Implicit and Explicit Coherence Relations

Maite Taboada

Department of Linguistics
Simon Fraser University
Vancouver, Canada
mtaboada@sfu.ca

7e Giornate svizzere della Linguistica
Lugano, Switzerland
September 14, 2012
Acknowledgements

• Joint work with Debopam Das (Simon Fraser University)

• Funding
  ▪ Natural Sciences and Engineering Research Council of Canada
  ▪ Ministry of Science and Innovation of Spain
Outline

- Coherence/rhetorical/discourse relations
- Rhetorical Structure Theory
- Signals for coherence relations
  - The Cooperative Principle
  - Psycholinguistic and cognitive evidence for coherence relations
  - Are there true ‘implicit’ relations?
- Corpus study
  - Taxonomy of cues
  - Annotation
  - Results
- Other corpus-based work
- An application
  - Discourse parsing for sentiment detection
- Conclusions and further work
Coherence relations

- Coherence/discourse/rhetorical relations
- Relations among propositions in discourse
- Building blocks of discourse
- Help explain coherence
- Most of what I will say applies across different theories of coherence relations
  - Rhetorical Structure Theory (Mann and Thompson 1988)
  - Rhetorical relations in SDRT (Asher and Lascarides 2003)
  - Linguistic Discourse Model (Polanyi 1988)
  - ...
Rhetorical Structure Theory

• Explains text coherence through the relations holding among parts of a text

• Components
  ▪ Text spans
  ▪ Relation schemas

• Hierarchical: spans and schemas might become part of other spans and enter in new relations.

• Based on hypotactic and paratactic relations in language
Text spans

• Minimal units
  ▪ Clauses
  ▪ Mostly adjunct clauses
  ▪ Subject and object clauses (i.e., complement clauses) are part of their matrix
  ▪ Treatment of relative clauses (adjuncts) varies
  ▪ In general, granularity can be adjusted depending on requirements
Relation schemas

- circumstance
- contrast
- joint
- motivation
- enablement
- sequence
- sequence
Two basic types of relations

- Paratactic
  
  I’ll go to the grocery store and I’ll also go to the bank

- Hypotactic
  
  I’ll go to the grocery store and I’ll also go to the bank while I’m there
Relations

- Relations hold between two non-overlapping text spans
- Text spans are called *nuclei* and *satellite*
- Most of the relations described hold between a nucleus and a satellite, although there are also multi-nuclear relations
- A relation consists of:
  1. Constraints on the Nucleus,
  2. Constraints on the Satellite,
  3. Constraints on the combination of Nucleus and Satellite,
  4. The Effect.
### One possible taxonomy of relations

<table>
<thead>
<tr>
<th>Circumstance</th>
<th>Antithesis and Concession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solutionhood</td>
<td>Antithesis</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Concession</td>
</tr>
<tr>
<td>Background</td>
<td></td>
</tr>
<tr>
<td>Enablement and Motivation</td>
<td>Condition and Otherwise</td>
</tr>
<tr>
<td>Enablement</td>
<td>Condition</td>
</tr>
<tr>
<td>Motivation</td>
<td>Otherwise</td>
</tr>
<tr>
<td>Evidence and Justify</td>
<td>Interpretation and Evaluation</td>
</tr>
<tr>
<td>Evidence</td>
<td>Interpretation</td>
</tr>
<tr>
<td>Justify</td>
<td>Evaluation</td>
</tr>
<tr>
<td>Relations of Cause</td>
<td>Restatement and Summary</td>
</tr>
<tr>
<td>Volitional Cause</td>
<td>Restatement</td>
</tr>
<tr>
<td>Non-Volitional Cause</td>
<td>Summary</td>
</tr>
<tr>
<td>Volitional Result</td>
<td>Other Relations</td>
</tr>
<tr>
<td>Non-Volitional Result</td>
<td>Sequence</td>
</tr>
<tr>
<td>Purpose</td>
<td>Contrast</td>
</tr>
</tbody>
</table>
How to do an RST analysis

1. Divide the text into units
2. Look at each unit, and its neighbours. Is there a clear relation holding between them?
3. If yes, then mark that relation (e.g., Condition)
4. If not, the unit might be at the boundary of a higher-level relation. Look at relations holding between larger units (spans)
5. Continue until all the units in the text are accounted for
6. Marking a relation involves satisfying all 4 fields (especially the Effect). The Effect is the plausible intention that the text creator had
This is the sixth book in the Stephanie Plum series. If you have never had the fun of reading a book in this series do not start with this one. Go to the library and start with One For The Money and work your way up to Hot Six.
This is the sixth book in the Stephanie Plum series. If you have never had the fun of reading a book in this series do not start with this one. Go to the library and start with One For The Money and work your way up to Hot Six.
Outline

• Coherence/rhetorical/discourse relations
• Rhetorical Structure Theory
• Signals for coherence relations
  ▪ The Cooperative Principle
  ▪ Psycholinguistic and cognitive evidence for coherence relations
  ▪ Are there true ‘implicit’ relations?
• Corpus study
  ▪ Taxonomy of cues
  ▪ Annotation
  ▪ Results
• Other corpus-based work
• An application
  ▪ Discourse parsing for sentiment detection
• Conclusions and further work
Signals for coherence relations

• Mann and Thompson (1988) found “no reliable, unambiguous signals for any of the relations” that they proposed

• Most researchers have focused on studying connectives/discourse markers as signals for relations
  ▪ Including research on determining a taxonomy of relations based on connectives (Knott and Dale 1994)
Examples of signals

• Unsignalled, unmarked, implicit or underspecified relation
  ▪ Tom quit his job.
  ▪ He was tired of the long hours.

• Signalled, marked or explicit relation
  ▪ Tom quit his job
    because he was tired of the long hours.

• With unsignalled relations, the implicature can be cancelled
  ▪ Tom quit his job.
  ▪ He was tired of the long hours, anyway.
Corpus studies of signalling

- Studies of frequency of signalling vary, but all seem to indicate that a large number of relations are unsignalled.

- Signalling levels in previous work
  - 39% - German (Schauer and Hahn 2001)
  - 28% - miscellaneous genres (RST website, Mann and Taboada 2012)
  - 31% - conversation (Taboada 2006)
  - 43% - newspaper articles (Taboada 2006)
The Cooperative Principle

- Grice (1975)
  - “Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.”

- Quantity maxim
  - Make your contribution as informative as is required, and not more

- If we can assume that speakers work under the Cooperative Principle, then unsignalled relations are such because no signal is necessary (Spooren 1997)

- Unsignalled relations may be more difficult to process, but, under the CP, not impossible
Psycholinguistic evidence for coherence relations

• Are coherence relations cognitive entities?
  ▪ If so, they are recognized by hearers and readers
  ▪ They must follow the Cooperative Principle

• Psycholinguistic evidence
  ▪ There are different types of relations
  ▪ Signalling does have an effect on processing
  ▪ Signalling effects on recall and comprehension?
Different types of relations

• Differences in processing between
  - Causal and additive relations (e.g., Keenan et al 1984, Trabasso and Sperry 1985, Myers et al 1987, …)
  - List and problem-solution relations (Sanders and Noordman 2000)
Effects of signalling

- On processing
  - Faster processing of signalled relations, especially causal and adversative (Haberlandt 1982, Sanders and Noordman 2000)

- On comprehension
  - Better answers to comprehension questions (Millis and Just 1994, Degand and Sanders 2002, Sanders et al. 2007)

- On recall
  - Mixed results
  - Better recall (Meyer et al. 1980)
  - No effect (Sanders and Noordman 2000)
Possible explanations

• Positive effect on processing and comprehension, but mixed in recall

• Sanders and Noordman 2000
  ▪ It could be because the effect of the marker decreases over time, just as the surface representation of the text tends to be lost (Anderson et al 2001, van Dijk and Kintsch 1983)

• Degand and Sanders 2002
  ▪ Different methodologies and reader backgrounds
  ▪ Many studies have shown the effects of signalling are different for different types of readers, tending to help average to underperforming readers (Meyer et al 1980, Britton et al 1982)
Another explanation

- Coherence relations are interpreted whether signalled or not
- Signalling helps immediate processing
- Signalling does not have such a clear effect with longer-term comprehension and recall, because the content is interpreted correctly
Interpretation without signalling

- Are relations correctly interpreted in the absence of signalling?
- Two answers
  - No, and no signalling means no relation, or an underspecified one
  - Yes, and ‘no signalling’ often means there are signals, but they have not been studied in previous research
In Construction Grammar, it is postulated that multiple signals for a phenomenon exist, but often one is sufficient to evoke the frame/construction.

- “In cognition, and therefore in language, we can often evoke a larger complex structure without being given all the component pieces” (Dancygier and Sweetser 2005: 25)

- Frame metonymy
  - A part of the frame may be enough to infer the entire frame (Sweetser and Fauconnier 1996)

- Constructional compositionality (Dancygier and Sweetser 2005)
  - Constructions, as frames, may use components of the construction/frame to stand for the entire construction
    - Certain combinations of verb tenses in juxtaposed clauses to evoke a conditional construction
Outline

• Coherence/rhetorical/discourse relations
• Rhetorical Structure Theory
• Signals for coherence relations
  ▪ The Cooperative Principle
  ▪ Psycholinguistic and cognitive evidence for coherence relations
  ▪ Are there true ‘implicit’ relations?
• Corpus study
  ▪ Taxonomy of cues
  ▪ Annotation
  ▪ Results
• Other corpus-based work
• An application
  ▪ Discourse parsing for sentiment detection
• Conclusions and further work
Back to implicit relations

- Signalling has been understudied
  - Restricted mostly to connectives/conjunctions/discourse markers
- It may well be the case that there are no implicit, i.e., unsignalled relations
- Hypothesis
  - There are no (or very few) ‘implicit’ relations
- Testing the hypothesis empirically
  - For any given relation, can we find clear signals?
Large-scale corpus study

• To test the hypothesis that most relations are in fact explicit, we have undertaken a large-scale corpus study

• RST Discourse Treebank (Carlson et al. 2002)
  ▪ Collection of Wall Street Journal articles
  ▪ 385 articles, 176,000 words
  ▪ A portion of the Penn Treebank
  ▪ Annotated with rhetorical relations
    • 21,789 discourse units, 17,004 relations
Goal

- Annotate every relation with signalling information
- Collate the information and find out
  - Whether there are any true implicit relations
  - Frequency of signalling for groups of relations
  - Co-occurrence of signals and relations
Pilot study

• Initial steps
  ▪ Create a taxonomy of signals
  ▪ Set up annotation software
  ▪ Annotate
  ▪ Extract information
### Taxonomy of signals (8 groups, 39 total)

<table>
<thead>
<tr>
<th>Group</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Discourse marker</td>
<td>Conjunction, adverbial, prepositional phrase</td>
</tr>
<tr>
<td>2. Reference</td>
<td>Personal, demonstrative, comparative</td>
</tr>
<tr>
<td>3. Lexical</td>
<td>Indicative phrase/word</td>
</tr>
<tr>
<td>4. Semantic</td>
<td>Synonym, antonym, hyponym, lexical chain</td>
</tr>
<tr>
<td>5. Morphological</td>
<td>Tense</td>
</tr>
<tr>
<td>6. Syntactic</td>
<td>Non-finite/relative clause, parallel structure</td>
</tr>
<tr>
<td>7. Graphical</td>
<td>Colon, dash, bullet</td>
</tr>
<tr>
<td>8. Genre</td>
<td>Attribution, pyramid scheme</td>
</tr>
</tbody>
</table>
Discourse markers

• The most widely studied
  ▪ If, because, although, however, in the event that, on the one hand

• High levels of signalling for certain relations… but not so high for others
  ▪ Study of the RST Discourse Treebank (Taboada 2006)

<table>
<thead>
<tr>
<th>Relation</th>
<th>Percentage marked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concession</td>
<td>90.35</td>
</tr>
<tr>
<td>Circumstance</td>
<td>66.42</td>
</tr>
<tr>
<td>Result</td>
<td>66.67</td>
</tr>
<tr>
<td>Background</td>
<td>26.56</td>
</tr>
<tr>
<td>Elaboration</td>
<td>9.79</td>
</tr>
<tr>
<td>Summary</td>
<td>4.00</td>
</tr>
</tbody>
</table>

\(n\) 1672 (727 marked)
• Concession
  - Some entrepreneurs say the red tape they most love to hate is red tape they would also hate to lose. They concede that much of the government meddling that torments them is essential to the public good, and even to their own businesses.

• Temporal
  - Nelson Holdings International Ltd. shareholders approved a 1-for-10 consolidation of the company's common stock at a special meeting. At the same time, shareholders approved the adoption of a rights plan and a super-majority voting approval requirement…
Lexical chains

• Contrast
  ▪ The Commerce Department says go, and the Defense Department says stop.

• Elaboration
  ▪ … Since then, deliveries have slumped. GM's car sales dropped 24.8% in mid-October to 69,980, while truck sales fell 26% to 37,860. GM also had dismal results in the first 10 days of the month, while other auto makers reported mixed results.
Until 1980, when Japan joined the $10,000 per capita GNP club of the advanced countries, it was a model developing nation.

- The government built ports, bridges, highways, schools, hospitals and railways.
- When industries were weak, it protected them.
- It gave the Japanese people a value system, based on the rationalization that given the country's lack of natural resources, they must work hard to create value through exports and buy food with the surplus. Individual prosperity inevitably would result.
Syntax: Finiteness

• Circumstance
  ▪ Insisting that they are protected by the Voting Rights Act, a group of whites brought a federal suit in 1987 to demand that the city abandon at-large voting for the nine-member City Council and create nine electoral districts, including four safe white districts.

• Purpose
  ▪ The Court of Appeals for the Federal Circuit was created in 1982 to serve, among other things, as the court of last resort for most patent disputes.
• Condition
  • Had he been a little less gung-ho, “I'd have gotten the thing on the ground and headed for the nearest bar,” Mr. Brown says.
  • For more vivid descriptions of people and places and events in an atmosphere of mystery, read Crichton, or Ian Fleming - or even Tom Clancy!

• List (syntactic parallelism and lexical chains)
  • Exxon's profit fell 9% in the third quarter, hurt by sagging results at two of its three main businesses. Phillips and Arco posted declines. Ashland had a loss. Amerada Hess and Occidental Petroleum had gains.
Elaboration

QUANTUM CHEMICAL Corp.’s plant in Morris, Ill., is expected to resume production in early 1990. The year was misstated in Friday’s editions. (See: “Dividend News: Payout Stalled at Quantum Chemical Corp.—Firm Posts Quarterly Loss, Plans a Stock Dividend to Take Place of Cash—WSJ Oct. 27, 1989)
Summary

I guess it is only befitting that since I wrote a review about a potato masher that I should write about the companion piece that usually goes along with it hand and hand. […]

When I was in Wal-Mart one day I discovered the Oxo 7.25" Steel Swivel Peeler quite by accident. I had ground my old Ecco peeler in the garbage disposal by accident and needed something to replace it. […]

The Oxo 7.25" Steel Swivel Peeler says it is made of steel but there is a major difference between it and the Ecco. The handle of the Oxo 7.25" Steel Swivel Peeler […]

The clean up is easy, like I said above, just place in the dishwasher and slip into the drawer when done.

The next time you are looking for a peeler give Oxo 7.25" Steel Swivel Peeler a try it’s worth the extra $2.00 that you will spend.

Thanks for reading.
Analysis so far

- Relations annotated
  - 1,306 (about 10% of the corpus)
- Annotation procedure
  - RSTTool
  - Extract XML
  - Add information onto XML
## Results

<table>
<thead>
<tr>
<th>Signalling</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relations signalled by DMs</td>
<td>251</td>
<td>19.22%</td>
</tr>
<tr>
<td>Relations signalled by other signals</td>
<td>878</td>
<td>67.23%</td>
</tr>
<tr>
<td>Unsignalled relations (implicit?)</td>
<td>177</td>
<td>13.55%</td>
</tr>
<tr>
<td>Total</td>
<td>1,306</td>
<td></td>
</tr>
</tbody>
</table>

- 251 relations signalled by DMs
  - 58 markers in total
- 878 signalled by other signals
  - Varied, spanning all categories
  - 64 relations with multiple signals
Results, signalling by relation

<table>
<thead>
<tr>
<th>Relation group</th>
<th>DM</th>
<th>Other signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribution</td>
<td>-</td>
<td>Syntactic, lexical, genre</td>
</tr>
<tr>
<td>Background</td>
<td>Yes</td>
<td>Lexical, semantic, morphological, syntactic</td>
</tr>
<tr>
<td>Cause</td>
<td>Yes</td>
<td>Lexical</td>
</tr>
<tr>
<td>Comparison</td>
<td>Yes</td>
<td>Lexical, semantic</td>
</tr>
<tr>
<td>Condition</td>
<td>Yes</td>
<td>Lexical, syntactic</td>
</tr>
<tr>
<td>Contrast</td>
<td>Yes</td>
<td>Syntactic, semantic, genre</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Yes</td>
<td>Lexical, semantic, syntactic, genre, graphical</td>
</tr>
<tr>
<td>Enablement</td>
<td>-</td>
<td>Syntactic, lexical</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Yes</td>
<td>Lexical, syntactic</td>
</tr>
<tr>
<td>Joint</td>
<td>Yes</td>
<td>Syntactic, lexical, semantic, graphical</td>
</tr>
<tr>
<td>Manner/means</td>
<td>Yes</td>
<td>Lexical, syntactic</td>
</tr>
<tr>
<td>Topic-comment</td>
<td>Yes</td>
<td>Semantic, lexical</td>
</tr>
<tr>
<td>Summary</td>
<td>-</td>
<td>Semantic, graphical</td>
</tr>
<tr>
<td>Temporal</td>
<td>Yes</td>
<td>Lexical</td>
</tr>
<tr>
<td>Topic change</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Same unit</td>
<td>Yes</td>
<td>Syntactic, graphical</td>
</tr>
</tbody>
</table>
## Results, ambiguous signals

<table>
<thead>
<tr>
<th>DM</th>
<th>Frequency</th>
<th>Common relation group</th>
</tr>
</thead>
<tbody>
<tr>
<td>And</td>
<td>70</td>
<td>Cause, Elaboration, Joint, Explanation, Temporal</td>
</tr>
<tr>
<td>But</td>
<td>35</td>
<td>Contrast, Elaboration</td>
</tr>
<tr>
<td>If</td>
<td>13</td>
<td>Conditional</td>
</tr>
<tr>
<td>When</td>
<td>10</td>
<td>Background, Circumstance</td>
</tr>
<tr>
<td>However</td>
<td>9</td>
<td>Contrast, Antithesis</td>
</tr>
<tr>
<td>Because</td>
<td>8</td>
<td>Cause, Consequence</td>
</tr>
</tbody>
</table>

... |

<table>
<thead>
<tr>
<th>Other signals</th>
<th>Frequency</th>
<th>Common relation group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic</td>
<td>573</td>
<td>Attribution, Elaboration, Enablement</td>
</tr>
<tr>
<td>Semantic</td>
<td>163</td>
<td>Elaboration</td>
</tr>
<tr>
<td>Reference</td>
<td>51</td>
<td>Background, Comparison, Elaboration</td>
</tr>
<tr>
<td>Graphical</td>
<td>37</td>
<td>Elaboration, Summary</td>
</tr>
</tbody>
</table>

...
Results, summary

- A variety of signals for each relation
- A variety of relations for each signal
- Clearly, no one-to-one mapping
- Yet, we solve these ambiguities in discourse
- Main conclusion
  - Signalling is more widespread than previously acknowledged
Implicit relations

• What about the 177 relations for which we could not find a signal?
  1. Error in the annotation; no relation actually exists
  2. Questionable discourse relations (not included in the original RST taxonomy)
     • Comment
     • Topic-shift
  3. We, as annotators, get a ‘sense’ that there is a signal
     • Tenuous entity relations
     • World knowledge that leads to an entity relation
     • Constructional compositionality
Outline

• Coherence/rhetorical/discourse relations
• Rhetorical Structure Theory
• Signals for coherence relations
  ▪ The Cooperative Principle
  ▪ Psycholinguistic and cognitive evidence for coherence relations
  ▪ Are there true ‘implicit’ relations?
• Corpus study
  ▪ Taxonomy of cues
  ▪ Annotation
  ▪ Results
• Other corpus-based work
• An application
  ▪ Discourse parsing for sentiment detection
• Conclusions and further work
Other corpus-based work

- Signalling one of the ‘hottest’ topics in discourse
- Efforts aimed at explaining
  - How relations are signalled
  - How signalling interacts with other phenomena
    - Cohesion and reference
    - Genre
- Corpus-based work
  - RST Treebank
  - Penn Discourse Treebank
The Penn Discourse Treebank (PDTB)

- Large corpus annotated for *local* discourse relations
  - Also Wall Street Journal
  - Full set, 2,000+ articles, 1 million words
  - 40,600 relations
- Relations classified as
  - Explicit
  - Implicit
    - Implicit marker that can easily be added
    - Alternative lexicalization: A marker is not simply omitted; other forms of signalling are used
    - Entity-based link (reference, cohesion)
Work using the PDBT

• Quite a few studies addressing
  ▪ Types of relations and their markers (Prasad et al. 2010)
  ▪ How marking and genre are related (Webber 2009)
  ▪ How ambiguous relation marking is (Pitler et al. 2009, Pitler and Nenkova 2009)

• A few conclusions
  ▪ Markers are relatively infrequent
    • About 45% of local relations are marked with connectives (Prasad et al. 2008)
  ▪ Markers are ambiguous between discourse and non-discourse uses (Pitler and Nenkova 2009)
    • Nevertheless, classification systems can be built to identify relations automatically with relative accuracy (Louis et al. 2010)
Other work

• By Gisela Redeker and colleagues
• On a corpus of Dutch texts, expository and persuasive genres
• Interaction between genre and relation marking (only discourse markers)
• Levels of marking: Intra-, inter- and multi-sentential relations
• Conclusions so far
  ▪ Connectives particularly useful for the identification of local relations
  ▪ Some relations do have DMs at the multi-sentential level
    • Conjunction, Concession, Non-Volitional Cause and Non-Volitional Result
Automatically labelling implicit relations (Marcu and Echihabi 2002)

- It is relatively easy to identify explicit relations
- A method: take explicit examples, remove the connective and use that data to train a classifier (Machine Learning)
  - Hypothesis: other signals in the relation will help identify it, in the absence of the connective
- Results: relative success

On the other hand, other work questions the method (Sporleder and Lascarides 2008)

- Explicit and implicit relations seem to be qualitatively different
- Using explicit examples will not tell us anything about implicit ones
Outline

• Coherence/rhetorical/discourse relations
• Rhetorical Structure Theory
• Signals for coherence relations
  ▪ The Cooperative Principle
  ▪ Psycholinguistic and cognitive evidence for coherence relations
  ▪ Are there true ‘implicit’ relations?
• Corpus study
  ▪ Taxonomy of cues
  ▪ Annotation
  ▪ Results
• Other corpus-based work
• An application
  ▪ Discourse parsing for sentiment detection
• Conclusions and further work
Why the interest in ‘implicit’ relations

• Theoretical
  ▪ If we do identify relations, we should know how it’s done
  ▪ What cues are there for us to identify them
  ▪ When can cues be misinterpreted

• Practical
  ▪ Automatically extract relations from text for various applications
    ▪ Summarization
    ▪ Sentiment detection
Sentiment detection

- Classification of texts based on subjective content (=sentiment)
  - Positive
  - Negative
- Input a text, and produce a numeric value that expresses its subjective content → the text’s sentiment
  - -3: quite negative
  - -5: very negative
  - 2: somewhat positive
  - …
Applications

- Search for a movie
  - Results are returned in categories: positive and negative reviews

- Consumer reviews
  - Products are ranked according to their reviews

- Market intelligence
  - What are people saying about a new product? What do they think about a company?

- Politics
  - Opinions about a candidate, a policy, a new piece of legislation (maybe also over time)
It could have been a great movie. It could have been excellent, and to all the people who have forgotten about the older, greater movies before it, will think that as well. It does have beautiful scenery, some of the best since Lord of the Rings. The acting is well done, and I really liked the son of the leader of the Samurai. He was a likeable chap, and I hated to see him die... But, other than all that, this movie is nothing more than hidden rip-offs.
Concession

It could have been a great movie. It could have been excellent, and to all the people who have forgotten about the older, greater movies before it, will think that as well. It does have beautiful scenery, some of the best since Lord of the Rings. The acting is well done, and I really liked the son of the leader of the Samurai. He was a likeable chap, and I hated to see him die.

But, other than all that, this movie is nothing more than hidden rip-offs.
We had a terrific time in Key West as long as we didn't spend a minute more than we had to in our so-called 'luxury' rooms. I will never recommend this hotel to any family, friend, co-worker, or stranger on the street. Wyndham has a lot of work to do if this is what is considered top-drawer.
Julia Roberts without a doubt, is a good actress depending on what film she does....

Sure, her smile will make you go to the theaters but does it help the film. No.

Current research

• Build a discourse parser that can automatically identify relations
  ▪ Using ‘standard’ discourse markers
  ▪ But also some of the other signals

• Classify relations
  ▪ Not only at the local level, but also more higher-level relations

• Use the resulting trees in sentiment detection
  ▪ Rely only on the nucleus part of the relation
  ▪ Calculate sentiment depending on relations
    • In Condition, average the two parts
    • In Concession, weaken or strengthen the opinion depending on the polarity of the components
  • Asher et al. 2009, Trnavac and Taboada 2012
Conclusions

• Clearly, there are fewer ‘implicit’ relations than we have acknowledged thus far

• It is possible that there are no implicit relations at all
  ▪ Large-scale corpus study
  ▪ Both local and global relations
  ▪ Large set of signals

• Ours is the only corpus-based work that looks at both local and global relations, and at signals beyond discourse markers

• A test: can we build a system that recognizes most relations in a text?
  ▪ Current work on discourse parsing
Future work

• Continuation of the annotation, entire RST-DT corpus

• Development of a discourse parsing system
  ▪ Starting with automatic segmentation into minimal discourse units (Tofiloski, Brooke and Taboada 2009)


Maite Taboada
mtaboada@sfu.ca
http://www.sfu.ca/~mtaboada
http://www.sfu.ca/~mtaboada/research/pubs.html