Computing Discourse Structure and Discourse Semantics
Bonnie Webber
School of Informatics
University of Edinburgh
Work carried out with Aravind Joshi (UPenn), Matthew Stone (Rutgers U), and Alistair Knott (U Otago), and members of the UPenn DL-TAG group (Cassie Creswell, Kate Forbes, Eleni Miltsakaki and Rashmi Prasad).

1 Assumption

There are several mechanisms used in computing discourse structure and discourse semantics:

2. Compositionality, involving clauses and discourse connectives (in English - "so", "because", "but", etc.).
3. Interactions between discourse relations.
5. Comparison with other accounts of discourse structure and semantics.
6. DIAG.
7. Discourse parsing with DLTAG.
8. Open problems.

The aim of research reported here include:

1. Understanding the extent to which anaphora is a mechanism for conveying meaning in discourse.
2. Understanding how meaning conveyed anaphorically interacts with meaning conveyed compositionally and inferentially.
3. Understanding the extent to which anaphora is a mechanism for conveying meaning in discourse.
4. Developing a way of representing and processing discourse that acknowledges the contributions of all three mechanisms.

Outline

1. Background: Abstract Objects and Discourse Relations
2. Distinguishing Discourse Adverbials from Structural Connectives
3. Interactions between Discourse Relations
4. Empirical Studies
5. Comparison with other accounts of discourse structure and semantics
6. DIAG
7. Discourse Parsing with DLTAG
8. Open Problems
I. Background: Abstract Objects

Introduction: Abstract Objects (AO) is the term that Asher (1993) uses for the possible referents of non-NP constituents, including a sequence of clauses (censed as a single object). Asher (1993) notes that for the possibility of abstract reference, the referent is not an entity in the world but rather a relation between appropriate sorts.

5. Classes of AO

Assumption: Discourse relations are relations between AO of appropriate sorts.

II. Discourse Adverbials and Structural Connectives

Structural Connectives: Coordinating and subordinating conjunctions, paired conjunctions (On the one hand...; On the other...).

A. Structural connectives allow stretching of dependency relations.

b. Although John is generous, he's very hard to find.

b. Although John is generous— if you need some money, you only have to ask him for it—he's very hard to find.

b. Although John is generous, his baby son kept him awake all night.

5. Discourse Relations

Although John broke his arm, so he can't cycle to work.

so he can't cycle to work.

6. Additional discourse relations are conveyed by discourse adverbials.

On the one hand, Sam likes beans. Not only does he eat them for dinner, but he also eats them for snacks. On the other hand, they make him ill.
Discourse adverbials admit crossing dependencies.

(9) a. Although John is very generous –
   b. you only have to ask him for it.
   c. he's very hard to find –
   d. you need some money –

b. Structural Conns do not allow crossing of Pred-Ary dependencies.

(10) a. On the one hand, Fred likes beans.
    b. Not only does he eat them for dinner.
    c. On the other hand, he's allergic to them.
    d. But he also eats them for breakfast and snacks.

    b. So he ordered three cases of the '97.
    c. But he had to cancel the order.
    d. Because he discovered he was broke.
a. High heels are fine for going to the theater.

b. But wear comfortable shoes if instead you plan to go to the zoo.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.

Anaphoric dependencies are generally not considered to be structural.
Discourse adverbials behave like anaphors:

- **Coreferential** – definite pronouns and NPs that denote one or more discourse entities in the current discourse context.

- **Indirect anaphora** – aka bridging anaphora, where the anaphor (usually a definite NP) denotes a discourse entity associated with a discourse entity in the current discourse context.

Consider:

1. Sue grabbed one phone as Tom darted to the other phone.
2. Sue grabbed the other phone as Tom darted to the other phone.

This suggests that discourse adverbials are accessing discourse context and are thus anaphoric. This doesn't exhaust the space of constructs that derive all or part of their sense from the discourse context and are thus anaphoric.

What kind of anaphor is a discourse adverbial? Between adjacent clauses:

- Possible cases of *adjacent objects* (other than standard structural connection)
- *Change* (e.g., *she was happy* changing to *she was sad*)
- *Every person selling the Big Issue*

Also consider:

1. Sue lifted the receiver as Tom darted to the other phone.
2. Sue lifted the other phone as Tom darted to the other phone.

This suggests that discourse adverbials are accessing discourse context and are thus anaphoric.

We have called these *lexically-specified anaphors*.

A discourse adverbial is an expression that maps directly onto discourse-contextual discourse entities.

Every farmer who owns a donkey feeds it rutabagas.

This is not associative anaphora, where $\alpha \text{ assoc}(\epsilon)$. This is not associative anaphora, where $\alpha \text{ assoc}(\epsilon)$. This is not associative anaphora, where $\alpha \text{ assoc}(\epsilon)$.

What kind of anaphor is a discourse adverbial?
If we now have three types of anaphora, not two:

- Coreference: $e_\alpha = e_r$
- Indirect anaphora: $e_\alpha \# \text{assoc}(e_r)$
- Lexically-specified anaphora: $e_\alpha = f_\alpha \% e_i \& \text{assoc}(e_r)$

Where do Discourse Adverbials fit in?

Nothing requires that the source of $e_3$ be in NPs, or that the antecedent be a pronoun.

What is the source of $e_3$?

For $e_3$ to be a type of lexically-specified anaphor, $e_3$ must be governed by $\text{agg}(e_2)$.
Then (27a) John loves Barolo.

b) So he ordered three cases of the '97.

c) But he had to cancel the order because he found he was broke.

α = the anaphoric expression –

Rα = the relation name linked with α –

Sα = the matrix clause/sentence containing α –

σ = the interpretant of α as an abstract object –

β = find(j, e5), where e5: broke(j)

Unresolved interpretation of (27d):

λx. Rα(x, EV)

σ

λx. after(x, EV)

e4:

Resolved interpretation:

EV

e2: order(j, c1) from (27b)

after(e4, EV)

+ after(e4, e2)

Example: "Otherwise"

"Otherwise" conveys that the complement of its anaphorically-derived argument serves as the condition under which the interpretation of its structural argument holds.

Otherwise: "Otherwise"

Example: "Otherwise"

If the light is red, stop.

Otherwise you'll get a ticket.

α = otherwise

Rα = if

S = you get a ticket

σ = e3, where e3: get ticket(you)

Unresolved interpretation:

λx. if(VE28, x)

e3:

Resolved interpretation:

EV

e1: red(light1)

If you do something other than stop, you'll get a ticket.

α = otherwise

Rα = if

S = stop(you)

σ = e2, where e2: stop(you)

Unresolved interpretation:

λx. if(VE29, x)

e3:

Resolved interpretation:

EV

e1: red(light)

If the light is not red, go straight on.

If you do something other than stop, you'll get a ticket."

Example: "Then"

Example: "Otherwise"
What does this treatment achieve?

Standard account: Both structural connectives and discourse adverbials indicate the relation holding between the AO from adjacent DUs—e.g. "because"—explanation relation, "but"—contrast relation, "in general"—generalization relation, "in other words"—elaboration relation, "otherwise"—otherwise relation, "therefore"—result relation.

Revised account: (1) The standard account holds for structural connectives. (2) A discourse adverbial conveys a separate relation between the AO associated with its matrix clause and an AO derived from the previous discourse. (3) These relations can interact in interesting ways.

III. Interactions between Discourse Relations

\[ R_a \subseteq \text{disourse adversarial relation} \]

\[ R_b \subseteq \text{disourse adversarial relation} \]

\[ R_c \subseteq \text{disourse adversarial relation} \]

\[ R_d \subseteq \text{disourse adversarial relation} \]

\[ R_e \subseteq \text{disourse adversarial relation} \]

\[ R_f \subseteq \text{disourse adversarial relation} \]

\[ R_g \subseteq \text{disourse adversarial relation} \]

\[ R_h \subseteq \text{disourse adversarial relation} \]

\[ R_i \subseteq \text{disourse adversarial relation} \]

\[ R_j \subseteq \text{disourse adversarial relation} \]

\[ R_k \subseteq \text{disourse adversarial relation} \]

\[ R_l \subseteq \text{disourse adversarial relation} \]

\[ R_m \subseteq \text{disourse adversarial relation} \]

\[ R_n \subseteq \text{disourse adversarial relation} \]

\[ R_o \subseteq \text{disourse adversarial relation} \]

\[ R_p \subseteq \text{disourse adversarial relation} \]

\[ R_q \subseteq \text{disourse adversarial relation} \]

\[ R_r \subseteq \text{disourse adversarial relation} \]

\[ R_s \subseteq \text{disourse adversarial relation} \]

\[ R_t \subseteq \text{disourse adversarial relation} \]

\[ R_u \subseteq \text{disourse adversarial relation} \]

\[ R_v \subseteq \text{disourse adversarial relation} \]

\[ R_w \subseteq \text{disourse adversarial relation} \]

\[ R_x \subseteq \text{disourse adversarial relation} \]

\[ R_y \subseteq \text{disourse adversarial relation} \]

\[ R_z \subseteq \text{disourse adversarial relation} \]

\[ R_{aa} \subseteq \text{disourse adversarial relation} \]

\[ R_{ab} \subseteq \text{disourse adversarial relation} \]

\[ R_{ac} \subseteq \text{disourse adversarial relation} \]

\[ R_{ad} \subseteq \text{disourse adversarial relation} \]

\[ R_{ae} \subseteq \text{disourse adversarial relation} \]

\[ R_{af} \subseteq \text{disourse adversarial relation} \]

\[ R_{ag} \subseteq \text{disourse adversarial relation} \]

\[ R_{ah} \subseteq \text{disourse adversarial relation} \]

\[ R_{ai} \subseteq \text{disourse adversarial relation} \]

\[ R_{aj} \subseteq \text{disourse adversarial relation} \]

\[ R_{ak} \subseteq \text{disourse adversarial relation} \]

\[ R_{al} \subseteq \text{disourse adversarial relation} \]

\[ R_{am} \subseteq \text{disourse adversarial relation} \]

\[ R_{an} \subseteq \text{disourse adversarial relation} \]

\[ R_{ao} \subseteq \text{disourse adversarial relation} \]

\[ R_{ap} \subseteq \text{disourse adversarial relation} \]

\[ R_{aq} \subseteq \text{disourse adversarial relation} \]

\[ R_{ar} \subseteq \text{disourse adversarial relation} \]

\[ R_{as} \subseteq \text{disourse adversarial relation} \]

\[ R_{at} \subseteq \text{disourse adversarial relation} \]

\[ R_{au} \subseteq \text{disourse adversarial relation} \]

\[ R_{av} \subseteq \text{disourse adversarial relation} \]

\[ R_{aw} \subseteq \text{disourse adversarial relation} \]

\[ R_{ax} \subseteq \text{disourse adversarial relation} \]

\[ R_{ay} \subseteq \text{disourse adversarial relation} \]

\[ R_{az} \subseteq \text{disourse adversarial relation} \]

\[ R_{ba} \subseteq \text{disourse adversarial relation} \]

\[ R_{bb} \subseteq \text{disourse adversarial relation} \]

\[ R_{bc} \subseteq \text{disourse adversarial relation} \]

\[ R_{bd} \subseteq \text{disourse adversarial relation} \]

\[ R_{be} \subseteq \text{disourse adversarial relation} \]

\[ R_{bf} \subseteq \text{disourse adversarial relation} \]

\[ R_{bg} \subseteq \text{disourse adversarial relation} \]

\[ R_{bh} \subseteq \text{disourse adversarial relation} \]

\[ R_{bi} \subseteq \text{disourse adversarial relation} \]

\[ R_{bj} \subseteq \text{disourse adversarial relation} \]

\[ R_{bk} \subseteq \text{disourse adversarial relation} \]

\[ R_{bl} \subseteq \text{disourse adversarial relation} \]

\[ R_{bm} \subseteq \text{disourse adversarial relation} \]

\[ R_{bn} \subseteq \text{disourse adversarial relation} \]

\[ R_{bo} \subseteq \text{disourse adversarial relation} \]

\[ R_{bp} \subseteq \text{disourse adversarial relation} \]

\[ R_{bq} \subseteq \text{disourse adversarial relation} \]

\[ R_{br} \subseteq \text{disourse adversarial relation} \]

\[ R_{bs} \subseteq \text{disourse adversarial relation} \]

\[ R_{bt} \subseteq \text{disourse adversarial relation} \]

\[ R_{bu} \subseteq \text{disourse adversarial relation} \]

\[ R_{bv} \subseteq \text{disourse adversarial relation} \]

\[ R_{bw} \subseteq \text{disourse adversarial relation} \]

\[ R_{bx} \subseteq \text{disourse adversarial relation} \]

\[ R_{by} \subseteq \text{disourse adversarial relation} \]

\[ R_{bz} \subseteq \text{disourse adversarial relation} \]

\[ R_{ca} \subseteq \text{disourse adversarial relation} \]

\[ R_{cb} \subseteq \text{disourse adversarial relation} \]

\[ R_{cc} \subseteq \text{disourse adversarial relation} \]

\[ R_{cd} \subseteq \text{disourse adversarial relation} \]

\[ R_{ce} \subseteq \text{disourse adversarial relation} \]

\[ R_{cf} \subseteq \text{disourse adversarial relation} \]

\[ R_{cg} \subseteq \text{disourse adversarial relation} \]

\[ R_{ch} \subseteq \text{disourse adversarial relation} \]

\[ R_{ci} \subseteq \text{disourse adversarial relation} \]

\[ R_{cj} \subseteq \text{disourse adversarial relation} \]

\[ R_{ck} \subseteq \text{disourse adversarial relation} \]

\[ R_{cl} \subseteq \text{disourse adversarial relation} \]

\[ R_{cm} \subseteq \text{disourse adversarial relation} \]

\[ R_{cn} \subseteq \text{disourse adversarial relation} \]

\[ R_{co} \subseteq \text{disourse adversarial relation} \]

\[ R_{cp} \subseteq \text{disourse adversarial relation} \]

\[ R_{cq} \subseteq \text{disourse adversarial relation} \]

\[ R_{cr} \subseteq \text{disourse adversarial relation} \]

\[ R_{cs} \subseteq \text{disourse adversarial relation} \]

\[ R_{ct} \subseteq \text{disourse adversarial relation} \]

\[ R_{cu} \subseteq \text{disourse adversarial relation} \]

\[ R_{cv} \subseteq \text{disourse adversarial relation} \]

\[ R_{cw} \subseteq \text{disourse adversarial relation} \]

\[ R_{cx} \subseteq \text{disourse adversarial relation} \]

\[ R_{cy} \subseteq \text{disourse adversarial relation} \]

\[ R_{cz} \subseteq \text{disourse adversarial relation} \]
If the light is red, stop, and otherwise go straight on.

Case 3: $R$ is parasitic on $R'$. It is the abstract object interpretation of "If the light is red, stop.

Inter-clausal "for example": Abstraction on discourse relation.

Relation $R$: exemplification $(e, X' \cdot \text{explanation}(X', e))$
Relation $R'$: exemplification $(e, X' \cdot \text{explanation}(X', e))$
Relation $R''$: because $(e, X' \cdot \text{explanation}(X', e))$

If the light is something other than red, go straight on.

The light is red stop and otherwise go straight on.
You shouldn't trust John. For example, he never returns what he borrows.

But what about cases where only exemplify seems to hold?

Case 2: \( R^2 \) is a defeasible rule that incorporates \( R^1 \).

Although Certo was called the pedestal of beauty, she never married.

The preceding discourse shows that only \( R^2 \) is applicable in this situation. Certo is the pedestal of beauty, but she never married. Hence, she is not the measure of beauty. The rule that incorporates \( R^1 \) seems to hold.
Interactions between Discourse Relations

1. $\sigma$ separately serves as an argument to both $R_{\alpha}$ and $R_{\beta}$; $R_{\alpha}$ is parasitic on $R_{\beta}$; $R_{\alpha}$ is a defeasible rule that incorporates $R_{\beta}$.

IV. Empirical Studies

If we want to claim that anaphor plays a larger part in discourse semantics than

Study 1: Differences between Connectives [Creswell et al., 2002]

Summary: Interactions between Discourse Relations
**Annotation Scheme**

**Features of left argument**
- Syntactic type: main, subordinate, phrasal constituent (XP), a sequence of main clauses

**Features of right argument**
- Combines with: punctuational one, conjunctions, other adverbial connectives
- Position of connective: initial, medial, clause-final

All features can be derived from a parsed version of the corpus.

**Inter-annotator consistency**

A second study was done with 4 annotators and the added 25 tokens of each 3 connectives.

**Results of Initial Annotation: Summary**

- Relative percentages of tags remained stable, indicating that the patterns may be systematic enough to allow automatic

- Position of connective: initial, medial or clause-final

- Conclusions: punctuation alone, conjunctions, other adverbial

- Syntactic type: main, subordinate, phrasal constituent (XP), a sequence of main clauses

- Connectives of left argument

- Connectives of right argument

- Syntactic type: main, subordinate, phrasal constituent (XP), a sequence of main clauses

- Connectives of left argument

- Connectives of right argument
Majority agreement (3-way or better) is 88% for
nevertheless, 92% for in addition, and 96% for as a result.

4-way agreement is 50% in all cases.

Study 2: Towards Feature-based Anaphor Resolution

With respect to procedures for resolving anaphoric discourse connectives, we are

Some Results
As a readverbial, “instead” gets its second argument anaphorically, from the discourse context. However, note every text span in the previous discourse appears able to provide an anaphoric reference of a modal auxiliary (should) or a monotonous-decreasing quantifier (many).

(47) a. John ate an apple. Instead, he ate a pear.
    b. John decided to eat a pear. Instead, he ate an apple.
    c. John won’t eat fruit. Instead, he eats only candy bars and potato chips.

But no one has established what kind of phrases/clauses suggest alternatives and which don’t. And presumably, the ability to suggest alternatives is only one of several factors relevant to identifying the intended antecedent of the anaphoric argument of “instead.”

Step 1: Annotate Antecedents

Pairs of annotators separately examined 100 successive instances of bare “instead” in the Penn TreeBank (Wall Street Journal) and recorded the minimal text span containing the antecedent of its anaphoric argument. There was agreement in 97/100 cases. The other 3 cases were excluded from further analysis.

Step 2: Choose Annotation Features

Features were chosen that were observed in examples of “instead” that we had observed in examples of “instead” that we had chosen.

Further analyses:

- whether the antecedent is embedded in a higher clause (Embed)
- presence of a modal auxiliary (should)
- presence of a monotonous-decreasing quantifier (many)
- clausal negation (48) John couldn’t sleep. Instead, he wrote code.
- verbal negation (49) No one could sleep. Instead, everyone wrote code.
- presence of a monotone-decreasing quantifier (51) Few students like to do homework. Instead, they would rather party.

- presence of a modal auxiliary (should)
- presence of a monotonous-decreasing quantifier (many)
- clausal negation (48) John couldn’t sleep. Instead, he wrote code.
- verbal negation (49) No one could sleep. Instead, everyone wrote code.
- presence of a monotone-decreasing quantifier (51) Few students like to do homework. Instead, they would rather party.
Step 3: Annotate Features of Agreed Upon Antecedents

<table>
<thead>
<tr>
<th>Feature</th>
<th>YES of 97</th>
<th>NO of 97</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal neg.</td>
<td>37 (38%)</td>
<td>60 (62%)</td>
</tr>
<tr>
<td>Subj. neg.</td>
<td>5 (5%)</td>
<td>92 (95%)</td>
</tr>
<tr>
<td>Obj. neg.</td>
<td>10 (10%)</td>
<td>82 (85%)</td>
</tr>
<tr>
<td>MDQ</td>
<td>1 (1%)</td>
<td>96 (99%)</td>
</tr>
<tr>
<td>Modal</td>
<td>12 (12%)</td>
<td>85 (88%)</td>
</tr>
<tr>
<td>Condit.</td>
<td>1 (1%)</td>
<td>96 (99%)</td>
</tr>
<tr>
<td>Embed.</td>
<td>57 (59%)</td>
<td>40 (41%)</td>
</tr>
</tbody>
</table>

N.B. Antecedents could display more than one of the above features.

Distribution of Features of PCAs

<table>
<thead>
<tr>
<th>Feature</th>
<th>YES of 169</th>
<th>NO of 169</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal neg.</td>
<td>21 (12%)</td>
<td>148 (88%)</td>
</tr>
<tr>
<td>Subj. neg.</td>
<td>8 (5%)</td>
<td>161 (95%)</td>
</tr>
<tr>
<td>Obj. neg.</td>
<td>6 (4%)</td>
<td>139 (82%)</td>
</tr>
<tr>
<td>MDQ</td>
<td>0 (0%)</td>
<td>169 (100%)</td>
</tr>
<tr>
<td>Modal</td>
<td>17 (10%)</td>
<td>152 (90%)</td>
</tr>
<tr>
<td>Condit.</td>
<td>0 (0%)</td>
<td>169 (100%)</td>
</tr>
<tr>
<td>Embed.</td>
<td>14 (8%)</td>
<td>155 (91%)</td>
</tr>
</tbody>
</table>

Obvious differences between Antecedents and PCAs

1. Negation of the verb or one of its arguments is over 2.5 times more common in the antecedent of “instead” than in potentially competing antecedents – 52/97 times (K 53%) versus 35/169 times (K 20%).

2. The antecedent of the anaphoric argument of “instead” is over 7 times more common than possibly competing antecedents – 3/97 times (5%) versus 23/169 times (14%).

Step 4: Annotate Features of Agreed Upon Antecedents
Other differences between antecedents and PCAs

1. In the case of antecedents, the main verb of the embedding clause included 
   
   - insist, 
   - abandon, 
   - doubt, 
   - expect, 
   - tell, 
   - say, 
   - concede, 
   - want, 
   - be appropriate, etc.

   Some embedded PCAs were also dominated by these main verbs.

2. Certain verbs themselves appear to suggest alternatives, independent of explicit
   
   - instead, 
   - rather, 
   - otherwise, 
   - otherwise.

   Such alternatives are often used to express a change in attitude or perspective.

3. Other lexis-syntactic elements may influence alternatives:

   - Location of the antecedent of the anaphoric argument of "instead" (cf. SS, PS, XP, etc.)
   - Syntactic nature of the anaphor (e.g., Main, Subordinate, XP, etc.)
   - Inferred character of the antecedent (e.g., "normal", "opportunity", etc.)

   After additional annotations, we should be able to
   
   - Induce a decision procedure based on the set of features and assess its
   - Accuracy and its dependency on these features
   - Synthesize the results of the antecedent (e.g., Main, Subordinate, XP, etc.)

   Need to broaden the range of features being considered: it does not correspond

   to any previously defined set!
V. Comparison with other approaches

All approaches to discourse structure and semantics address the obvious fact that discourse conveys more than the meaning of its component clauses/sentences. In several approaches, such as Rhetorical Structure Theory (RST) (Hobbs, Stickel, et al., 1993; Carlson et al., 2003), Linguistic Discourse Model (Polanyi, 1988; Polanyi & vanden Berg, 1996) and Structured DRT (Lascarides & Asher, 1991, 1993), at least part of the additional meaning is associated with some type of discourse relation holding between non-overlapping, adjacent text spans. This earned mixed reviews from Wall Street analysts. Xerox Corporation’s third-quarter net income grew 6.2% on 7.3% higher revenue. This earned mixed reviews from Wall Street analysts.

More specifically, a discourse relation is a propositional connection between text spans. In RST, the minimal units of discourse are text spans that represent the thematic roles of participants in a situation, where situations are grounded in syntax and semantics. This has led to a tension between theory and practice. In order to implement these approaches to discourse, practitioners have found it difficult to separate discourse and clause-level syntax. This led to a tension between theory and practice. The 6.2% growth of Xerox Corporation’s third-quarter net income on 7.3% higher revenue, which earned mixed reviews from Wall Street analysts, is an example. RST makes the following assumptions:

1. The terminal nodes of discourse structure are text spans that represent the minimal units of discourse.
2. Schemas (abstract patterns) define how adjacent, non-overlapping text spans relate to one another. Non-terminal nodes come from schema applications.
3. Most schemas specify one or more discourse relations between non-overlapping, adjacent text spans.
4. Some schemas specify one or more discourse relations between overlapping, adjacent text spans.

Implementation of RST

Implementation of RST includes:

- Text generation: Hovy, 1988; Moore, 1993
- Text validation: McNamara, 1999; McNamara & Edmonds, 2002; Carlson et al., 2003
- Consequence

Example: RST makes the following assumptions:

A comparison with other approaches
Nominalised clauses are not treated as minimal discourse units. So (69) is annotated as a single discourse unit, while (66)–(68), as two units.

We believe this is simpler to recognize. That is, the same range of meaning can be captured within a single discourse unit.

Another assumption about discourse syntax: They are assumptions that RST is attempting to set out a compositional discourse semantics, based on the structure of the discourse.

We believe that it is difficult to adhere to both in analysis and in generation.

The results underscore Sears’s difficulties in implementing the “everyday low pricing” strategy.

Instead, the Treasury announced it would sell $2 billion of 51-day cash management bills today and said that the weekly sale of $15.6 billion of three-month and six-month bills will take place today, as usual.

There is just so much going on that it’s difficult to pick just one. There is just so much going on that it’s difficult to pick just one.

The results underscore Sears’s difficulties in implementing the strategy.

We believe that it is simpler to recognize.

Some, possibly all, the same range of meaning can be captured within a single discourse unit.

As the theory, RST is attempting to set out a compositional discourse semantics, based on the structure of the discourse.

That is, the same range of meaning can be captured within a single discourse unit.
VI. DLTAG: A Lexicalized Grammar for Discourse

How to incorporate this treatment of discourse relations into a computational approach to discourse semantics and discourse structure?

Incorporate into a sentence-level grammar since, syntactically, both discourse adverbials and structural connectives fall within the sentence.

Downside: Sentence-level grammars don't provide for forming the meaning of multi-clausal units that cross sentence-level punctuation.

Take a different approach to discourse-level grammar and semantics, one for discourse anchor has = associated tree structures (environmental needs) one for each anchor.

73

Extends sentence-level grammar and its associated semantic mechanisms to discourse, using either


- the unique foot node, used for adorning.

- a sentence-level grammar and its associated semantic mechanisms to discourse.

Downside: This requires discourse semantics reaching deeper and deeper into a discourse anchor.

Downside: Sentence-level grammars don't provide for forming the meaning of multi-clausal units that cross sentence-level punctuation.

Downside: Sentence-level grammars don't provide for forming the meaning of discourse relations into a computational model.

- a Lexicalized Categorial Grammar for Discourse.
Unlikethewidevarietyoftreesneededattheclauselevel,extendingLTAG
todiscoursehasrequiredonlyafewelementarytreestructures,possiblybecause
discoursedoesn'texploitstructuralvariationinwaysthatclause-levelsyntaxdoes.

Treefamily

ofinitialtreesformain/subordinateconstructions,allofwhichshare
thesamepredicate-argumentdependencies."subconj"standstheparticular
subordinateconjunctionthatanchorsthetree.

On the one hand
On the other

Otherinitialtreesexistforparallel
additiveconstructions("notonly...butalso...")andparallel
contrastiveconstructions("ontheonehand...ontheotherhand").

Otherinitialtreesexistforparallel

andparallel
additive

and
parallel
contrastive
constructions;
notonly...butalso...
"ontheonehand...ontheotherhand."

\begin{diagram}
\node{\textit{study}};
\node{\textit{come}};
\node{\textit{help}};
\node{\textit{phone}};
\end{diagram}

\begin{diagram}
\node{\textit{John}};
\node{\textit{zoo}};
\node{\textit{phone}};
\end{diagram}

\begin{diagram}
\node{\textit{generous}};
\node{\textit{donates}};
\node{\textit{money}};
\node{\textit{charity}};
\node{\textit{you}};
\node{\textit{help}};
\end{diagram}

\begin{diagram}
\node{\textit{John}};
\node{\textit{zoo}};
\node{\textit{phone}};
\end{diagram}

\begin{diagram}
\node{\textit{ontheonehand}};
\node{\textit{ontheotherhand}};
\node{\textit{help}};
\end{diagram}

\begin{diagram}
\node{\textit{John}};
\node{\textit{zoo}};
\node{\textit{phone}};
\end{diagram}

\begin{diagram}
\node{\textit{ontheonehand}};
\node{\textit{ontheotherhand}};
\node{\textit{help}};
\end{diagram}

\begin{diagram}
\node{\textit{John}};
\node{\textit{zoo}};
\node{\textit{phone}};
\end{diagram}

\begin{diagram}
\node{\textit{ontheonehand}};
\node{\textit{ontheotherhand}};
\node{\textit{help}};
\end{diagram}
Derivation of Example 17 using AuxTrees (i)-(ii)

Example Derivations

Our aim is to explain discourse semantics in terms of a product of the same three

Composition Interp: Explanation interp(T2 interp(T1))

Interpretation: How (79a-b) and (79c-d) receive similar interpretations, despite their
different derivations.

Derivation of Example 79a

Composition Interp: Explanation interp(T2 interp(T1))

Interpretation: How (79a-b) and (79c-d) receive similar interpretations, despite their
different derivations.

Derivation of Example 79a

Composition Interp: Explanation interp(T2 interp(T1))

Interpretation: How (79a-b) and (79c-d) receive similar interpretations, despite their
different derivations.
Derivation of Example 79b

\[ \text{Compositional Interpretation: exemplify} \ (\text{interp}(T_2), \lambda X. \ \text{explanation} (\text{interp}(T_1))) \]

Defeasible inference: exemplify \ (\text{interp}(T_2), \lambda X. \ \text{explanation} (\text{interp}(T_1)))

Why defeasible? Because the inference can be contradicted.

You shouldn't trust John. He never returns what he borrows. But that's not why you shouldn't trust him.

VII. Discourse Parsing with DLTAG

Initial two-pass implementation. Each pass using the same chain-based LTAG parser.
LEM is used to parse first these sequence of words and punctuation making up a sentence, and later these sequence of clause units and connectives (realized and unrealized) making up a discourse.

At the sentence level, LEM is designed to produce a single parse according to the XTAGE English grammar [XTAG-Group, 2001], using heuristics to decide which elementary tree to assign to each word:

- Structural information alone is insufficient: Structurally, there is no difference of a peel, "an" vs. an NP constituent.
- Lexical information alone is insufficient: Discourse connectives usually have no difference: otherwise, "otherwise".
- There is difference both lexically and syntactically.

The TreeExtractor extracts a single clause deviation and one elementary tree anchored in a discourse connective. This appears to be a discourse connective.

\[ (i) \text{while she was eating lunch, she saw a dog.} \]

\[ (ii) \text{while she was eating lunch, she saw a dog.} \]

\[ (iii) \text{while she was eating lunch, she saw a dog.} \]

\[ (iv) \text{while she was eating lunch, she saw a dog.} \]

Future work will involve using a statistical version of LEM, based on training on the Penn Treebank.

\[ (i) \text{While she was eating lunch, she saw a dog.} \]

\[ (ii) \text{While she was eating lunch, she saw a dog.} \]

\[ (iii) \text{While she was eating lunch, she saw a dog.} \]

\[ (iv) \text{While she was eating lunch, she saw a dog.} \]

\[ (v) \text{While she was eating lunch, she saw a dog.} \]
While Mary was eating lunch, a dog approached her. It barked and she gave it a sandwich. Then it barked again.

The Tree Mapper maps sentence-level structural descriptions of connective elements to discourse-level structural descriptors. While we assume no lexical ambiguity associated with these sections (but cf. Open Problems), there are structural ambiguities associated with the auxiliary trees for structural connectives (including the empty connective, $\phi$). Where we are not certain whether a connective is structural or anaphoric, we assume the former since it is a less powerful (i.e., more constraining) mechanism.

Example of Tree Extraction

Discourse Input Generation

While we are not certain whether a connective is structural or anaphoric, we assume the former since it is a less powerful (i.e., more constraining) mechanism.

Discourse Input Generation

While we are not certain whether a connective is structural or anaphoric, we assume the former since it is a less powerful (i.e., more constraining) mechanism.
Example 84 has five possible derivations: (ii) (v) (i) (iv) (iii) 

\[ \begin{align*} 
\tau_4 \tau_4 \beta: \tau_3 \\
\tau_2 \tau_3 \beta: \phi \\
\tau_1 \tau_4 \beta: \phi \\
\tau_1 \tau_4 \tau_4 \beta: \phi \\
\tau_1 \tau_2 \beta: \phi \\
\end{align*} \]

Currently, LEM only considers the unique derivation that satisfies the following criteria: 1. Adjunction in initial trees is only allowed at the root node. 2. For all other trees, only the lowest adjunction is allowed.

According to the following derivations, (paraphrased) 1. John and his sister are stubborn. His parents are stubborn. So they [his parents] are always argumenting. 2. John and his sister and his parents are stubborn. So all of them are always argumenting. 3. John and his sister and his parents are stubborn. His sister and parents are stubborn. So they [his sister and parents] are always argumenting.
Mary was eating lunch. A dog approached her and she gave it a sandwich. Then it barked again.

---

Open Problem: Lexical Ambiguity

Just as discourse adverbials like "otherwise" may have non-discourse roles (e.g., as adjectival modifiers), they may also have more than one discourse role. They may also have more than one discourse role.

**Open Problems**

- Integrated Incremental Processing
- Treatment of LINE Semantics Embedding
- Lexical Ambiguity
Thus, when one of these discourse connectives appears, there may be ambiguity as to which connective tree they serve to anchor in the current discourse—parallel initial trees or an auxiliary tree—cf. (87).

Brooklyn College students have an ambivalent attitude toward their school. On the one hand, there is a good deal of self-congratulation at attending a good college. On the other hand, they know they’re saving money by living at home.

N.B. Webber, Knott & Joshi, 1999 follow Knott and Mellish, 1996 in treating connectives and anchors as features structures: Any connective whose features can unify with an anchor can realize that anchor.

Our current DLTAG parser misinterprets this. It does not currently treat "by" as a (structural) discourse connective (easily fixed), and it does not treat the clause (structural) discourse connective (easily fixed), and if it does not treat the clause.

On the other hand, the pilots could play hardball by noting that they are crucial to any sale or restructuring because they can refuse to fly hardball.

The pilots’ ability to refuse is what makes them crucial to any sale or restructuring.

A discourse can occur within direct or indirect speech. The effect of either is that an antecedent that an antecedent can realize that antecedent.

Weber, Knorr & Jost, 1996 follow Knorr and Mellish, 1996 in meaning the influence of their high school. Weber, Knorr & Jost, 1996 follow Knorr and Mellish, 1996 in meaning the influence of their high school. Any connective whose features can realize an antecedent can realize that antecedent.

The pilots could play hardball by noting that they are crucial to any sale or restructuring because they can refuse to fly hardball.

The pilots’ ability to refuse is what makes them crucial to any sale or restructuring.

A discourse can occur within direct or indirect speech. The effect of either is that an antecedent that an antecedent can realize that antecedent.

Weber, Knorr & Jost, 1996 follow Knorr and Mellish, 1996 in meaning the influence of their high school. Weber, Knorr & Jost, 1996 follow Knorr and Mellish, 1996 in meaning the influence of their high school. Any connective whose features can realize an antecedent can realize that antecedent.
While there are still open problems, hope to have shown that the approach of structural variation in discourse is probably less

We can have a common grammatical framework for the two, though the range

Open Problem: Integrated Incremental Parsing

The holy grail of discourse processing is to have a realistic model that is computed

Let 1-2 rival, in parallel with incrementally sentence-discourse processing.

There have been efforts to make anaphoric, rather than a structural connection to

There was both cleavage and merit.

Refusal and non-refusal inference, co-reified, to compute the

Defeasible and non-defeasible inference, co-reified, to compute the

Byron, 2002] and are the immediate reduction of discourse adverbials.

1999], but extended both to deep pronouns as in [Eckert & Strube, 2001; An incremental anaphor resolution mechanism similar to that in [Eckert & Strube, 2001; A LR parser that would simultaneously compute increments in both

...
Bibliography


