically Equivalent Paraphrasing

Copy editing
- Improvement of coherence
- Text simplification
- Sentence compression
- Sentence expansion

Data augmentation
- Adversarial example generation

Linguistic steganography
- Acrostification
- Natural language watermarking

Style adjustment
- Author obfuscation
- Plagiarizing
- Style transfer

Improvement of coherence
Paraphrasing to ease reading across sentences.

Original:  
In the lungs, carbon dioxide leaves the circulating blood and oxygen enters it.

Paraphrase:  
In the lungs, carbon dioxide that has been collected from cells as blood has passed around the body, leaves the circulating blood and oxygen enters it.
Text simplification

Improving readability of text through meaning-preserving rewrites. [Maddela et al., 2021]

Original: He settled in London, devoting himself chiefly to practical teaching.

Paraphrase: He lived in London. He was a teacher.
Sentence compression / expansion

Generating a shorter / longer paraphrase of a sentence. [Filippova et al., 2013]

Original: The future of the nation is in your hands.
Paraphrase: The nation’s future is in your hands.
Adversarial example generation

Modifying a text preserving the ground truth label but changing the prediction label. [Szegedy et al., 2014]

Original:  *There is no pleasure in watching a child suffer.*

Paraphrase:  *In watching the child suffer, there is no pleasure.*

Predicted sentiment:  ■ negative  ■ positive
Paraphrasing Task Taxonomy
Semantically Equivalent Paraphrasing

Acrostification

Rewriting a text such that it encodes an acrostic. [Stein et al., 2014]

Original:  
*To achieve your dreams, stay optimistic and persistent despite doubts. Embrace high expectations and let your light shine.*

Paraphrase:  
H old onto your dream while mindful of time  
O ptimism required, let your light shine  
P ersistence prevails, while it may cast doubt  
E xpectation desired is what it’s about.*
Author obfuscation

Preventing detection of whether two texts were written by the same author (authorship verification) by altering the text style. [Bevendorff et al., 2019]

Original:  
*Three billion people watching the home fleet take off, knowing the skies were open for all the hell that a savage enemy could send!*

Paraphrase:  
*Three billion people watching the home fleet take off, deciding the skies were resort for all the mischief that a savage enemy could send!*

Detected author:  
- [ ] Lester del Rey  
- [ ] ?
Paraphrasing Task Taxonomy

Semantically Equivalent Paraphrasing

Paraphrase generation task → Semantically equivalent paraphrasing → Semantically similar paraphrasing

Copy editing
- Improvement of coherence
- Text simplification
- Sentence compression
- Sentence expansion

Data augmentation
- Adversarial example generation

Linguistic steganography
- Acrostification
- Natural language watermarking

Style adjustment
- Author obfuscation
- Plagiarizing
- Style transfer

Style transfer

Changing the style of a text while preserving its semantics. [Krishna et al., 2020]

Original:  

*Yall kissing before marriage?*

Paraphrase:  

*And you kiss’d before nuptial?*

Style:  

- Tweet
- Shakespeare
Paraphrasing Task Taxonomy

Semantically Similar Paraphrasing

Paraphrase generation task ← Semantically equivalent paraphrasing ← Semantically similar paraphrasing

Context change
- Image recaptioning
- Positive reframing
- Text localization

Conversational interaction
- Argument repetition
- Question dodging
- Rogerian rhetoric application
- Utterance clarification

Textual entailment generation

Information disguise

Query suggestion

Image recaptioning

Assigning a new image caption to fit an image into a new context.

Original: Twelfth century illustration of a man digging.

Paraphrase: An English serf at work digging, circa 1170.

Context:
- Digging
- Economics of English agriculture in the Middle Ages
Positive reframing

Shifting to positive perspective without distorting the underlying meaning.

[Ziems et al., 2022]

Original:  This was a bland dish.

Paraphrase: I’ve made dishes that are much tastier than this one.
**Paraphrasing Task Taxonomy**

**Semantically Similar Paraphrasing**

- Paraphrase generation task
- Semantically equivalent paraphrasing
- Semantically similar paraphrasing

**Argument repetition**

Repeating the same argument within a discourse to make it more convincing.

**Original:**  *The movie “Die Hard” deserves an Oscar.*

**Paraphrase:**  *Other films have potential, but they do not deserve an Oscar like “Die Hard” does.*
Textual entailment generation

Generating a text which is implied by an original text.

Original: A woman is doing a cartwheel in the sand next to the beach.

Paraphrase: A woman is doing a cartwheel.

Original ⊨ Paraphrase
Query expansion/suggestion

Suggesting search queries that retain the original search intent.

Original:  *why do we yawn*

Paraphrase:  *why do we yawn so much*
Paraphrasing Task Classification
Experiment: Human Paraphrase Classification

Considered tasks and datasets:

- **Sentence compression**
  - Google Sentence Compression [Filippova and Altun, 2013]
  - Microsoft Abstractive Compression [Toutanova et al., 2016]

- **Sentence simplification**
  - TurkCorpus [Xu et al., 2016]
  - WikiLarge [Zhang and Lapata, 2017]

- **Style transfer**
  - ParaDetox [Logacheva et al., 2022]
  - Bible style transfer [Carlson et al., 2018]

- **Image recaptioning**
  - MSCOCO [Lin et al., 2015]
  - VizWiz [Gurari et al., 2020]

- **Textual entailment generation**
  - SciTail [Bowman et al., 2015]
  - HELP [Yanaka et al., 2019]
Sampling: 50 random paraphrases per dataset with 100 to 180 characters each. Performance: F1=0.73
Paraphrasing Task Classification
Experiment: Machine Paraphrase Classification

Paraphrase sampling:
- 50,000 task-specific paraphrase pairs
- 80:20 split of training and testing data
- Even distribution of examples from the five paraphrasing tasks

Feature engineering:
- Lexical similarity: ROUGE1, BLEU
- Syntax features: POS tag n-gram frequencies
- Semantic similarity: Sentence-BERT

Train a classifier:
- Random Forest classifier
- Max. depth of 15
Paraphrasing Task Classification

Machine Performance

Sampling: 1,000 random paraphrases per dataset with 100 to 180 characters each.
Performance: Test set: F1=0.81, 5-fold cross-validation: F1=0.82
Paraphrasing Task Classification
Experiment: Paraphrase Corpora Inhomogeneity

Analyzing inhomogeneity of known paraphrase corpora:

- Microsoft Research Paraphrase dataset (MSRPC) [Dolan and Brockett, 2005]
- TaPaCo [Scherrer, 2020]
- PAWS [Zhang et al., 2019]
- Wikipedia-IPC\textsubscript{silver} [Gohsen et al., 2023]
Paraphrasing Task Classification

Experiment: Paraphrase Corpora Inhomogeneity

Analyzing inhomogeneity of known paraphrase corpora:

- Microsoft Research Paraphrase dataset (MSRPC) [Dolan and Brockett, 2005]
- TaPaCo [Scherrer, 2020]
- PAWS [Zhang et al., 2019]
- Wikipedia-IPC\textsubscript{silver} [Gohsen et al., 2023]

<table>
<thead>
<tr>
<th>Paraphrase Dataset</th>
<th>Image Recaptioning</th>
<th>Sentence Compression</th>
<th>Sentence Simplification</th>
<th>Style Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSRPC</td>
<td>6.7%</td>
<td>32.0%</td>
<td>38.6%</td>
<td>11.4%</td>
</tr>
<tr>
<td>PAWS</td>
<td>5.2%</td>
<td>24.7%</td>
<td>62.7%</td>
<td>3.7%</td>
</tr>
<tr>
<td>TaPaCo</td>
<td>1.8%</td>
<td>8.4%</td>
<td>1.0%</td>
<td>76.8%</td>
</tr>
<tr>
<td>Wikipedia-IPC\textsubscript{silver}</td>
<td>16.3%</td>
<td>62.0%</td>
<td>19.8%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Total</td>
<td>8.6%</td>
<td>34.1%</td>
<td>17.3%</td>
<td>33.5%</td>
</tr>
</tbody>
</table>
Conclusions

Contributions:
- Extensive paraphrase literature review
- Paraphrasing task taxonomy grounded in the literature
- Paraphrasing classifier labeling one of five tasks to a paraphrase

Findings:
- Humans distinguish task-specific paraphrases with F1=0.73
- Machines distinguish task-specific paraphrases with F1=0.82
- Paraphrase corpora are biased towards different paraphrasing tasks

Future work:
- Classifier: More tasks, multi-label, test on unseen task-specific paraphrases
- Investigate paraphrase recognition bias for specific tasks
Conclusions

Contributions:

- Extensive paraphrase literature review
- Paraphrasing task taxonomy grounded in the literature
- Paraphrasing classifier labeling one of five tasks to a paraphrase

Findings:

- Humans distinguish task-specific paraphrases with F1=0.73
- Machines distinguish task-specific paraphrases with F1=0.82
- Paraphrase corpora are biased towards different paraphrasing tasks

Future work:

- Classifier: More tasks, multi-label, test on unseen task-specific paraphrases
- Investigate paraphrase recognition bias for specific tasks

Thank you!