Recognizing Implied Predicate-Argument Relationships in Textual Inference

Asher Stern and Ido Dagan

Bar Ilan University, Ramat Gan, Israel

ACL 2014 (extended version)
Motivation: implied predicate-argument relationships

Example

Prius sales plunged after Toyota’s recall announcement.
Motivation: implied predicate-argument relationships

Example

_Prius_ sales plunged after Toyota’s _recall_ announcement.

- The predicate-argument relationship “recall”–“Prius” is implied from the text.
Motivation: implied predicate-argument relationships

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- The predicate-argument relationship “recall”–“Prius” is implied from the text.
- Not expressed in the syntactic structure.
Motivation: implied predicate-argument relationships

Example

*Prius* sales plunged after Toyota’s *recall* announcement.

- The predicate-argument relationship “recall”–“Prius” is implied from the text.
- Not expressed in the syntactic structure.
  - Not detected by parsers.
Motivation: implied predicate-argument relationships

Example

_Prius_ sales plunged after Toyota’s _recall_ announcement.

- The predicate-argument relationship “recall”—“Prius” is implied from the text.
- Not expressed in the syntactic structure.
  - Not detected by parsers.
  - Mostly beyond SRLs.
Motivation: implied predicate-argument relationships

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

Why is this important?
Motivation: implied predicate-argument relationships

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

Why is this important?

- Question Answering (QA)
Motivation: implied predicate-argument relationships

Example

Prius sales plunged after Toyota’s recall announcement.

Why is this important?
- Question Answering (QA)
  - Which model has been recalled?
Motivation: implied predicate-argument relationships

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

Why is this important?

- Question Answering (QA)
  - Which model has been recalled?
- Recognizing Textual Entailment (RTE)
Motivation: implied predicate-argument relationships

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

Why is this important?

- **Question Answering (QA)**
  - Which model has been recalled?
- **Recognizing Textual Entailment (RTE)**
  - Hypothesis “Toyota recalled Prius”.
Motivation: implied predicate-argument relationships

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

Why is this important?

- Question Answering (QA)
  - Which model has been recalled?
- Recognizing Textual Entailment (RTE)
  - Hypothesis “Toyota recalled Prius”.
- Information Extraction (IE)
Motivation: implied predicate-argument relationships

Example

 Prius sales plunged after Toyota’s recall announcement.

Why is this important?

- Question Answering (QA)
  - Which model has been recalled?
- Recognizing Textual Entailment (RTE)
  - Hypothesis “Toyota recalled Prius”.
- Information Extraction (IE)
  - Recall (\textit{Firm}, \textit{Model})
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse

Recall:

Agent = "Toyota"
Theme = "Prius"

Empirically very difficult

$\text{F1} < 1.5\%$ in SemEval challenge

State of the art F1 $< 20\%$
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse

Recall:
Agent = “Toyota”
Theme = “Prius”
...
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse

Recall:
Agent = “Toyota”
Theme = “Prius”
...

- Annotate everything, regardless of application needs
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse

Recall:
Agent = “Toyota”
Theme = “Prius”
...

- Annotate everything, regardless of application needs
- Empirically very difficult
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse

Recall:
Agent = “Toyota”
Theme = “Prius”
...

- Annotate everything, regardless of application needs
- Empirically very difficult
  - F1 < 1.5% in SemEval challenge
Prior work: Labeling uninstantiated roles

(Ruppenhofer et al. 2010), (Gerber and Chai, 2012), (Silberer and Frank, 2012), (Roth and Frank, 2013), (Laparra and Rigau, 2012)

- SRL extension (SemEval 2010)
- Labeling syntactically-uninstantiated roles implied from discourse

Recall:
Agent = “Toyota”
Theme = “Prius”
...

- Annotate everything, regardless of application needs
- Empirically very difficult
  ▶ F1< 1.5% in SemEval challenge
  ▶ State of the art F1< 20%
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- Observation: in inference applications terms for the predicate and the argument are pre-detected.
Annotation vs. Recognition

Example

Prius sales plunged after Toyota’s recall announcement.

- Observation: in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- **Observation:** in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.
Annotation vs. Recognition

Example

Prius sales plunged after Toyota’s recall announcement.

- Observation: in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.

  “Toyota recalled Prius”
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- Observation: in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.
    - **Prius** sales plunged after Toyota’s **recall** announcement.
  - “Toyota recalled Prius”
- QA: match question’s predicate and argument answer type.
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- Observation: in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.
    - “**Prius** sales plunged after Toyota’s **recall** announcement.”
  - QA: match question’s predicate and argument answer type.
  - IE: predicate detection, argument entity extractor.
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- Observation: in inference applications, terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.
    - “**Prius** sales plunged after Toyota’s **recall** announcement.”
  - “**Toyota** recalled **Prius**”
  - QA: match question’s predicate and argument answer type.
  - IE: predicate detection, argument entity extractor.
Annotation vs. Recognition

Example

**Prius** sales plunged after Toyota’s **recall** announcement.

- Observation: in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.
    - Prius sales plunged after Toyota’s **recall** announcement.
  - “Toyota recalled Prius”
  - QA: match question’s predicate and argument answer type.
  - IE: predicate detection, argument entity extractor.

Recognition approach

Don’t annotate. Verify.
## Task Definition

**Text**

Prius sales plunged after Toyota’s recall announcement.

**Hypothesis**

Toyota recalled Prius.
Task Definition

Text
Prius sales plunged after Toyota’s recall announcement.

Hypothesis
Toyota recalled Prius.

Textual Entailment

Stern and Dagan
Recognizing Implied Predicate-Argument Relationships in Textual Inference
ACL 2014 5 / 13
## Task Definition

<table>
<thead>
<tr>
<th>Text</th>
<th>Prius sales plunged after Toyota’s recall announcement.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis</td>
<td>Toyota recalled Prius.</td>
</tr>
</tbody>
</table>

### Textual Entailment
- **Input:** Text and Hypothesis

Compared to the annotation task, recognition is more feasible. Recognition covers more cases. ▶ Details next.
Prius sales plunged after Toyota’s recall announcement.

Toyota recalled Prius.

**Textual Entailment**
- **Input:** Text and Hypothesis
- **Output:** Does the Text entail the Hypothesis?
Task Definition

Text

Prius sales plunged after Toyota’s recall announcement.

Hypothesis

Toyota recalled Prius.

Recognizing Implied Predicate Argument Relationships
Task Definition

Text

Prius sales plunged after Toyota’s recall announcement.

Hypothesis

Toyota recalled Prius.

Recognizing Implied Predicate Argument Relationships

- **Input**: Explicit relationship in the hypothesis (recall–Prius)
Task Definition

Text

Prius sales plunged after Toyota’s recall announcement.

Hypothesis

Toyota recalled Prius.

Recognizing Implied Predicate Argument Relationships

- **Input:** Explicit relationship in the hypothesis (recall–Prius)
- **Input:** Candidate Predicate and Candidate Argument (recall, Prius) in the Text
Task Definition

Text

| Prius sales plunged after Toyota’s recall announcement. |

Hypothesis

| Toyota recalled Prius. |

Recognizing Implied Predicate Argument Relationships

- **Input:** Explicit relationship in the hypothesis (recall–Prius)
- **Input:** Candidate Predicate and Candidate Argument (recall, Prius) in the Text
- **Output:** Does the Hypothesis relationship hold also in the Text?
**Task Definition**

**Text**

*Prius* sales plunged after Toyota’s *recall* announcement.

**Hypothesis**

Toyota *recalled* *Prius*.

**Recognizing Implied Predicate Argument Relationships**

- **Input:** Explicit relationship in the hypothesis (recall–*Prius*)
- **Input:** *Candidate Predicate* and *Candidate Argument* (recall, *Prius*) in the Text
- **Output:** Does the Hypothesis relationship hold also in the Text?

- Compared to the annotation task, recognition is more feasible.
Task Definition

Text

Prius sales plunged after Toyota’s recall announcement.

Hypothesis

Toyota recalled Prius.

Recognizing Implied Predicate Argument Relationships

- **Input**: Explicit relationship in the hypothesis (recall–Prius)
- **Input**: Candidate Predicate and Candidate Argument (recall, Prius) in the Text
- **Output**: Does the Hypothesis relationship hold also in the Text?

- Compared to the annotation task, recognition is more feasible.
- Recognition covers more cases.
Task Definition

Text

Prius sales plunged after Toyota’s recall announcement.

Hypothesis

Toyota recalled Prius.

Recognizing Implied Predicate Argument Relationships

- **Input**: Explicit relationship in the hypothesis (recall–Prius)
- **Input**: Candidate Predicate and Candidate Argument (recall, Prius) in the Text
- **Output**: Does the Hypothesis relationship hold also in the Text?

Compared to the annotation task, recognition is more feasible.
Recognition covers more cases.

Details next.
A negative example

Text
Sheehan’s protest is misguided and is hurting troop morale. . . .
Sheehan never wanted Casey to join the military.

Hypothesis
Barbara Cummings heard the tale of a woman who was coming to Crawford to join Cindy Sheehan’s protest.
Better coverage

- Some cases fall beyond the implied-SRL task
Better coverage

- Some cases fall beyond the implied-SRL task
- But are covered by our task
Better coverage

- Some cases fall beyond the implied-SRL task
- But are covered by our task
Better coverage

- Some cases fall beyond the implied-SRL task
- But are covered by our task

Modifiers (adjuncts)

5 days after he arrived in Iraq last year, Casey Sheehan was killed.
Better coverage

- Some cases fall beyond the implied-SRL task
- But are covered by our task

Modifiers (adjuncts)

5 days after he arrived in Iraq last year, Casey Sheehan was killed.

Filled roles

Hurricane Rita was upgraded from a tropical storm as it threatened the southeastern United States, forcing an alert in southern Florida and scuttling plans to repopulate New Orleans after Hurricane Katrina turned it into a ghost city 3 weeks earlier.
### Dataset

<table>
<thead>
<tr>
<th>Text</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Prius</em> sales plunged after Toyota’s <em>recall</em> announcement.</td>
<td>Toyota <em>recalled</em> <em>Prius</em>.</td>
</tr>
</tbody>
</table>
Dataset

Text

**Prius** sales plunged after Toyota’s **recall** announcement.

Hypothesis

**Toyota** **recalled** **Prius**.

- Semi-automatic dataset construction, based on RTE-6 dataset (Bentivogli et al., 2010)
Dataset

Text
Prius sales plunged after Toyota’s recall announcement.

Hypothesis
Toyota recalled Prius.

- Semi-automatic dataset construction, based on RTE-6 dataset (Bentivogli et al., 2010)
  - Explicit predicate-argument relationship in the Hypothesis
Stern and Dagan
Recognizing Implied Predicate-Argument Relationships in Textual Inference

Dataset

Text

**Prius** sales plunged after Toyota’s **recall** announcement.

Hypothesis

Toyota **recalled** **Prius**.

- Semi-automatic dataset construction, based on RTE-6 dataset (Bentivogli et al., 2010)
  - Explicit predicate-argument relationship in the Hypothesis
  - Candidate predicate and argument in the Text
Text

*Prius* sales plunged after Toyota’s *recall* announcement.

Hypothesis

Toyota *recalled* *Prius*.

- Semi-automatic dataset construction, based on RTE-6 dataset (Bentivogli et al., 2010)
  - Explicit predicate-argument relationship in the Hypothesis
  - Candidate predicate and argument in the Text
- Yes/No annotation by human annotator
Dataset

Text

Prius sales plunged after Toyota’s recall announcement.

Hypothesis

Toyota recalled Prius.

- Semi-automatic dataset construction, based on RTE-6 dataset (Bentivogli et al., 2010)
  - Explicit predicate-argument relationship in the Hypothesis
  - Candidate predicate and argument in the Text
- Yes/No annotation by human annotator
- Larger than all prior work datasets (4022 instances)
Dataset

Text

**Prius** sales plunged after Toyota’s **recall** announcement.

Hypothesis

Toyota **recalled** **Prius**.

- Semi-automatic dataset construction, based on RTE-6 dataset (Bentivogli et al., 2010)
  - Explicit predicate-argument relationship in the Hypothesis
  - Candidate predicate and argument in the Text
- Yes/No annotation by human annotator
- Larger than all prior work datasets (4022 instances)
- Random Forest learning algorithm
- 15 features
Random Forest learning algorithm

15 features

<table>
<thead>
<tr>
<th>#</th>
<th>Category</th>
<th>Feature</th>
<th>Prev. work</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>statistical discourse</td>
<td>co-occurring predicate</td>
<td>New</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>co-occurring argument</td>
<td>New</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>co-reference: whether an explicit argument of ( p ) co-refers with ( a ).</td>
<td>New</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>last known location: If the NE of ( a ) is “location”, and it is the last location mentioned before ( p ) in the document.</td>
<td>New</td>
</tr>
<tr>
<td>5</td>
<td>local discourse</td>
<td>argument prominence: The frequency of the lemma of ( a ) in a two-sentence windows of ( p ), relative to all entities in that window.</td>
<td>S&amp;F</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>predicate frequency in document: The frequency of ( p ) in the document, relative to all predicates appear in the document.</td>
<td>G&amp;C</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>statistical argument frequency: The Unigram-model likelihood of ( a ) in English documents, calculated from a large corpus.</td>
<td>New</td>
</tr>
<tr>
<td>8</td>
<td>local candidate properties</td>
<td>definite NP: Whether ( a ) is a definite NP</td>
<td>G&amp;C</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>indefinite NP: Whether ( a ) is an indefinite NP</td>
<td>G&amp;C</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>quantified predicate: Whether ( p ) is quantified (i.e., by expressions like “every . . .”, “a good deal of . . .”, etc.)</td>
<td>G&amp;C</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>NE mismatch: Whether ( a ) is a named entity but the corresponding argument in the hypothesis is not, or vice versa.</td>
<td>New</td>
</tr>
<tr>
<td>12</td>
<td>predicate-argument relatedness</td>
<td>predicate-argument frequency: The likelihood of ( a ) to be an argument of ( p ) (formally: ( Pr(a</td>
<td>p) )) in a large corpus.</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>sentence distance: The distance between ( p ) and ( a ) in sentences.</td>
<td>G&amp;C, S&amp;F</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>mention distance: The distance between ( p ) and ( a ) in entity-mentions.</td>
<td>S&amp;F</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>shared head-predicate: Whether ( p ) and ( a ) are themselves arguments of another predicate.</td>
<td>G&amp;C</td>
</tr>
</tbody>
</table>
Statistical Discourse Features

Co-occurring predicate
Statistical Discourse Features

Co-occurring predicate

Example

At least 10 people were killed . . . in the [crash]_{cand-pred} . . . Alvarez is accused of . . . causing the derailment of one [train]_{cand-arg} . . .
Statistical Discourse Features

Co-occurring predicate

Example

At least 10 people were killed . . . in the [crash]_{cand-pred} . . . Alvarez is accused of . . . causing the derailment of one [train]_{cand-arg} . . .

- Does derailment–train indicate crash–train?
Statistical Discourse Features

Co-occurring predicate

Example

At least 10 people were killed . . . in the [crash]_{cand-pred} . . . Alvarez is accused of . . . causing the derailment of one [train]_{cand-arg} . . .

- Does derailment–train indicate crash–train?
- Assessed by collecting statistics from a large corpus.
Statistical Discourse Features

Co-occurring predicate

Example

At least 10 people were killed . . . in the [crash]_{cand-pred} . . . Alvarez is accused of . . . causing the derailment of one [train]_{cand-arg} . . .

- Does derailment–train indicate crash–train?
- Assessed by collecting statistics from a large corpus.
Statistical Discourse Features

Co-occurring predicate

Example

At least 10 people were killed ... in the [crash]_{cand-pred} ... Alvarez is accused of ... causing the derailment of one [train]_{cand-arg} ...

- Does derailment–train indicate crash–train?
- Assessed by collecting statistics from a large corpus.

Similar idea: Co-occurring argument
### Statistical Discourse Features

**Co-occurring predicate**

**Example**

| At least 10 people were killed . . . in the [crash]_{cand-pred} . . . Alvarez is accused of . . . causing the derailment of one [train]_{cand-arg} . . . |

- Does derailment–train indicate crash–train?
- Assessed by collecting statistics from a large corpus.

**Similar idea: Co-occurring argument**

**Example**

| A senior official defended the [PATRIOT Act]_{cand-arg} . . . President Bush has urged Congress to [renew]_{cand-pred} the law . . . |
### Results

#### First experiment: accuracy of our method

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full algorithm</td>
<td>81.0</td>
</tr>
</tbody>
</table>

#### Ablation tests

- Major category (all true): 56.5
- Union of prior work: 78.0
- no statistical discourse: 79.9
- no local discourse: 79.3
- no local candidate properties: 79.2
- no predicate-argument relatedness: 79.7
Results

- First experiment: accuracy of our method
First experiment: accuracy of our method

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full algorithm</td>
<td><strong>81.0</strong></td>
</tr>
<tr>
<td>Ablation tests</td>
<td></td>
</tr>
<tr>
<td>Major category (all true)</td>
<td>56.5</td>
</tr>
<tr>
<td>Union of prior work</td>
<td>78.0</td>
</tr>
<tr>
<td>no statistical discourse</td>
<td>79.9</td>
</tr>
<tr>
<td>no local discourse</td>
<td>79.3</td>
</tr>
<tr>
<td>no local candidate properties</td>
<td>79.2</td>
</tr>
<tr>
<td>no predicate-argument relatedness</td>
<td>79.7</td>
</tr>
</tbody>
</table>
Results

- Second experiment: incorporate into an RTE system
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:

  1. Lexical coverage
  2. Predicate-argument relationships coverage
     - Explicit only
     - Explicit + implicit, taking gold-standard annotations
     - Explicit + implicit, taking only relationships recognized by the recognition algorithm.

Configuration

<table>
<thead>
<tr>
<th>F1 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit only</td>
</tr>
<tr>
<td>Gold-standard annotations</td>
</tr>
<tr>
<td>Algorithm recognition</td>
</tr>
</tbody>
</table>

Evaluate on RTE-6 dataset
Train & classify using F1 optimized logistic regression classifier of Jansche (2005).
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage

Evaluate on RTE-6 dataset
Train & classify using F1 optimized logistic regression classifier of Jansche (2005).
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage
     - Explicit only
  3. Explicit + implicit, taking gold-standard annotations
  4. Explicit + implicit, taking only relationships recognized by the recognition algorithm.

Configuration:
- F1 %
  - Explicit only: 44.4
  - Gold-standard annotations: 46.8
  - Algorithm recognition: 45.2

Evaluate on RTE-6 dataset
Train & classify using F1 optimized logistic regression classifier of Jansche (2005).
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage
     1. Explicit only
     2. Explicit + implicit, taking gold-standard annotations

Evaluate on RTE-6 dataset
Train & classify using F1 optimized logistic regression classifier of Jansche (2005).
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage
     1. Explicit only
     2. Explicit + implicit, taking gold-standard annotations
     3. Explicit + implicit, taking only relationships recognized by the recognition algorithm.
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage
     1. Explicit only
     2. Explicit + implicit, taking gold-standard annotations
     3. Explicit + implicit, taking only relationships recognized by the recognition algorithm.
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage
     1. Explicit only
     2. Explicit + implicit, taking gold-standard annotations
     3. Explicit + implicit, taking only relationships recognized by the recognition algorithm.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>F1 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit only</td>
<td>44.4</td>
</tr>
<tr>
<td>Gold-standard annotations</td>
<td>46.8</td>
</tr>
<tr>
<td>Algorithm recognition</td>
<td>45.2</td>
</tr>
</tbody>
</table>
Results

- Second experiment: incorporate into an RTE system
- Basic system with two features:
  1. Lexical coverage
  2. Predicate-argument relationships coverage
     1. Explicit only
     2. Explicit + implicit, taking gold-standard annotations
     3. Explicit + implicit, taking only relationships recognized by the recognition algorithm.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>F1 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit only</td>
<td>44.4</td>
</tr>
<tr>
<td>Gold-standard annotations</td>
<td>46.8</td>
</tr>
<tr>
<td>Algorithm recognition</td>
<td>45.2</td>
</tr>
</tbody>
</table>

- Evaluate on RTE-6 dataset
- Train & classify using F1 optimized logistic regression classifier of Jansche (2005).
Takeouts
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
New recognition task: Recognizing Implied Predicate-Argument Relationships

Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
  - Better coverage

- New and freely available dataset
  - Constructed semi-automatically.
  - Larger than all earlier work datasets.

- Recognition features
  - Promising results
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
  - Better coverage
- New and freely available dataset
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
  - Better coverage
- New and freely available dataset
  - Constructed semi-automatically.
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
  - Better coverage
- New and freely available dataset
  - Constructed semi-automatically.
  - Larger than all earlier work datasets.
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
  - Better coverage
- New and freely available dataset
  - Constructed semi-automatically.
  - Larger than all earlier work datasets.
- Recognition features
Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships
- Targets the needs of textual inference tasks, like RTE, QA, IE, better than earlier work annotation task
  - Easier: candidates are pre-detected
  - Better coverage
- New and freely available dataset
  - Constructed semi-automatically.
  - Larger than all earlier work datasets.
- Recognition features
- Promising results