Paraphrase Recognition via Dissimilarity Significance Classification

Long Qiu, Min-Yen Kan and Tat-Seng Chua

School of Computing
National University of Singapore
Paraphrase

Sentences that are “semantically equivalent” are a (Sentence-Level) Paraphrase.

- “It’s probably not the easiest time to take over the shuttle program,” he added, “but I look forward to the challenge.”

- “It is probably not the easiest time to come in and take over the shuttle program, but then again, I look forward to the challenge,” he said.
Sentences that are “semantically equivalent” are a (Sentence-Level) Paraphrase.

- “It’s probably not the easiest time to take over the shuttle program,” he *added*, “but I look forward to the challenge.”

- “It is probably not the easiest time to come in and take over the shuttle program, but then again, I look forward to the challenge,” he *said*.
Paraphrase

Sentences that are “semantically equivalent” are a (Sentence-Level) Paraphrase.

- “It’s probably not the easiest time to take over the shuttle program,” he added, “but I look forward to the challenge.”
- “It is probably not the easiest time to come in and take over the shuttle program, but then again, I look forward to the challenge,” he said.
Related Work

- Bag-of-words: Corley and Mihalcea, 2005
- Sequence of Tokens: Barzilay and Lee, 2003 (Multiple-Sequence Alignment)
- Syntactic Tree: Wu, 2005 (Inversion Transduction Grammar)
Related Work

- Bag-of-words: Corley and Mihalcea, 2005
- Sequence of Tokens: Barzilay and Lee, 2003 (Multiple-Sequence Alignment)
- Syntactic Tree: Wu, 2005 (Inversion Transduction Grammar)
Motivation

- Paraphrase (+pp):
  - “It’s probably not the easiest time to take over the shuttle program,” he *added*, “but I look forward to the challenge.”
  - “It is probably not the easiest time to come in and take over the shuttle program, but then again, I look forward to the challenge,” he *said*. 

Non-Paraphrase (-pp):

- The technology-laced Nasdaq Composite Index added 1.92 points, or 0.12 percent, at 1,647.94.
- The technology-laced Nasdaq Composite Index dipped 0.08 of a point to 1,646.
Motivation

• Paraphrase ($+pp$):
  • “It’s probably not the easiest time to take over the shuttle program,” he added, “but I look forward to the challenge.”
  • “It is probably not the easiest time to come in and take over the shuttle program, but then again, I look forward to the challenge,” he said.
Motivation

• Paraphrase (+pp):
  • “It’s probably not the easiest time to take over the shuttle program,” he added, “but I look forward to the challenge.”
  • “It is probably not the easiest time to come in and take over the shuttle program, but then again, I look forward to the challenge,” he said.

• Non-Paraphrase (−pp):
  • The technology-laced Nasdaq Composite Index added 1.92 points, or 0.12 percent, at 1,647.94.
  • The technology-laced Nasdaq Composite Index dipped 0.08 of a point to 1,646.
Motivation

Two attributes of paraphrasing sentences:

- **Similarity**: they share a substantial amount of information nuggets;
- **Dissimilarities**: extraneous words, predicate argument tuples, sentences.
Motivation

Two attributes of paraphrasing sentences:

- **similarity**: they share a substantial amount of information nuggets;

  *dissimilarities are extraneous*: if extra information in the sentences exists, the effect of its removal is not significant.
Motivation

Two attributes of paraphrasing sentences:

- **similarity**: they share a substantial amount of information nuggets;

- **dissimilarities are extraneous**: if extra information in the sentences exists, the effect of its removal is not significant.
Motivation

Two attributes of paraphrasing sentences:

- **similarity**: they share a substantial amount of information nuggets;
- **dissimilarities are extraneous**: if extra information in the sentences exists, the effect of its removal is not significant.

Information nuggets $\leftrightarrow$ words, predicate argument tuples, sentences.
Motivation

Two attributes of paraphrasing sentences:

- **similarity**: they share a substantial amount of information nuggets;
- **dissimilarities are extraneous**: if extra information in the sentences exists, the effect of its removal is not significant.

Information nuggets $\leadsto$ **words**, predicate argument tuples, sentences.
Outline

- Related Work
- Motivation
- Two-phase Framework
- Evaluation
- Discussion
0. Preprocessing
0. Preprocessing

1. Phase 1: Similarity Detector

- **Preprocessing:** Sentence Pairs
  - Sentence pairs are preprocessed to prepare for further analysis.

- **Phase 1:** Similarity Detector
  - Tuples Labeled
    - Tuples labeled with predicate and argument annotations.
  - Tuples Matched
    - Some tuples remain unpaired after matching.

- **Phase 2:** Dissimilarity Classifier
  - Further analysis to determine dissimilarity.

- **Paraphrase Judgement**
  - Final step to judge the paraphrase quality.
0. Preprocessing

1. Phase 1: Similarity Detector

2. Phase 2: Dissimilarity Significance Classifier

Diagram:

- Thesaurus
- Charniak Parser & ASSERT
- Sentence Pairs
- Predicate Argument Tuples Labeled
- Predicate Argument Tuples Matched (some remain unpaired)
- Dissimilarity Classifier
- Paraphrase Judgement

Flow:

- Preprocessing:
- Similarity Detector
- Dissimilarity Significance Classifier
0. Preprocessor: Semantic Role Labeler

Authorities said a young man injured Richard Miller.
0. Preprocessor: Semantic Role Labeler

Authorities said a young man injured Richard Miller.

- Authorities$_{ARG0}$ said$_{PREDICATE}$ a young man injured$_{ARG1}$ Richard Miller$_{ARG1}$
- Authorities said a young man$_{ARG0}$ injured$_{PREDICATE}$ Richard Miller$_{ARG1}$
1. Similarity Detector (SD)

- The technology-laced Nasdaq Composite Index added 1.92 points or 0.12 percent at 1,647.94
- Rose 0.08 of a point to 1,646
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94
- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET rose] [ARG2 0.08 of a point] to [ARG4 1,646]
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]
1. Similarity Detector (SD)

- [**ARG1** The technology-laced Nasdaq Composite Index] [**TARGET added**] [**ARG2** 1.92 points or 0.12 percent] at 1,647.94
- [**ARG1** The technology-laced Nasdaq Composite Index] [**TARGET dipped**] [**ARG2** 0.08 of a point] to [**ARG4** 1,646]
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94
- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]

It’s probably not the easiest time to take over the shuttle program he added but I look forward to the challenge.
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94
- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]

- It’s probably not [ARGM-TMP the easiest time] to [TARGET take] over [ARG1 the shuttle program] he added but I look forward to the challenge.
- It is probably not the easiest time to come in and [TARGET take] over [ARG1 the shuttle program] but then again I look forward to the challenge he said.
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94
- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]

- [ARG0 It’s probably not [ARGM-TMP the easiest time] to [TARGET take] over [ARG1 the shuttle program] he added but I look forward to the challenge
- It is probably not the easiest time to come in and [TARGET take] over [ARG1 the shuttle program] but then again I look forward to the challenge he said
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94
- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]

- [ARG0 It]’s probably not [ARGM-TMP the easiest time] to [TARGET take] over [ARG1 the shuttle program] he added but I look forward to the challenge
- It is probably not the easiest time to [TARGET come] [ARGM-DIR in] and [TARGET take] over [ARG1 the shuttle program] but then again I look forward to the challenge he said
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added] [ARG2 1.92 points or 0.12 percent] at 1,647.94
- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]

- [ARG0 It] ’s probably not [ARGM-TMP the easiest time] to [TARGET take] over [ARG1 the shuttle program] he added but I look forward to the challenge
- It is probably not the easiest time to [TARGET come] [ARGM-DIR in] and [TARGET take] over [ARG1 the shuttle program] but then again I look forward to the challenge he said
1. Similarity Detector (SD)

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET added ] [ARG2 1.92 points or 0.12 percent] at 1,647.94

- [ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped ] [ARG2 0.08 of a point] to [ARG4 1,646]

- [ARG0 It]’s probably not [ARGM-TMP the easiest time] to [TARGET take ] over [ARG1 the shuttle program] he added but I look forward to the challenge

- It is probably not the easiest time to [TARGET come ] [ARGM-DIR in] and [TARGET take ] over [ARG1 the shuttle program] but then again I look forward to the challenge he said
The technology-laced Nasdaq Composite Index dipped 0.08 of a point to 1,646.

2. Dissimilarity Classifier (DC)
2. Dissimilarity Classifier (DC)

[ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]

It is probably not the easiest time to [TARGET come] [ARGM-DIR in] and take over the shuttle program but then again I look forward to the challenge he said.
2. Dissimilarity Classifier (DC)

[ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped ] [ARG2 0.08 of a point] to [ARG4 1,646]
It is probably not the easiest time to [TARGET come ] [ARGM-DIR in] and take over the shuttle program but then again I look forward to the challenge he said

Automatically annotated instances:

- *Insignificant Tuples* in paraphrasing sentence pairs where only one sentence has extra tuples;
- *Significant Tuples* in non-paraphrasing sentence pairs where only one sentence has only one extra tuple.
2. Dissimilarity Classifier (DC)

[ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped ] [ARG2 0.08 of a point] to [ARG4 1,646]
It is probably not the easiest time to [TARGET come ] [ARGM-DIR in] and take over the shuttle program but then again I look forward to the challenge he said

Machine Learner : Support Vector Machine

• Linear Kernel
2. Dissimilarity Classifier (SVM)

[ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped] [ARG2 0.08 of a point] to [ARG4 1,646]
2. Dissimilarity Classifier (SVM)

[ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped ] [ARG2 0.08 of a point] to [ARG4 1,646]

Internal Features:

- # numeric expressions: 2
- # named entities: 1
- # words: 12
- # semantic roles: 4
- similar to other tuples in the same sentence: false
2. Dissimilarity Classifier (SVM)

[ARG1 The technology-laced Nasdaq Composite Index] [TARGET dipped ] [ARG2 0.08 of a point] to [ARG4 1,646]

Contextual Features:

- *hosting/opposing sentence length*: 13/14
- *# paired tuples*: 0
- *etc.*
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

• come [↑ VB, ↑ VP, − CC, ↓ VP, ↓ VB] take
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

- come [↑ VB, ↑ VP, ¬ CC, ↓ VP, ↓ VB] take
  - come
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

- \( \text{come \; in} \) \text{ADVP}\;
- \( \text{and} \) \text{ADVP}\;
- \( \text{over} \) \text{ADVP}\;
- \( \text{the} \) \text{ADVP}\;
- \( \text{shuttle} \) \text{ADVP}\;
- \( \text{program} \) \text{ADVP}\;

- \( \text{come \; take} \) \text{ADVP}\;
- \( \text{in} \) \text{ADVP}\;
- \( \text{take} \) \text{ADVP}\;
- \( \text{over} \) \text{ADVP}\;
- \( \text{the} \) \text{ADVP}\;
- \( \text{shuttle} \) \text{ADVP}\;
- \( \text{program} \) \text{ADVP}\;

- \( \text{take} \) \text{ADVP}\;
- \( \text{come} \) \text{ADVP}\;
- \( \uparrow V B \)
- \( \uparrow V P \)
- \( \uparrow V B \uparrow V P \)
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
come [↑VB, ↑VP, ↓CC, ↓VP, ↓VB] take
  • come
  • ↑VB
  • ↑VB ↑VP
  • ↑VB ↑VP − CC
```
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

\[
\text{come \ 
\uparrow VB, \ 
\uparrow VP, \ 
\rightarrow CC, \ 
\downarrow VP, \ 
\downarrow VB} \text{ take}
\]

- come
- \(\uparrow VB\)
- \(\uparrow VB \uparrow VP\)
- \(\uparrow VB \uparrow VP - CC\)
- \(\uparrow VB \uparrow VP - CC \downarrow VP\)
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

- come $[\uparrow VB, \uparrow VP, -CC, \downarrow VP, \downarrow VB]$ take
  - come
  - $\uparrow VB$
  - $\uparrow VB \uparrow VP$
  - $\uparrow VB \uparrow VP - CC$
  - $\uparrow VB \uparrow VP - CC \downarrow VP$
  - ...
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
VP
  / 
VP  CC
  /  /
VB  ADVP   VB  PRT
  /   /
RB  RB     RP  DT
```

```
come [↑ VB, ↑ VP, -CC, ↓ VP, ↓ VB] take
- come
- ↑ VB
- ↑ VB ↑ VP
- ↑ VB ↑ VP - CC
- ↑ VB ↑ VP - CC ↓ VP
- ...
```
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

- come \([\uparrow VB, \uparrow VP, \neg CC, \downarrow VP, \downarrow VB]\) take
  - come
  - \(\uparrow VB\)
  - \(\uparrow VB \uparrow VP\)
  - \(\uparrow VB \uparrow VP - CC\)
  - \(\uparrow VB \uparrow VP - CC \downarrow VP\)
  - ...

2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
come [↑ VB, ↑ VP, ¬CC, ↓ VP, ↓ VB] take
- come
- ↑ VB
- ↑ VB ↑ VP
- ↑ VB ↑ VP - CC
- ↑ VB ↑ VP - CC ↓ VP
- ...
```

```
come in and take over the shuttle program
```
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
V P
  / \  / \
VP  CC VP
  / \  / |
VB ADVP VB PRT NP
  / \ /  |
RB RP DT NN NN
```

- come \([\uparrow VB, \uparrow VP, \not CC, \downarrow VP, \downarrow VB]\) take
  - come
  - \(\uparrow VB\)
  - \(\uparrow VB \uparrow VP\)
  - \(\uparrow VB \uparrow VP - CC\)
  - \(\uparrow VB \uparrow VP - CC \downarrow VP\)
  - ...

2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
come [↑ VB, ↑ VP, ~CC, ↓ VP, ↓ VB] take
  - come
  - ↑ VB
  - ↑ VB ↑ VP
  - ↑ VB ↑ VP - CC
  - ↑ VB ↑ VP - CC ↓ VP
  - ...
```
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
```
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
come [↑ VB, ↑ VP, –CC, ↓ VP, ↓ VB] take
  • come
  • ↑ VB
  • ↑ VB ↑ VP
  • ↑ VB ↑ VP – CC
  • ↑ VB ↑ VP – CC ↓ VP
  • ...
```
2. Dissimilarity Classifier (SVM)

Features that show performance gain:

- lemma of the predicate;
- n-grams from Syntactic Parse Path.

```
<table>
<thead>
<tr>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP</td>
</tr>
<tr>
<td>CC</td>
</tr>
<tr>
<td>VP</td>
</tr>
<tr>
<td>VB</td>
</tr>
<tr>
<td>VB</td>
</tr>
<tr>
<td>NP</td>
</tr>
<tr>
<td>RB</td>
</tr>
<tr>
<td>RP</td>
</tr>
<tr>
<td>NN</td>
</tr>
</tbody>
</table>

come [↑ VB, ↑ VP, − CC, ↓ VP, ↓ VB] take
- come
- ↑ VB
- ↑ VB ↑ VP
- ↑ VB ↑ VP − CC
- ↑ VB ↑ VP − CC ↓ VP
- ...
- ↓ VB
```

Paraphrase Judgement

IF Sentence pairs with perfectly paired tuples
    THEN Paraphrase

ELSE {
    IF Sentence pairs with insignificant unpaired tuples
        THEN Paraphrase
    IF Sentence pairs with significant unpaired tuples
        THEN Non-paraphrase
}
Outline

• Related Work
• Motivation
• Two-phase Framework
• Evaluation
• Discussion
Evaluation

• Goals of the evaluation: do they work?
  1. Similarity Detector (SD)
  2. Dissimilarity Classifier (DC)
  3. The whole PR system (SD + DC)

• Data Set: Microsoft Research Paraphrase Corpus
  • 4076 sentence pairs in training set (2753 +pp)
  • 1725 sentence pairs in test set (1147 +pp)
1. Similarity Detection

- Statistics for 200 further annotated sentence pairs in the test set (200set):

<table>
<thead>
<tr>
<th>Description</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td># sentence pairs with tuple pairs (by SD)</td>
<td>157</td>
</tr>
<tr>
<td># correctly paired (annotators agree)</td>
<td>144</td>
</tr>
<tr>
<td># sentence pairs with missed tuple pairs (by annotators)</td>
<td>31</td>
</tr>
</tbody>
</table>
1. Similarity Detection

Statistics for 200 further annotated sentence pairs in the test set (200set):

<table>
<thead>
<tr>
<th>Description</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td># sentence pairs with tuple pairs (by SD)</td>
<td>157</td>
</tr>
<tr>
<td># correctly paired (annotators agree)</td>
<td>144</td>
</tr>
<tr>
<td># sentence pairs with missed tuple pairs (by annotators)</td>
<td>31</td>
</tr>
</tbody>
</table>

- **Precision** = \( \frac{Positive_{true} \cap Positive_{system}}{Positive_{system}} \) = \( \frac{144}{157} \) = 92%

- **Recall** = \( \frac{Positive_{true} \cap Positive_{system}}{Positive_{true}} \) = \( \frac{144}{144 + 31} \) = 82%
2. Dissimilarity Classification

- insignificant tuples well captured
- significant tuples evenly distributed

<table>
<thead>
<tr>
<th>SVM Prediction</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; .5</td>
<td>0</td>
</tr>
<tr>
<td>.5 - .25</td>
<td>20</td>
</tr>
<tr>
<td>.25 - 0</td>
<td>40</td>
</tr>
<tr>
<td>0 - .25</td>
<td>60</td>
</tr>
<tr>
<td>.25 - .5</td>
<td>80</td>
</tr>
<tr>
<td>.5 - .75</td>
<td>100</td>
</tr>
<tr>
<td>.75 - 1</td>
<td>120</td>
</tr>
<tr>
<td>&lt; 1</td>
<td>140</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Insignificant</th>
<th>Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>significant</td>
<td>112</td>
<td>33</td>
</tr>
<tr>
<td>insignificant</td>
<td>263</td>
<td>5</td>
</tr>
</tbody>
</table>

Insignificant by classifier

Significant by classifier
3. Overall

The system’s ability of pinpointing paraphrase barriers:

- In the 200set, 55 $-pp$ cases are correctly recognized;
- For 40 (73%), significant unpaired tuples are agreed to be the reason for non-paraphrasing by human.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Overall Performance (100% of Test set)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acc</td>
</tr>
<tr>
<td>Majority Classifier</td>
<td>66.5%</td>
</tr>
<tr>
<td>SimFinder</td>
<td>72.9%</td>
</tr>
<tr>
<td>CM05</td>
<td>71.5%</td>
</tr>
<tr>
<td>Our System</td>
<td>72.0%</td>
</tr>
</tbody>
</table>
Outline

- Related Work
- Motivation
- Two-phase Framework
- Evaluation
- Discussion
Discussion & Future Work

- significant tuples appear as:
  - (40%) The nucleus of the sentence (often the matrix tuple):
    Michael Hill, a Sun reporter who is a member of the
    Washington-Baltimore Newspaper Guild’s bargaining
    committee, estimated meetings to last late Sunday.
Discussion & Future Work

- significant tuples appear as:
  - (40%) The nucleus of the sentence (often the matrix tuple):
  - (30%) A part of a coordination:
    *Security lights have also been installed and police have swept the grounds for booby traps.*
Discussion & Future Work

- significant tuples appear as:
  - (40%) The nucleus of the sentence (often the matrix tuple):
  - (30%) A part of a coordination:
  - (13%) A predicate of a modifying clause:
    Westermayer was 26 then, and a friend and former manager who knew she was unhappy in her job tipped her to another position.
Discussion & Future Work

- **significant tuples appear as:**
  - (40%) The nucleus of the sentence (often the matrix tuple):
  - (30%) A part of a coordination:
  - (13%) A predicate of a modifying clause:
  - (7%) An adjunct: *While waiting for a bomb squad to arrive, the bomb exploded, killing Wells.*
Discussion & Future Work

**significant tuples appear as:**

- (40%) The nucleus of the sentence (often the matrix tuple):
- (30%) A part of a coordination:
- (13%) A predicate of a modifying clause:
- (7%) An adjunct:
- (7%) An embedded sentence:

Dean told reporters traveling on his 10-city “Sleepless Summer” tour that he considered campaigning in Texas a challenge.
Discussion & Future Work

**significant tuples appear as:**

- (40%) The nucleus of the sentence (often the matrix tuple):
- (30%) A part of a coordination:
- (13%) A predicate of a modifying clause:
- (7%) An adjunct:
- (7%) An embedded sentence:
- (3%) Or factual content that conflicts with the opposing sentence:

*Total sales for the period declined 8.0 percent to USD1.99 billion from a year earlier.*

*Wal-Mart said sales at stores open at least a year rose 4.6 percent from a year earlier.*
Discussion & Future Work

- **significant** tuples appear as:
  - (40%) The nucleus of the sentence (often the matrix tuple):
  - (30%) A part of a coordination:
  - (13%) A predicate of a modifying clause:
  - (7%) An adjunct:
  - (7%) An embedded sentence:
  - (3%) Or factual content that conflicts with the opposing sentence:
Discussion & Future Work (Cont’d)

• Problematic cases
  • Non-literal language issues such as implicature, idiom, metaphor, *etc.* are not addressed in our current system:

  \[ +pp \]

  *Later in the day, a standoff developed between French soldiers and a Hema battlewagon that attempted to pass the UN compound.*

  *French soldiers later threatened to open fire on a Hema battlewagon that tried to pass near the UN compound.*
Discussion & Future Work (Cont’d)

- Problematic cases
  - Non-literal language issues such as implicature, idiom, metaphor, etc. are not addressed in our current system;
  - A paraphrasing pair may exceed the system’s threshold for syntactic difference:
    +pp
    With the exception of dancing, physical activity did not decrease the risk.
    Dancing was the only physical activity associated with a lower risk of dementia.
Discussion & Future Work (Cont’d)

- Problematic cases
  - Non-literal language issues such as implicature, idiom, metaphor, etc. are not addressed in our current system;
  - A paraphrasing pair may exceed the system’s threshold for syntactic difference;
  - One or more unpaired tuples exist, but their significance is not inferred correctly:
    \[ +pp \]
    Inhibited children tend to be timid with new people, objects, and situations, while uninhibited children spontaneously approach them.
    Simply put, shy individuals tend to be more timid with new people and situations.
Discussion & Future Work (Cont’d)

- Problematic cases
  - Non-literal language issues such as implicature, idiom, metaphor, etc. are not addressed in our current system;
  - A paraphrasing pair may exceed the system’s threshold for syntactic difference;
  - One or more unpaired tuples exist, but their significance is not inferred correctly.
Conclusion

• Proposed a PR framework focusing on dissimilarity
  • Similarity Detector: Matches similar tuples and detects extra ones;
  • Dissimilarity Classifier: Judges whether extra tuples are significant.

• Implemented a system that shows:
  • what information makes the sentences non-paraphrasing;
  • high accuracy in matching similar tuples;
  • robust dissimilarity classification;
  • comparable overall PR performance.