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` $(\bullet_{\texttt{Firm}}, \bullet_{\texttt{Model}})$

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

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.

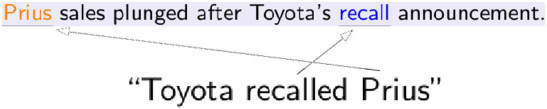
- Observation: in inference applications terms for the predicate and the argument are pre-detected.
  - RTE: match Hypothesis predicate and arguments in the Text.

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.

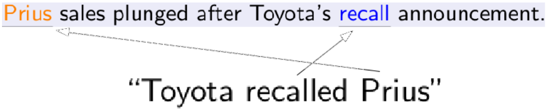


Annotation vs. Recognition

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.



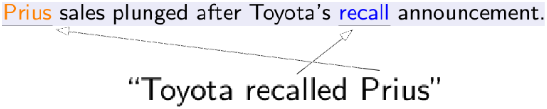
  - QA: match question's predicate and argument answer type.

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.



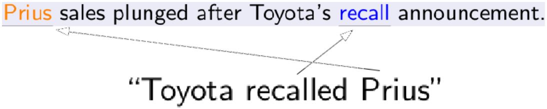
  - QA: match question's predicate and argument answer type.
  - IE: predicate detection, argument entity extractor.

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.



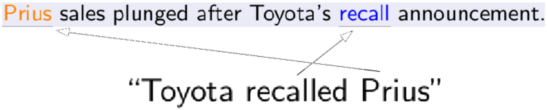
  - QA: match question's predicate and argument answer type.
  - IE: predicate detection, argument entity extractor.

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.



  - 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

Task Definition

Text

Prius sales plunged after Toyota's recall announcement.

Hypothesis

Toyota recalled Prius.

Textual Entailment

- **Input:** Text and Hypothesis

Task Definition

Text

Prius sales plunged after Toyota's recall announcement.

Hypothesis

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

Text

Prius sales plunged after Toyota's recall announcement.

Hypothesis

Toyota recalled Prius.

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

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

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

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)
- Freely available `www.cs.biu.ac.il/~nlp/resources/downloads/implied-relationships/`

- Random Forest learning algorithm
- 15 features

- Random Forest learning algorithm
- 15 features

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

Statistical Discourse Features

Co-occurring predicate

Statistical Discourse Features

Co-occurring predicate

Example

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

Statistical Discourse Features

Co-occurring predicate

Example

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

- Does derailment–train indicate crash–train?

Statistical Discourse Features

Co-occurring predicate

Example

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

- 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]$_{\text{cand-pred}}$ . . . Alvarez is accused of . . . causing the derailment of one [train]$_{\text{cand-arg}}$ . . .

- 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]$_{\text{cand-pred}}$ ... Alvarez is accused of ... causing the derailment of one [train]$_{\text{cand-arg}}$ ...

- 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]$_{\text{cand-pred}}$ ... Alvarez is accused of ... causing the derailment of one [train]$_{\text{cand-arg}}$ ...

- 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]$_{\text{cand-arg}}$ ... President Bush has urged Congress to [renew]$_{\text{cand-pred}}$ the law ...

Results

Results

- First experiment: accuracy of our method

Results

- First experiment: accuracy of our method

Configuration	Accuracy %
Full algorithm	**81.0**
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

- Second experiment: incorporate into an RTE system

Results

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

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

Results

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

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

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.

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

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.

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).

Takeouts

Takeouts

- New recognition task: Recognizing Implied Predicate-Argument Relationships

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

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

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