NLP
Text Similarity

Entailment and Paraphrasing
<table>
<thead>
<tr>
<th>1</th>
<th>Amrozi accused his brother, whom he called &quot;the witness&quot;, of deliberately distorting his evidence. Referring to him as only &quot;the witness&quot;, Amrozi accused his brother of deliberately distorting his evidence.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Yucaipa owned Dominick's before selling the chain to Safeway in 1998 for $2.5 billion. Yucaipa bought Dominick's in 1995 for $693 million and sold it to Safeway for $1.8 billion in 1998.</td>
</tr>
<tr>
<td>1</td>
<td>They had published an advertisement on the Internet on June 10, offering the cargo for sale, he added. On June 10, the ship's owners had published an advertisement on the Internet, offering the explosives for sale.</td>
</tr>
<tr>
<td>0</td>
<td>Around 0335 GMT, Tab shares were up 19 cents, or 4.4%, at A$4.56, having earlier set a record high of A$4.57. Tab shares jumped 20 cents, or 4.6%, to set a record closing high at A$4.57.</td>
</tr>
<tr>
<td>1</td>
<td>Revenue in the first quarter of the year dropped 15 percent from the same period a year earlier. With the scandal hanging over Stewart's company, revenue the first quarter of the year dropped 15 percent from the same period a year earlier.</td>
</tr>
<tr>
<td>1</td>
<td>The DVD-CCA then appealed to the state Supreme Court. The DVD CCA appealed that decision to the U.S. Supreme Court.</td>
</tr>
<tr>
<td>0</td>
<td>That compared with $35.18 million, or 24 cents per share, in the year-ago period. Earnings were affected by a non-recurring $8 million tax benefit in the year-ago period.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Entailment</strong></td>
<td></td>
</tr>
</tbody>
</table>
| **1** | YES | The sale was made to pay Yukos' US $27.5 billion tax bill, Yuganskneftegaz was originally sold for US $9.4 billion to a little known company Baikalfinansgroup which was later bought by the Russian state-owned oil company Rosneft.  

Baikalfinansgroup was sold to Rosneft. |
| **2** | NO | The sale was made to pay Yukos' US $27.5 billion tax bill, Yuganskneftegaz was originally sold for US $9.4 billion to a little known company Baikalfinansgroup which was later bought by the Russian state-owned oil company Rosneft.  

Yuganskneftegaz cost US $27.5 billion. |
| **3** | NO | Loraine besides participating in Broadway's Dreamgirls, also participated in the Off-Broadway production of "Does A Tiger Have A Necktie". In 1999, Loraine went to London, United Kingdom. There she participated in the production of "RENT" where she was cast as "Mimi" the understudy.  

"Does A Tiger Have A Necktie" was produced in London. |
| **4** | YES | "The Extra Girl" (1923) is a story of a small-town girl, Sue Graham (played by Mabel Normand) who comes to Hollywood to be in the pictures. This Mabel Normand vehicle, produced by Mack Sennett, followed earlier films about the film industry and also paved the way for later films about Hollywood, such as King Vidor's "Show People" (1928).  

"The Extra Girl" was produced by Sennett. |
| **5** | YES | A bus collision with a truck in Uganda has resulted in at least 30 fatalities and has left a further 21 injured.  

30 die in a bus collision in Uganda. |
| **6** | NO | Take consumer products giant Procter and Gamble. Even with a $1.8 billion Research and Development budget, it still manages 500 active partnerships each year, many of them with small companies.  

500 small companies are partners of Procter and Gamble. |
| **7** | YES | After his release, the clean-shaven Magdy el-Nashar told reporters outside his home that he had nothing to do with the July 7 transit attacks, which killed 52 people and the four bombers.  

52 people and four bombers were killed on July 7 |
| **8** | NO | Mrs. Bush's approval ratings have remained very high, above 80%, even as her husband's have recently dropped below 50%.  

80% approve of Mr. Bush. |
| **9** | NO | Recent Dakosaurus research comes from a complete skull found in Argentina in 1996, studied by Diego Pol of Ohio State University, Zulma Gasparini of Argentina's National University of La Plata, and their colleagues.  

A complete Dakosaurus was discovered by Diego Pol. |
Entailment and Presupposition

• Entailment
  – One fact follows from another
  – “All cats have whiskers” and “Martin is a cat” entail the statement “Martin has whiskers”
  – “Martin has whiskers and a tail” entails “Martin has whiskers”

• Presupposition
  – “The Queen of Utopia is dead” presupposes that Utopia has a queen
Entailment and Presupposition

• NACLO problem from 2010
• Author: Aleka Blackwell
  • http://www.nacloweb.org/resources/problems/2010/M.pdf
  • http://www.nacloweb.org/resources/problems/2010/MS.pdf
Think about the meaning of the following sentence:

(1) The 2010 Winter Olympics were in Canada.

Assuming that we only know sentence 1 to be true, is sentence 2 necessarily true?

(2) The 2010 Winter Olympics were in Vancouver.

The answer is no. Assuming we only know sentence 1 to be true, the 2010 Winter Olympics could have taken place in any Canadian city, but not necessarily in Vancouver.

Now examine the relationship between sentences 3 and 4. Assuming sentence 3 is true, is sentence 4 now necessarily true?

(3) The 2010 Winter Olympics were in Vancouver.
(4) The 2010 Winter Olympics were in Canada.

Now the answer is yes. Since Vancouver is a Canadian city, any event which occurs in Vancouver necessarily occurs in Canada.

The logical relationship which holds between sentences 3 and 4 is called an entailment. In formal terms, sentence A entails sentence B if whenever A is true, B is necessarily true. The entailment relationship is typically represented graphically this way: $A \models B$.

Here are some more examples of the entailment relationship between sentences:

(5) Shaun White is a Winter Olympian $\models$ Shaun White is an Olympian
(6) Shaun White is an Olympian $\models$ Shaun White is an athlete
(7) Shaun White won a gold medal $\models$ Someone won a gold medal

Notice that the entailment relationship must hold in the specified direction but will not necessarily hold in both directions. So, sentence 3 entails sentence 4 even though sentence 4 does not entail sentence 3.
Entailment and Presupposition

Now examine the relationship between sentences 8 and 9.

(8) I did not see Shaun White win the gold medal in the 2010 Winter Olympics.
(9) Shaun White won the gold medal in the 2010 Winter Olympics.

Sentences 8 and 9 illustrate a relationship called presupposition. In this pair of sentences, the information presented in sentence 9 is what the speaker assumes (or presupposes) to be the case when uttering sentence 8. That is, to say “I did not see Shaun White win the gold medal” assumes the belief that Shaun White won a gold medal. In formal terms, sentence A presupposes sentence B if A not only implies B but also implies that the truth of B is somehow taken for granted. A presupposition of a sentence is thus part of the background against which its truth or falsity is judged. The presupposition relationship is typically represented graphically this way: A >> B

Here are some more examples of presuppositions (where the first sentence in each pair presupposes the second):

(10) I regret not seeing Shaun White’s gold medal run >> Shaun White had a gold medal run
(11) Shaun White continues to rule the halfpipe >> Shaun White had been ruling the halfpipe
(12) Snowboarding is now an Olympic sport >> Snowboarding was once not an Olympic sport
Entailment and Presupposition

For any given pair of sentences, the entailment and presupposition relationships may or may not hold, together or separately.

For each of the following possible combinations, your task is to provide one example of a pair of sentences with an explanation of your reasoning for proposing your pair of sentences as a valid and convincing example in each case.

a. A pair of sentences in which sentence A neither entails nor presupposes sentence B.

b. A pair of sentences in which sentence A entails and presupposes sentence B.

c. A pair of sentences in which sentence A presupposes but does not entail sentence B.

d. A pair of sentences in which sentence A entails but does not presuppose sentence B.
Answers

For any given pair of sentences, the entailment and presupposition relationships may or may not hold, together or separately.

a. A pair of sentences in which sentence A neither entails nor presupposes sentence B.
   A. Shaun White is a Winter Olympian.
   B. The 2010 Winter Olympics were in Vancouver.
   Explanation: Sentences A and B are unrelated.

   Entailment: Given that sentence A is true, there is no way to know whether sentence B is true or false. If Shaun White is a Winter Olympian, the 2010 Winter Olympics may or may not have taken place in Vancouver. Thus, there is no entailment relationship between these two sentences.

   Presupposition: When uttering sentence A, a speaker would not take sentence B for granted (or assume that sentence B is background information against which the truth or falsity of sentence A would be judged). A speaker would not utter “Shaun White is a Winter Olympian” and assume the belief/take for granted that the 2010 Winter Olympics were in Vancouver.

b. A pair of sentences in which sentence A entails and presupposes sentence B.
   A. Shaun White continues to rule the halfpipe
   B. Shaun White had been ruling the halfpipe.

   Entailment: If sentence A is true, sentence B is necessarily true. The entailment relationship between these sentences relies on the meaning of the verb continue — to continue to rule the halfpipe, Shaun White had to be ruling the halfpipe already. Thus, sentence A entails sentence B.

   Presupposition: When uttering sentence A, a speaker would take sentence B for granted (or assume that sentence B is background information against which the truth or falsity of sentence A would be judged). A speaker who utters “Shaun White continues to rule the halfpipe” assumes the belief/takes for granted that Shaun White had been ruling the halfpipe. Thus, sentence A presupposes sentence B.
Entailment and Presupposition

c. A pair of sentences in which sentence A presupposes but does not entail sentence B.
   A. I did not see Shaun White win the gold medal in the 2010 Winter Olympics.
   B. Shaun White won the gold medal in the 2010 Winter Olympics.

Entailment: If sentence A is true, sentence B may or may not be true. The absence of an entailment relationship between these sentences relies on the words “did not see” – if it is true that I did not see Shaun White win the gold medal, then Shaun White may or may not have won the gold medal. Thus, sentence A does not entail sentence B.

Presupposition: When uttering sentence A, a speaker would take sentence B for granted (or assume that sentence B is background information against which the truth or falsity of sentence A would be judged). Specifically, a speaker who utters “I did not see Shaun White win the gold medal in the 2010 Winter Olympics” assumes the belief that Shaun White did actually win the gold medal in the 2010 Winter Olympics. Thus, sentence A presupposes sentence B.
Entailment and Presupposition

d. A pair of sentences in which sentence A **entails but does not presuppose** sentence B.

A. Shaun White did not win the gold medal in the 2010 Winter Olympics.

B. Shaun White did not both win the gold medal in the 2010 Winter Olympics and injure his ankle.

Entailment: If Shaun White did not win the gold medal in the 2010 Winter Olympics, then he necessarily did not *both* win that gold medal *and* injure his ankle, since he definitely did not win the gold medal. If one fact is not the case (the fact presented in sentence A), then both facts cannot be the case, either (the fact presented in sentence A + the new fact added to it in sentence B). Thus if sentence A is true, sentence B is necessarily true. Thus, sentence A entails sentence B.

Presupposition: When uttering sentence A, a speaker would not take sentence B for granted (or assume that sentence B is a background against which the truth or falsity of sentence A would be judged). Specifically, by uttering “Shaun White did not win the gold medal in the 2010 Winter Olympics” a speaker could not assume the belief that Shaun White did not both win the gold and injure his ankle, or that Shaun White either won a gold medal or injured his ankle. Whether Shaun White injured his ankle would not be information taken for granted when uttering “Shaun White did not win the gold medal in the 2010 Winter Olympics.” Thus, sentence A does not presuppose sentence B.
Using Multiple Translations

- The opening sentence of Madame Bovary by Gustave Flaubert (1857)
- Original
  - Nous étions à l’Étude, quand le Proviseur entra, suivi d’un nouveau habillé en bourgeois et d’un garçon de classe qui portait un grand pupitre. Ceux qui dormaient se réveillèrent, et chacun s’éléva comme surpris dans son travail.
- Translation 1 (Eleanor Marx-Aveling)
  - We were in class when the head-master came in, followed by a "new fellow," not wearing the school uniform, and a school servant carrying a large desk. Those who had been asleep woke up, and every one rose as if just surprised at his work.
- Translation 2
  - We were in the study-hall when the headmaster entered, followed by a new boy not yet in school uniform and by the handy man carrying a large desk. Their arrival disturbed the slumbers of some of us, but we all stood up in our places as though rising from our work.
- Translation 3
  - We were in the prep.-room when the Head came in, followed by a new boy in 'mufti' and a beadle carrying a big desk. The sleepers aroused themselves, and we all stood up, putting on a startled look, as if we had been buried in our work. [Barzilay]
A mathematician solved the problem

The problem was solved by a young mathematician

[Androutsopoulos and Malakasiotis 2010]
Use in information extraction

\[ S_1 = \text{Protesters seized several pumping stations, holding 127 Shell workers hostage.} \]

\[ S_2 = \text{Troops recently have raided churches, warning ministers to stop preaching.} \]

Figure 1: Sentences as dependency graphs.

<table>
<thead>
<tr>
<th>Relation Instance</th>
<th>Shortest Path in Undirected Dependency Graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$: protesters AT stations</td>
<td>protesters $\rightarrow$ seized $\leftarrow$ stations</td>
</tr>
<tr>
<td>$S_1$: workers AT stations</td>
<td>workers $\rightarrow$ holding $\leftarrow$ protesters $\rightarrow$ seized $\leftarrow$ stations</td>
</tr>
<tr>
<td>$S_2$: troops AT churches</td>
<td>troops $\rightarrow$ raided $\leftarrow$ churches</td>
</tr>
<tr>
<td>$S_2$: ministers AT churches</td>
<td>ministers $\rightarrow$ warning $\leftarrow$ troops $\rightarrow$ raided $\leftarrow$ churches</td>
</tr>
</tbody>
</table>

Table 1: Shortest Path representation of relations.

[Bunescu and Mooney 2005]
Use in information extraction

Figure 1: The dependency tree of the sentence “The results demonstrated that KaiC interacts rhythmically with KaiA, KaiB, and SasA.”

1. PROTJI - nsubj - interacts - prep_with - PROTJ2
2. PROTJI - nsubj - interacts - prep_with - PROTJ0 - conj_and - PROTJ2
3. PROTJI - nsubj - interacts - prep_with - PROTJ0 - conj_and - PROTJ2
4. PROTJI - conj_and - PROTJ2
5. PROTJI - conj_and - PROTJ2
6. PROTJI - conj_and - PROTJ0 - conj_and - PROTJ2

[Erkan et al. 2007]
Using lattices for paraphrasing

Figure 6: Word lattices obtained from sentence clusters in Barzilay and Lee’s method.

[figure from Androutsopoulos and Malakasiotis 2010, based on Barzilay and Lee 2003]
NLP