Textual inference: Methods, open source platform and applications

Ido Dagan
Bar-Ilan University, Israel

Bernardo Magnini
Foundation Bruno Kessler, Trento

Guenter Neumann
German Research Center for Artificial Intelligence, Saarbrucken

Sebastian Pado
University of Heidelberg

Excitement project
What is applied *textual inference*?

“Match” different text fragments where:

- One text has the **same meaning** as the other
- One text **implies the meaning** of the other

**pepper may trigger sneezing**

**pepper can cause sneezing**

**pepper may trigger sneezing**

**allergies can be produced by hot spices**
What is applied *textual inference*?

"Match" different text fragments where:

- One text has the **same meaning** as the other
- One text **implies the meaning** of the other

**paraphrasing**

**bi-directional entailment**

pepper may trigger sneezing

pepper can cause sneezing

(allergies can be produced by hot spices)

pepper may trigger sneezing

(directional) textual entailment
Example Applications

Question Answering

- Which foods are allergenic?
  - Allergies can be produced by hot spices
  - Pepper may trigger sneezing
  - Many people are allergic to peanuts

Information Extraction

- Extract pairs of foods and symptoms

Search

- Allergenic foods

Summarization

- Summarize documents about allergies
Novel Application: Text Exploration

- coffee in economy is awful
- coffee is awful
- they have horrible coffee
disgusting coffee is served
- sandwiches are too expensive
sandwiches are overpriced
you charge too much for sandwiches
- no vegetarian food
provide veggie meals
- no refreshments
- no enough food selection
expand meal options
- not happy with the catering
- food on train is too expensive
- food is bad
food quality is disappointing
- bad food in premier
- journey is too slow
- no clear information
- not happy with the service
- not happy with the staff
- staff is unfriendly
- not enough food selection
expand meal options
- food is bad
food quality is disappointing
- bad food in premier
- journey is too slow
- no clear information
- not happy with the service
- not happy with the staff
- staff is unfriendly
The EXCITEMENT Project

• Scientific goals
  • Advance textual entailment research
  • Provide a flexible open platform for textual inference (EOP)

• Industrial goals
  • Advance customer interaction analytics, via
    • textual inference technologies

EXCITEMENT: EXploring Customer Interactions via TExtual entailMENT
Outline

• Entailment recognition algorithm
  • Alignment based

• Entailment knowledge resources

• The EXCITEMENT Open Platform (EOP)

• Entailment graphs
Alignment-based Entailment Recognition
Alignment-based Entailment

- Various algorithms proposed to recognize textual entailment
- Recent work in EXCITEMENT: Alignment-based entailment
- Intuition: The more material in the hypothesis can be “explained” / ”covered” by the premise, the more likely entailment is

P: Peter was Susan’s husband
H: Peter was married to Susan

P: Peter did not know Susan
H: Peter was married to Susan
Alignment-based Entailment: The Algorithmic Level

• **Step 1**: Automatic linguistic analysis (Optional)
  - Normalize surface forms, detect structure

```
H: Peter was married to Susan
```

```
P: Peter was Susan’s husband
```

- Part-of-speech tagger
- Lemmatizer
- Parser
- ...
Alignment-based Entailment: The Algorithmic Level

• **Step 2**: Identify links between words or phrases across the two texts
  • What words/phrases of P can explain words/phrases of H?

Lexical and Paraphrase Resources

P: Peter was Susan’s husband

H: Peter was married to Susan
Lexical and Paraphrase Alignment Resources

• Broad-coverage knowledge needed to align words/phrases
  • Align identical words
  • Align lexically related **words**: use lexical resources (WordNet, distributional similarity)
  • Align equivalent/related **phrases**: use paraphrase resources

- Peter ➔ Peter
- dog ➔ mammal
  Paris ➔ France
- was ➔ used to
  husband ➔ married to
Alignment-based Entailment: The Algorithmic Level

• **Step 3**: Computation of features over alignment
  
  • Formulate features that capture typical properties of valid entailments

P: Peter was not married to Susan

H: Peter was married to Susan
Concrete features

• Current implementation uses just four simple features

• **Word coverage**: What % of hypothesis words is covered?
• **Content word coverage**: What % of content words (N,V, A) covered?
• **Verb coverage**: What % of verbs is covered?
  • Verbs express the relations
• **Proper Noun coverage**: What % of proper nouns is covered?
  • Proper nouns express participants, typically require explicit mentions

• More features under development
  • E.g compatibility of negations
Alignment-based Entailment: The Algorithmic Level

• **Step 3**: Computation of features over alignment

```
H: Peter was married to Susan
   NE  V  V  P  NE
P: Peter was Susan’s husband
   NE  V  NE  NN
```

- **Word Coverage**: 5/5 = 100%
- **Content Word Coverage**: 4/4 = 100%
- **Verb Coverage**: 1/1 = 100%
- **Proper Noun Coverage**: 2/2 = 100%
Alignment-based Entailment: The Algorithmic Level

• **Step 4:** Classification (logistic regression, with training examples)

  
  
  
  Classification Model

  - **Yes / No**

  Word Coverage: 4/5 = 100%
  Content Word Coverage: 4/4 = 100%
  Verb Coverage: 1/1 = 100%
  Proper Noun Coverage: 2/2 = 100%

  P: Peter was Susan’s husband

  H: Peter was married to Susan

  NE V NE NN

  NE V V P NE
Why Alignment-based Entailment Recognition?

• Efficient
• (Almost completely) language-agnostic
• Robust: Can deal with noisy input data
  • Shallow linguistic cues
• Adaptable to new domains
  • Encode domain knowledge as alignment resource
• Extensible
• State of the art useful accuracy

• Will be included in EOP release in December 2014
Extensibility

- Sentence Pair
- Aligned Sentence Pair
- Feature Vector
- Pluggable aligners (one or more): Aligner A, Aligner B
- Pluggable scorers (one or more): Scorer (feature extractor) A, Score function B
- Classifier
- Visualization
- ENTAILMENT
- DECISION
# Performance at state-of-the-art

**Dataset: RTE-3**

<table>
<thead>
<tr>
<th>Language</th>
<th>Best Alignment-based EDA settings</th>
<th>Best previous EOP result</th>
</tr>
</thead>
<tbody>
<tr>
<td>EN</td>
<td>67.0</td>
<td>66.8 (BIUTEE transformation)</td>
</tr>
<tr>
<td>IT</td>
<td>65.4</td>
<td>63.5% (EDITS transformation)</td>
</tr>
<tr>
<td>DE</td>
<td>63.9</td>
<td>63.5 (TIE matching features)</td>
</tr>
</tbody>
</table>

- Used for entailment graph construction on customer interactions data
- Results seem useful
Entailment Knowledge Resources
Various Resources Types

• Wordnet
  • *pepper* ➔ *spice*  *stock* ➔ *share*

• Derivational morphology
  • *allergenic* ➔ *allergy*  *acquire* ➔ *acquisition*

• Corpus-based distributional similarity
  • As seen in tutorial
  • Similar to word2vec type of output; limited correlation with entailment/equivalence
    • *Directional* similarity, usually somewhat better

• Wikipedia derived
  • *Madonna* ➔ *singer*

• Paraphrasing – bilingual based

Tools for constructing knowledge resources for domain corpora and languages
Extraction from Wikipedia

(Shnarch et al., 2009)

E.T. the Extra-Terrestrial

From Wikipedia, the free encyclopedia

(Redirected from E.T. (film))

*E.T. the Extra-Terrestrial* is a 1982 science fiction film co-produced and directed by Steven Spielberg, written by Melissa Mathison and starring Henry Thomas, Robert MacNaughton, Drew Barrymore, Dee Wallace, and Peter Coyote. It tells the story of Elliott (played by Thomas), a lonely boy who befriends a friendly alien, dubbed "E.T.," who is stranded on Earth. Elliott and his

- Be-complement
- Redirect
- Parenthesis
- Link
Bilingual-based Paraphrases

• Intuition: p and p’ are paraphrases if both translate into same phrase t (a “pivot”)

• Procedure:
  1. Word- and phrase-align parallel corpus (e.g. English-German)
  2. Extract bilingual translation table
  3. Hop from English to German and back to obtain paraphrase table (plus probability)
Excitement Open Platform
Excitement Open Platform (EOP)

- **Excitement Project**: develop generic entailment platform
  - Step 1: Decouple preprocessing and actual entailment computation
  - Step 2: Decompose inference into components

Excitement EU project: [http://www.excitement-project.eu](http://www.excitement-project.eu)
Magnini et al.: The Excitement Open Platform, ACL demo 2014
Pado et al.: Journal Natural Language Engineering, 2014
EXCITEMENT Platform for Textual Inference

Configurator

**ITALIAN**
- Tokenization, Lemma, POS, dependency parsing

**GERMAN**
- Token, POS, Lemma, dependency parsing

**ENGLISH**
- Token, Lemma, POS, dependency parsing

**Lexical component**
- Entailment rules

**Alignment Component**

**Algorithms**
- Distance-based (EDITS)
- Classification-based (TIE)
- Transformation-based (BIUTEE)
- Alignment-based (P1EDA)

**Scoring Component**
- Bag of Words similarity
- Edit Distance

**Distance Component**

**Distribution Similarity**
- English
- German
- Italian

**WordNet**
- Italian
- German
- English

**Wikipedia**
- Italian
- English

**Derivational Morphology**
- Italian
- English
- German

**Phrase Tables**
- Italian
- English
- German

Smoking causes lung cancer.

entails?

Smoke hurts your lungs.

Y/N
EOP Users

• **Textual Entailment Researchers**
  • Evaluate algorithms to find out their strengths and weaknesses
  • Implement algorithmic ideas
  • Remove influence of resources, preprocessing, ...
  • Extend existing system OR build new system from scratch

• **Textual Entailment End Users**
  • Compare various TE algorithms for applications
  • Does not want to touch code
  • Clear interface (package):
    • Flexible, usable & configurable system
  • Fast prototype to setup simple TE system (Bulgarian)
EOP Distribution
The EXCITEMENT Open Platform (EOP) is an open source software platform containing state-of-the-art algorithms for recognizing textual entailment relations.

Major changes of release 1.1.4 compared to the previous release 1.1.3:

New Features:

- MaltParser for Italian language

Bug Fixes:

- English MaltParser pipeline made wrong results due to POS tag mismatch.
- Italian TreeTagger missed "canonical" POS tag.
- Util submodule used a wrong version of LAP.
Open Source Distribution of EOP

- Quick Code Integration
  - Git, Github, Maven, Jenkins

- Quality Control
  - Code quality tools (e.g. check style, find bugs)

- Additional Highlights
  - Archive for Experiments
  - GitHub wiki pages (release-specific documentation)
  - Two Distributions: API and Command Line Interface

- License: General Public License (GPL) version 3
Overview – Release Management

• Keeping several code versions (master branch, releases)
• Automatic methods for
  • creating new releases and resource distributions
  • maintenance of release-specific documentation
  • Generating Web Page (EOP web site)
• Separate documentations for end users and developers
EOP in Numbers (08/09/2014)

• EOP GitHub repository:
  • 52 Members (people who forked the EOP Repository)

• Mailing lists:
  • developers: 21
  • users: 24 (12 external users)

• EOP v1.1.3
  • Downloads: 77
  • Experiments Archive: 13 experiments
    • 96 experiments in the current developers version EOP v1.1.5
  • Download + Installation: 10 min by a shell script
Learn More

• EXCITEMENT project web site: http://www.excitement-project.org


Building Entailment Graphs
**Customer Interactions Scenario**

**Int-448:**
Efficient service. Quick through security and check in. Staff could have been a bit more friendly though and leg room in standard class was quite poor.

**Int-202:**
Everything ran smoothly and well. Only complaint is lack of leg room with seating with tables. Very cramped when all seats are taken.

**Int-275:**
The leg room in economy class is not enough I was constantly being kicked by opposite passenger I travel by train lots and this compares badly to other trains

**Int-303:**
My only gripes, not enough leg room in standard and I think it would be chic to have refreshments served in carriages, either trolley or trays like in theatres.
EXCITEMENT application scenario

Requirements

• Need for customer interaction analytics
  • Compact representation (show just relevant information)
  • Informative representation: general categories (e.g. “food”, “internet”) are not enough
• Need to manage streams of data
• Multiple channels: e-mail, speech, social media
• Noisy data: automatic transcriptions, social media style, etc.
• Multiple languages
  • Excitement: English, Italian, German

Challenge

• Core technology: entailment graphs based on the EOP platform
• Current experiments based on the Alignement-based algorithm
TOPIC: Reasons for dissatisfaction in railway service

Int-448: Efficient service. Quick through security and check in. But leg room in standard class was quite poor.

Int-202: Everything ran smoothly and well. Only complaint is lack of leg room with seating with tables.

Int-275: Seating is very cramped – my journey has been very uncomfortable with the person next to me taking up most of the space we have.

Int-303: My only gripes r not enough leg room in standard and I think it would be chic to have refreshments served in carriages, either trolley or trays like in theatres.
Leg room in standard class was quite poor.
Leg room in standard class was quite poor.

Leg room in standard class was poor.

Leg room was quite poor.
Leg room in standard class was quite poor

Leg room was poor in standard class

Leg room was quite poor
Leg room in standard class was quite poor.

Leg room in standard class was poor.

Leg room was quite poor.

Leg room was poor.
Building Fragment Graphs

Result: a DAG
- rooted in Fragment
- Base predicate (fragment without all modifiers as only leaf)
Merging Graphs with the EOP

Int-303

Not enough leg room in standard

Not enough leg room

Int-448

Leg room in standard class was quite poor

F2

F2_S1

leg room in standard class was poor

F2_S2

leg room was quite poor

F2_S3

leg room was poor
Merging Graphs with the EOP

- Not enough leg room in standard
- Leg room in standard class was quite poor
- Leg room was quite poor
- Not enough leg room
- Merging Graphs with the EOP

Int-448

Entails ?

F2_S1

Entails ?

F2

F2_S2

F2_S3

Entails ?

Int-303

Entails ?

Entails ?

Entails ?
Merging Graphs with the EOP

Not enough leg room in standard

Not enough leg room

Leg room in standard class was quite poor

leg room in standard class was poor

leg room was quite poor

leg room was poor
Merging Graphs with the EOP

- **Int-202 F1**: lack of leg room with seating with tables
- **Int-275 F1_S1**: seating is cramped
- **Int-448 F2_S3**: leg room was poor
- **Int-303 F1_S1**: not enough leg room
- **Int-202 F1_S1**: lack of leg room
- **Int-275 F1_S1**: seating is cramped
- **Int-448 F2_S2**: leg room was quite poor
- **Int-448 F2_S1**: leg room in standard class was poor
- **Int-303 F1**: Not enough leg room in standard
Conclusion

- **Textual Entailment** provides a generic perspective for inference over textual expressions.
- **Textual inference technology** is still in early stages, with limited yet potentially useful performance.
- The **EXCITEMENT Open Platform** offers available technology for research.
- **Entailment Graphs** have a potential for text exploration applications.
- **Datasets and baseline results** for customer interactions are available for further research.