Considering Discourse References in Textual Entailment Annotation

L. Bentivogli¹, I. Dagan², H. T. Dang³, D. Giampiccolo⁴, M. Lo Leggio⁴, B. Magnini¹

¹FBK-irst, Trento, Italy
²Bar-Ilan University, Ramat Gan, Israel
³NIST, Gaithersburg, Maryland, USA
⁴CELCT, Trento, Italy
Outline

• Introduction
• The Entailment Search Task
• Knowledge required to perform entailment annotation
• The Data set
• Creation of hypotheses
• Entailment annotation and co-reference resolution: entities, time, space, events
• Conclusions and future work
Outline

• Introduction
  • The Entailment Search Task
  • Knowledge required to perform entailment annotation
• The Data set
• Creation of hypotheses
• Entailment annotation and co-reference resolution: entities, time, space, events
• Conclusions and future work
Textual Entailment

**Textual entailment:** directional relationship between pairs of text expressions, denoted by T - the entailing “Text”, and H - the entailed “Hypothesis”

T entails H if, typically, a human reading T would infer that H is most likely true

**T:** A shootout at the Guadalajara airport in May, 1993, killed Cardinal Juan Jesus Posadas Ocampo and six others.

**H:** Juan Jesus Posadas Ocampo died in 1993.
The RTE Campaign

Recognizing Textual Entailment (RTE)

• annual evaluation campaign started in 2005, now at 5\textsuperscript{th} edition

• the \textit{traditional task}: given a manually created H-T pair, recognize whether T entails H

• \textbf{RTE-5}: a new pilot task has been introduced to answer the need of moving towards more realistic scenarios
Outline

• Introduction

• The Entailment Search Task
  • Knowledge required to perform entailment annotation

• The Data set

• Creation of hypotheses

• Entailment annotation and co-reference resolution: entities, time, space, events

• Conclusions and future work
The RTE-5 Search Pilot Task

Consists in finding all the sentences that entail a given H in a given set of documents

Hypotheses Set

H1: The AS-28 mini-submarine was trapped underwater
H2: Seven submariners were onboard the AS-28
H3: The AS-28 accident happened in eastern Russia
H4: Russia requested international help to rescue the AS-28
H5: The AS-28 crew was rescued in satisfactory conditions.

Sentences Set

S1: Effort seen to raise stricken Russian sub
S2: The effort to attach the cables marked the start of an operation to raise the vessel trapped below the surface with seven crewmen on board.
S3: The submarine, used in rescue, research and intelligence-gathering missions, became stuck at the bottom of the Bay of Berezovaya off Russia's Far East.
S4: He was subsequently quoted by Interfax news agency as saying that attachment of the cables to the mini-submarine was not certain and that a video camera was being lowered to find out what the cables were attached to.
S5: "Then we can tell 100 percent that it's the submarine. S6: And only then can we say that we carried out our task," Fyodorov said.
S7: Fyodorov said earlier that the seven crewmen were in satisfactory condition...

S0: Japanese coast help to trapped Russian sub
S1: Japan on Friday dispatched four military ships to help Russia rescue seven crew members aboard a small submarine trapped on the seabed in the Far East.
S2: The assistance comes despite rocky relations between Japan and Russia, which have yet to formally end World War II amid Japan's claims to four islands off Hokkaido that Soviet troops seized in August 1945.
S3: The Russian Pacific Fleet said the crew had 120 hours of oxygen reserves on board when the submarine submerged at midday Thursday 300 GMT Wednesday.
S4: Japan dispatched a submarine-rescue vessel, and a depot ship from Yokosuka port and two minesweepers from Hokkaido, the defense spokesman said.
S5: "We will do our utmost efforts to rescue them. S6: We are hopeful," he said.
S7: There were widely varying accounts of how much oxygen was remaining for the crew.

S6: There are seven crew members aboard the vessel, stranded on the ocean floor in a bay off the coast of the Kamchatka peninsula in Russia's Far East region.
## Traditional vs Pilot Task

<table>
<thead>
<tr>
<th>Aim</th>
<th>Traditional Task</th>
<th>Search Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judging whether T entails H, for each manually created T-H pair</td>
<td>Retrieving all the sentences that entail a given H in a set of documents about a topic</td>
<td></td>
</tr>
<tr>
<td>The distribution of entailing T’s is determined a priori</td>
<td>Reflects the natural distribution of entailment in a corpus</td>
<td></td>
</tr>
<tr>
<td>Both T and H do not contain any textual reference outside the pair itself</td>
<td>Both T and H are to be interpreted within the context of the topic</td>
<td></td>
</tr>
<tr>
<td>Verb tenses are ignored</td>
<td>Verb tenses must be interpreted in the context</td>
<td></td>
</tr>
</tbody>
</table>

### Data Set

- Bentivogli et al.

### References

- GL09 - Semantic annotation WS

### Pisa, September 17 2009
Outline

• Introduction
• The Entailment Search Task
• Knowledge required to perform entailment annotation
• The Data set
• Creation of hypotheses
• Entailment annotation and co-reference resolution: entities, time, space, events
• Conclusions and future work
Knowledge Required for the Entailment Judgment

The standard definition of entailment assumes prior knowledge:

Knowledge + Text $\rightarrow$ Hypothesis

Three kinds of knowledge:

- **Linguistic Knowledge**: to interpret H and T
- **Common background World Knowledge**: to make additional inferences
- **Text and Corpus Knowledge**: to resolve local and cross-document co-references
S0: US scientists announce discovery of possible ‘10th planet’
S1: A US astronomer announced Friday he had discovered what could be a 10th planet in the outer reaches of the solar system nearly 15 billion kilometers (nine billion miles) from earth.
S2: If confirmed, the discovery by Professor Mike Brown of the respected California Institute of Technology would be the first of a planet since Pluto was identified in 1930.
S3: The body 'the 10th planet of the solar system,' said Brown, a professor of planetary astronomy in a phone conference from CalTech in the town of Pasadena.
S4: It is the farthest object ever discovered to
S5: "It's definitely bigger than Pluto".

H: 2003 UB313 is bigger than Pluto.

S0: Scientists discover moon orbiting so-called 10th planet
S1: The astronomers who claim to have discovered the 10th planet in the solar system have made another intriguing announcement: it has a moon.
S2: While observing the new, so-called planet from Hawaii last month, a team of astronomers led by Michael Brown of the California Institute of Technology spotted a faint object trailing next to it.
S4: The discovery may also help scientists determine the new planet's mass.
S5: In July, Brown announced the discovery of an icy, rocky object larger than Pluto in the Kuiper Belt, a disc of icy bodies beyond Neptune.
S6: Brown labeled the object a planet and nicknamed it Xena after the lead character in the former TV series."Xena: Warrior Princess."
S7: By determining the moon's distance and orbit around Xena, scientists can calculate how heavy it is.
S8: "It is guaranteed bigger than Pluto," said Michael E. Brown, a professor of planetary astronomy at Caltech.
Outline

- Introduction
- The Entailment Search Task
- Knowledge required to perform entailment annotation

- The Data set
- Creation of hypotheses
- Entailment annotation and co-reference resolution: entities, time, space, events
- Conclusions and future work
The Pilot Search Task of the 5th RTE campaign is situated in the Summarization application setting.

The Development Set has been created from the TAC 2008 Update Summarization Task Data Set:

- **Documents:** 10 topic newswire collections randomly selected from the 48 TAC topics
- **Hypotheses:** based on Summary Content Units (SCUs) that have been created from human-authored summaries of the topic collections
The Data Set: Composition

• Description:
  – 10 Topics: 10 newspaper articles manually split into numbered sentences
  – Between 6 and 10 Hypotheses
  – Three annotators

• Numbers:
  – 80 H’s
  – 2,538 sentences
  – 20,104 sentence annotations (annotation of each sentence against each of the related hypotheses)
  – 810 entailing sentences
Outline

- Introduction
- The Entailment Search Task
- Knowledge required to perform entailment annotation
- The Data set
- Creation of hypotheses
  - Entailment annotation and co-reference resolution: entities, time, space, events
- Conclusions and future work
Creation of Hypotheses

Criteria:

• **Explicitness:**
  to reduce ambiguities and facilitate the correct interpretation of H’s

• **Conciseness:**
  to maintain linguistic fluency

• **Time anchoring:**
  if not otherwise specified, the content of H is conventionally fixed at the day after the publication of the last document in the topic collection
Examples of Hypotheses

• **Dennis Rader** was arrested on **February 25, 2005**
  
  ENTITY (person): complete proper name  
  TIME: explicit absolute date

• **The ice is melting** in the Arctic
  
  TIME: conventional anchoring  
  → the day after the last document: 2005/08/05  
  SPACE: explicit mention of location

• **The Kansas Bureau of Investigation** collected hundreds of DNA swabs related to the **BTK case**
  
  EVENT: use of a short phrase
Outline

• Introduction
• The Entailment Search Task
• Knowledge required to perform entailment annotation
• The Data set
• Creation of hypotheses
• Entailment annotation and co-reference resolution: entities, time, space, events
• Conclusions and future work
H: Michael Brown discovered 2003 UB313

Michael Brown mentions:

“Michael Brown”
“Michael A. Brown”
“Michael E. Brown”
“Mike Brown”
“Brown”
“he”
“the astronomer”
“the Caltech professor”
...

2003 UB313 mentions:

“2003 UB313”
“UB313”
“Xena” (nick-name)
“it”
“the planet”
“the new planet”
“the possible new planet”
“the object”
“a lump of rock and ice”
“the 10th planet of the solar system”
...
H: **2003 UB313** has a moon.

---

... 2003 UB313...

✅ **T1:** The astronomers who claim to have discovered the 10th planet in the solar system have made another intriguing announcement: **it has a moon.**

---

... EL61...

❌ **T2:** It is different from the previously discovered Kuiper bodies in that **it has a moon** that circles it every 49 days in a highly elliptical orbit.
**Time References/1**

- Conventional anchorage in H / Deictic expression in T

**H: The ice is melting in the Antarctic**

[Time anchor: 2005/08/15]

**T1:** Last month, scientists sounded an alarm bell on the effect of global warming on Antarctica, saying that more than 200 coastal glaciers are in retreat.

**T2:** Of the 244 marine glaciers that drain inland ice on the Antarctic Peninsula, 87 percent have fallen back over the last half century. (date of the article: 2005/06/06)

**T3:** Since the end of the last Ice Age, some 11,000 years ago, the iceshelf had slowly thinned, by several dozen metres. (date of the article: 2005/08/03)

*May 2005* (date of the article: 2005/06/06)
• Past tense in H / Future tense in T: modality

H: Orhan Pamuk went to court on 16 December 2005

T1: "It is not Orhan Pamuk who will stand trial tomorrow but Turkey."

T2: Pamuk, 53, Turkey's best-known novelist, is expected to go on trial Friday [...].

16 December 2005
(date of the article: 2005/12/15)
Ha: The ice is melting in the Arctic.
Hb: The ice is melting in the Antarctic.

... the Exit Glacier in the Arctic ...

• **T1:** [...] people moved ever closer to the rope line near the glacier as it shied away, practically groaning and melting before their eyes.

✓ Entailment for Ha: YES
✗ Entailment for Hb: NO

... in the Antarctic Peninsula...

• **T2:** Of the 244 glaciers that drain inland ice and feed these shelves, 87 percent have fallen back since the mid-1950s, according to a British study published in April.

✗ Entailment for Hb: NO
✓ Entailment for Hb: YES
H: The AS-28 was at a depth of about 600 feet when it was trapped.

...the AS-28 mini-submarine was freed on Sunday from its undersea entanglement...

T: The vessel rose to the surface at 4:26 p.m. local time, ending the crew's ordeal in the cold and darkness more than 600 feet below the surface off the Kamchatka Peninsula.
Outline

- Introduction
- The Entailment Search Task
- Knowledge required to perform entailment annotation
- The Data set
- Creation of hypotheses
- Entailment annotation and co-reference resolution: entities, time, space, events
- Conclusions and future work
Conclusions

- Preliminary qualitative evaluation of the impact of corpus knowledge and reference resolution on the textual entailment annotation in a natural occurring corpus
- Task: innovative way of performing textual entailment recognition as a search task against a corpus
- Data set: useful resource that reflects the “natural” distribution of entailment
Future work

• Realization of an enriched data-set containing all the entailing sentences annotated with:
  – Type(s) of co-reference involved in the entailment
  – Resolution of implicit/deictic/phoric references, relevant for the entailment

• Contradiction detection in the future RTE campaigns
Thank you for your attention
The Traditional Task

Given a couple H-T, the task consists in deciding if
- T entails H
- T contradicts H
- the veracity of H can’t be inferred from T

T: Internet media company Yahoo Inc. announced Monday it is buying Overture Services Inc. in a $1.63-billion (U.S.) cash-and-stock deal that will bolster its on-line search capabilities.

H1: Yahoo bought Overture  \rightarrow  Entailment
H2: Overture bought Yahoo  \rightarrow  Contradiction
H3: Yahoo manufactures cars  \rightarrow  Unknown
• Linguistic Knowledge

\textit{H: Mine accidents cause deaths in China.}

\textit{T: Four mine disasters in China have claimed the lives of at least 60 workers.}

\textit{to claim the lives = to cause the death of some people}

• World Knowledge

\textit{H: The ice is melting in the Arctic.}

\textit{T: The scene at the receding edge of the Exit Glacier in Alaska was part festive gathering, part nature tour with an apocalyptic edge.}

\textit{Alaska is in the Arctic}

\textit{The edge of the glacier is receding because the ice is melting}
Creation of Hypotheses/2

• **Entities:**  
  *Michael Brown discovered 2003 UB313*

• **Time:**  
  - explicit date:  
    *Dennis Rader was arrested on February 25, 2005*
  - conventional anchorage:  
    *The ice is melting in the Arctic*  
    (anchored on the day after the last document, 2005/08/15)

• **Locations:**  
  *Mine accidents cause deaths in China*

• **Events:**  
  *The Kansas Bureau of Investigation collected hundreds of DNA swabs related to the BTK case*  
  *About 50 people were killed in the attack*
Time References

- Absolute expression in H / Deictic expression in T

H: BTK resurfaced in 2004 after a period of silence.

T: The killer has sent several letters to police since resurfacing last March after years of silence.

**Time Reference:** publication date of the article: 2005/01/06

→ last March = March 2004

**Entailment:** YES
H: The conditions in subway tunnels after the attack hampered rescue operations
[Time anchor: 2005/07/12]

T: "It is extremely hot and very dusty and it is a great challenge for them to continue their work to recover the remaining bodies from the train underground," British Transport Police Chief told.

**Time Reference**: publication date of the article: 2005/07/09
H: Each bomb used in the attack contained less than 10 pounds of explosives.

✓ T: Forensic evidence indicates that the bombs each contained less than 10 pounds of high explosives – the Madrid bombs weighed 17 to 22 pounds - which could have kept them small enough to fit in rucksacks, Commissioner Blair said.

Event co-reference (implicit): the first mention of bombs refers to the (London) attack.

Entailment: YES
H: Sea levels are rising.

• T: But what has happened since did.

**Discourse analysis:** previous sentence:

“Because an ice "shelf" already floats on the sea, displacing its weight in water, Larsen B's disintegration - and that of the smaller, nearby Larsen A in 1995 - didn't raise ocean levels.”

→ did = raise ocean levels.

**Entailment:** ann1: NO, ann2: NO, ann3: YES