Computing Discourse Structure and Discourse Semantics

Bonnie Webber
School of Informatics
University of Edinburgh

carried out with Aravind Joshi (UPenn), Matthew Stone (Rutgers U), and ir Knott (U Otago), and members of the UPenn DL-TAG group (Cassie vell, Kate Forbes, Eleni Miltsakaki and Rashmi Prasad).

Assumption

There are several mechanisms used in conveying (speaker's perspective) and in computing (hearer's perspective) meaning in discourse:

1. Compositional semantics, involving clausal meaning and the presense of explicit discourse connectives (in English – "so", "because", "but", etc.)
2. Inference triggered by adjacency, and the need to relate the adjacent units – e.g.

   I can't wait for lunch to start. *I didn't have time to eat breakfast.*
   I was practicing my talk. *I didn't have time to eat breakfast.*
3. Anaphora, involving units whose sense and/or reference have to be understood (resolved) in the discourse context.

ims of research reported here include:

Understanding the extent to which anaphora is a mechanism for conveying neaning in discourse.

Understanding how meaning conveyed anaphorically interacts with meaning onveyed through compositional semantics and inference.

Developing a way of representing and processing discourse that cknowledges 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. DLTAG
7. Discourse parsing with DLTAG
8. Open problems

I. Background: Abstract Objects

act object (AO) is the term that [Asher, 1993] uses for the possible retations of non-NP constituents, including (a sequence of) clauses (tensed ensed) and verb phrases.

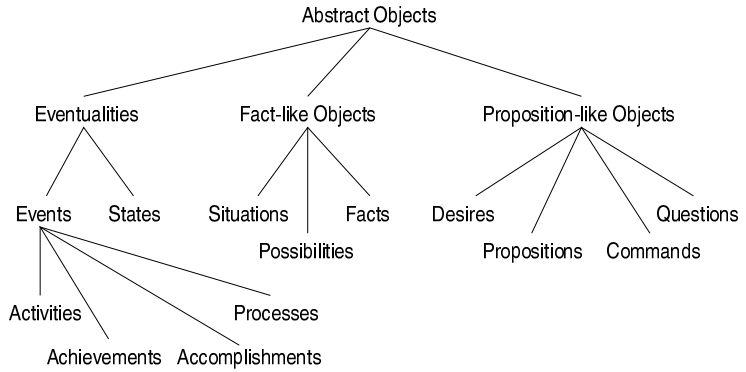
urse-salient AOs often serve as the referents of demonstrative pronouns ber, 1991] – e.g.

ohn took Biology 101 last year. *That* means he can take Biology 102 this ear.

woke up and brushed my teeth. I went downstairs and ate breakfast, and then went to work. *That*'s all I did today.

he instructions say to avoid cutting into the pattern. *That*'s what I am having ouble with.

Classes of AOs [Asher, 1993]



Assumption: Discourse relations are relations between AOs of appropriate sorts. That is their importance to this enterprise.

I. Background: Discourse Relations

ome discourse relations between AOs from adjacent discourse units (DUs) re explicitly marked by discourse connectives.

4) *Although* John is generous, he's hard to find.

5) John broke his arm. *So* he can't cycle to work.

ome discourse relations between AOs are simply inferred, based of the djacency of their DUs.

6) John fell asleep in class. His baby son kept him awake all night.

laim: Additional discourse relations are conveyed by discourse adverbials. Here, the relation is between the AO from the adverbially-marked DU and an AO recovered anaphorically from the previous discourse.

II. Discourse Adverbials and Structural Connectives

Structural connectives: coordinating and subordinating conjunctions, paired conjunctions ("On the one hand...On the other...").

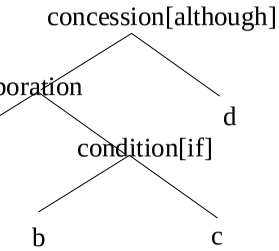
A. Structural connectives allow **stretching** of pred-arg dependencies.

(7) a. *Although* John is generous, he's hard to find.

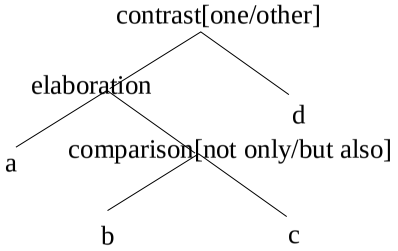
b. *Although* John is generous – if you need some money, you only have to ask him for it – he's very hard to find.

(8) a. *On the one hand*, Sam likes beans. *On the other hand*, they make him ill.

b. *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.



(i)



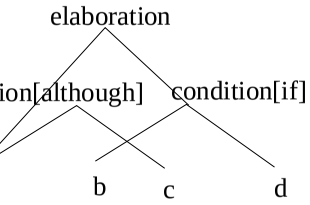
(ii)

ble discourse structures for Ex. 2b and Ex. 3b.

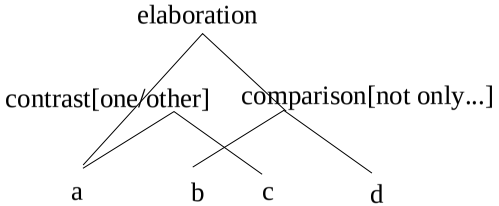
B. Structural Conns do not allow **crossing** of Pred-Arg dependencies.

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

(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.



(i)

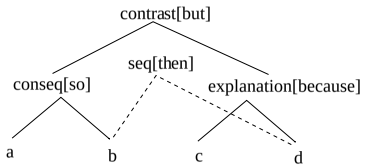


(ii)

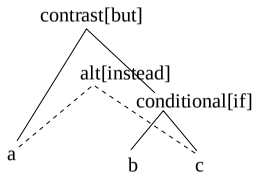
ossible) discourse structures for crossed Ex. 4 and Ex 5.

Discourse adverbials admit crossing dependencies

(11) a. John loves Barolo.
b. So he ordered three cases of the '97.
c. But he had to cancel the order
d. because *then* he discovered he was broke.



a. High heals are fine for going to the theater.
. But wear comfortable shoes
. if *instead* you plan to go to the zoo.



Crossing dependencies are characteristic of anaphors

(13) Every man$_i$ tells every woman$_j$ he$_i$ meets that she$_j$ reminds him$_i$ of his$_i$ mother.

(14) Sue$_i$ drives an Alfa Romeo. She$_i$ drives too fast. Mary$_j$ races her$_i$ on weekends. She$_j$ often beats her$_i$. [Strube, 1998]

Anaphoric dependencies are generally not considered to be structural.

Discourse Adverbials behave like Anaphors: 1

naphoric NPs, discourse adverbials are able to take implicit material as ents:

Stack five blocks on top of one another. Now close your eyes and try nocking {*the tower*, *this tower*} over with your nose.

Do you want an apple? *Otherwise* you can have a pear.

Discourse Adverbials behave like Anaphors: 2

While the left-hand argument of a structural connective is constrained to be a discourse unit on the *right frontier* of the evolving discourse structure, discourse adverbials have a wider range of options, constrained by salience.

(17) If the light is red, stop. *Otherwise* you'll get a ticket.
(*If you do something other than stop, you'll get a ticket.*)

Paraphrased using the conjunction "or": If the light is red, stop, or you'll get a ticket.

(18) If the light is red, stop. *Otherwise* go straight on.
(*If the light is not red, go straight on.*)

Not paraphrasable using "or": *If the light is red, stop, or go straight on.

Discourse Adverbials behave like Anaphors: 3

urse adverbials, like other anaphors, may require semantic reps in which rgs are *bound variables* ranging over discourse entities.

although(p,q) $p \wedge$ *nevertheless(e,q)*

tructure of "donkey sentences" blocks a direct syntactic relation between a un and its antecedent: *discourse semantics* has to provide variables to ate pronouns and *discourse mechanisms* have to treat them as bound.

Every farmer who owns a donkey feeds *it* rutabagas.

Discourse adverbials occur in their own version of "donkey sentences":

(20) Anyone who has developed innovative new software, has *then* had to hire a lawyer to protect his/her interests. (i.e., *after developing innovative new software*)

(21) Many people who have developed innovative new software, have *nevertheless* never gotten very rich. (i.e., *despite having developed innovative new software*)

(22) Every person selling "The Big Issue" might *otherwise* be asking for spare change. (i.e., *if s/he weren't selling "The Big Issue"*)

This suggests that discourse adverbials are accessing discourse entities (in particular, ones of type *abstract object*) rather than signalling structural connection between adjacent clauses.

What kind of Anaphor is a Discourse Adverbial?

mplest discourse anaphors are *coreferential* – definite pronouns and NPs that e one or more discourse entities in the current discourse context.

what more complex is *indirect anaphora* – aka *bridging anaphora* – where aphor (usually a definite NP) denotes a discourse entity *associated with* at one discourse entity in the current discourse context – e.g.,

Myra darted to *a phone* and picked up *the receiver*.

erence and *indirect anaphora* can be uniformly modelled by saying that the urse entity e_α denoted by an anaphoric expression α is either equal to or ated with an existing discourse entity e_r

$$e_\alpha = e_r$$
$$e_\alpha \in \text{assoc}(e_r).$$

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

Consider "other NPs" (Bierner 2001; Modjeska 2001, 2002):

(24) Sue grabbed one phone as Tom darted to *the other phone*.

This is not *associative anaphora*, where $e_\alpha \in \text{assoc}(e_r)$.

Also consider:

(25) Sue lifted the receiver as Tom darted to *the other phone*.

This suggests that an anaphor can convey an idiosyncratic function f_α that may be applied to e_r or assoc(e_r) to yield another discourse entity e_α.

We have called these *lexically-specified* anaphors.

Summary

ow have three types of anaphora, not two:

oreference: $e_\alpha = e_r$

ndirect anaphora: $e_\alpha \in \text{assoc}(e_r)$

exically-specified anaphora: $e_\alpha = f_\alpha(e_i)$ where $e_i = e_r$ or $e_i \in \text{assoc}(e_r)$.

Where do Discourse Adverbials fit in?

Nothing requires that the source of e_r be an NP, or that the anaphor be a pronoun or NP – e.g. The antecedent of "this" or "that" is often a clause or larger unit in the recent discourse.

We take discourse adverbials to be a type of *lexically-specified* anaphor.

Each discourse adverbial has at least one function f_α that maps a discourse entity (of semantic type *abstract object* – AO) in the current discourse context, to a *function* that applies to an AO interpretation of the adverbial's matrix clause.

The result is a *binary relation* that is added to the discourse context and that holds between the two AO entities.

e this, consider:

John didn't have enough money to buy a mango. *Instead*, he bought a guava.

ally, we represent the function contributed by a discourse adverbial α , as a ression

$x \,.\, R_\alpha(x, EV)$

r α=*instead* we have something like:

x . *alternative*(x,EV)

ment x is bound on application to α's matrix clause *S* interpreted as an AO σ:

$\lambda x \,.\, R_\alpha(x, EV)]\sigma \equiv R_\alpha(\sigma, EV)$

So here we would have something like:

$[\lambda$ x . *alternative*(x,EV)] $e_3 \equiv$ *alternative*(e_3,EV)
where $e_3 \equiv \|$John buy a guava$\|$

EV is resolved anaphorically to an AO entity e_i derived from the discourse context either directly or by association, yielding the proposition

$R_\alpha(\sigma, e_i)$

So here we would have something like:

alternative(e_3,e_2), where $e_2 \equiv \|$John buy a mango$\|$

N.B. This formal model is meant to have no implications for how processing takes place.

Example: “Then”

a. John loves Barolo.
. So he ordered three cases of the ’97.
. But he had to cancel the order
. because he *then* found he was broke.

= the anaphoric expression – *then*
$_\alpha$ = the relation name linked with α – *after*
= the matrix clause/sentence containing α – “he [John] found he was broke”
= the interp of *S* as an abstract object – e_4:find(j,e_5), where e_5:broke(j)

solved interpretation of (27d):

λx . R_α(x,*EV*)]σ ≡ [λx . *after*(x,*EV*)]e_4 ≡ *after*(e_4,*EV*)

ved Interpretation: *EV* ≡ e_2:order(j, c_1) from (27b)

fter(e_4,*EV*) → *after*(e_4,e_2)

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.

λ x . *if*(*VE*, x) where *complement*(*VE*,*EV*)
[λx . *if*(*VE*,x)] σ ≡ *if*(*VE*,σ)

f the light is red, stop. *Otherwise* you’ll get a ticket.

= otherwise
$_\alpha$ = *if*
$_{28}$ = you get a ticket
$_{28}$ = e_3, where e_3:get_ticket(you)

solved interpretation:

λx . *if*(VE_{28},x)] e_3 ≡ *if*(VE_{28},e_3), where *complement*(VE_{28},EV_{28})

ved interpretation: *EV* ≡ e_2:stop(you)

(e_4,e_3), where *complement*(e_2,e_4) and e_2:stop(you)
f you do something other than stop, you’ll get a ticket.)

(29) If the light is red, stop. *Otherwise* go straight on.

α = otherwise
R_α = *if*
S_{29} = go straight on
σ_{29} = $e_{3'}$, where $e_{3'}$:go_straight(you)

Unresolved interpretation:

[λx . *if*(VE_{29},x)] $e_{3'}$ ≡ *if*(VE_{29},$e_{3'}$), where *complement*(VE_{29},EV_{29})

Resolved interpretation: *EV* ≡ e_1:red(light1)

if($e_{4'}$, $e_{3'}$), where *complement*(e_1,$e_{4'}$) and e_1:red(light)
(*If the light is not red, go straight on.*)

What does this treatment achieve?

lard account: Both structural connectives and discourse adverbials indicate lation holding between the AOs 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

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

III. Interactions between Discourse Relations

- α = discourse adverbial;
- R_α = the name of the relation associated with α;
- S = the matrix clause/sentence of α;
- σ = the logical form (LF) interpretation of S
- D = discourse unit that is left-adjacent to S, to which a relationship holds either by inference or a structural connective;
- δ = the LF interpretation of D;
- R = the name of the relation between σ and δ triggered by adjacency or given by an explicit structural connective between them.

se 1: σ separately serves as an argument to both R_α and R.

John loves Barolo. So he ordered three cases of the '97. But he had to cancel he order *because* he *then* discovered he was broke.

$_\alpha$ = *after*
= e_4:find(j,e_5), where e_5:broke(j)
λx . *after*(x,*EV*)]$e_4 \equiv$ *after*(e_4, *EV*)

ved interpretation: $EV \equiv e_2$:order(j, c_1)

fter(e_4, e_2)

on R conveyed by "because" is: *explanation*(e_4,e_3), where e_3:cancel(j, o_1)

nary: Two separate propositions added to the discourse:

xplanation(e_4, e_3)
fter(e_4, e_2)

Case 2: $R_\alpha(\sigma, e_i)$ is itself an argument of R.

Look again at Examples 28–29, but here with an *explicit* structural connective.

(31) If the light is red, stop, because *otherwise* you'll get a ticket.

Resolving "otherwise" contributes the relation

> e_6: *if*(e_4,e_3), where *complement*(e_4,e_2), e_2:stop(you) and e_3:get_ticket(you)
> (*If you do something other than stop, you'll get a ticket.*)

This abstract object e_6 serves as one argument to the *explanation* relation contributed by "because", with e_2 being the other.

Together "because" and "otherwise" end up contributing *explanation*(e_2,e_6).

(*Your needing to stop is explained by the fact that if you do something other than stop, you'll get a ticket*).

f the light is red, stop, and/but *otherwise* go straight on.
ving "otherwise" contributes the relation

$_{5'}$:*if*($e_{4'}$, $e_{3'}$), where *complement*($e_{4'}$,e_1), e_1:red(light) and
$_{3'}$:go_straight_on(you)
f the light is something other than red, go straight on.)

is an example of Case 2, what relation R is $e_{6'}$ an argument to?

conjuncts describe (elaborate) alternative continuations of some situation e_0
uced earlier in the discourse – e.g.,
Go a mile to the bridge. If the light is red, stop. Otherwise go straight on.
is the case, then

laboration($e_{6'}$,e_0)
laboration(e_5,e_0)

e_5 is the abstract object interpretation of "If the light is red, stop".

Case 3: R_π is parasitic on R.

The adverbial "for example" is neither anaphoric nor a structural connective. Its relational interpretation (R_π) derives *parasitically* from another relation R through *abstraction*.

Intra-clausal "for example": Abstraction on verb-arg relation.

(34) The box contains, *for example*, a piece of hematite.

(a) contain(box1,hematite1)

(b) *exemplify*(hematite1, {X | contain(box1,X)})

(c) *exemplify*(hematite1, λX . contain(box1,X))

(b) and (c) are alternative ways of representing the abstraction introduced by the adverbial. Here we adopt (c).

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

John just broke his arm. *So, for example*, he can't cycle to work now.

e_1=‖John just broke his arm‖

e_2=‖ he [John] cannot cycle to work now‖

elation R_{so}: *result*(e_2,e_1)

elation R_π: *exemplify*(e_2, λX . *result*(X, e_1))

You shouldn't trust John *because, for example*, he never returns what he orrows.

e_3=‖You should not trust John‖

e_4=‖he [John] never returns what he borrows‖

elation $R_{because}$: *explanation*(e_4,e_3)

elation R_π: *exemplify*(e_4, λX . *explanation*(X, e_3))

(37) Let's go to the Lincoln Memorial. *Then, for example*, we can go to the White House.

e_5=‖we go to the Lincoln Memorial‖

e_6=‖we can go to the White House‖

Relation R_{then}: *after*(e_6,e_5)

Relation R_π: *exemplify*(e_6, λX . *after*(X, e_5))

You shouldn't trust John. *For example,* he never returns what he borrows.

e_7=‖You should not trust John‖

e_8=‖he [John] never returns what he borrows‖

elation $R_{because}$: *explanation*(e_8,e_7)

elation R_π: *exemplify*(e_8, λX . *explanation*(X, e_7))

hat about cases where only *exemplify* seems to hold?

The Herald's accomplishments were notable. *For example,* it consistently eat other papers on disclosures about the Mayor's financial dealings.

e_9=‖The Herald's accomplishments were notable‖

e_{10}=‖it consistently beat . . . ‖

exemplify(e_{10},e_9) versus *elaboration*(e_{10},e_9)

exemplify(e_{10}, λX . *elaboration*(X, e_9))

ne that *exemplify*(e_{10}, λX . *elaboration*(X, e_9)) ⇒ *exemplify*(e_{10},e_9)

Case 4: R_α is a Defeasible Rule that incorporates R.

Occurs with discourse adverbials such as "nevertheless" and "though" that carry the same presupposition as "although" and concessive "while" (Lagerwerf, 1998).

(40) Although Garbo was called the yardstick of beauty, she never married.

What is presupposed (or conventionally implicated) is a defeasible rule (PDR) that fails to hold in the current situation – e.g.

If a woman is called the measure of beauty, she will marry.

Here antecedent and consequent of the PDR derive *structurally* from (40).

With "nevertheless" and "though", the *consequent* of the PDR derives from the adverbial's matrix clause, while the *antecedent* derives anaphorically from the previous discourse.

While John is discussing politics, he is *nevertheless* thinking about fish.

σ=‖john thinks about fish‖

δ=‖john discusses politics‖

while: *during*(σ,δ)

α: *during*(X, E) ∧ E:discuss(Y, politics) > ¬X:think_about(Y, fish)

Normally, whatever one does during the time one is discussing politics, it is not thinking about fish.

he defeasible implication operator from (Asher and Morreau, 1991).

(42) Even *after* John has had three glasses of wine, he is *nevertheless* able to solve difficult algebra problems.

σ=‖john is able to solve difficult algebra problems‖

δ=‖John drinks three glasses of wine‖

R: *after*(σ,δ)

R_α: *after*(X, E) ∧ E:drink(Y,wine) > ¬X:solve(Y, hard_problems)

Normally, whatever one is able to do after one has had three glasses of wine, it is not solving difficult algebra problems.

Summary: Interactions between Discourse Relations

separately serves as an argument to both R_α and R;

$_\alpha(\sigma, e_i)$ is an argument of R;

$_\alpha$ is parasitic on R;

$_\alpha$ is a defeasible rule that incorporates R.

IV. Empirical Studies

If we want to claim that anaphora plays a larger part in discourse semantics than originally though, it is incumbent on us to:

- Identify the range of expressions that are anaphoric;
- Characterise the distribution of the antecedents of each type of anaphor (i.e., what it is in the text that gives them meaning);
- Characterise the relationship between the meaning of their antecedent and their own meaning;
- Develop algorithms for resolving them in Natural Language text, and decision procedures for employing them in Natural Language generation.

mpirical studies have been aimed at:

tudy 1: Verifying the distinction between structural and anaphoric onnectives [Creswell et al., 2002]

tudy 2 (ongoing): Developing feature-based anaphor resolution algorithms or anaphoric connectives [Miltsakaki et al., 2003]

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

Our approach predicts that the arguments to structural and anaphoric connectives have different properties with respect to their location. Is this in fact the case?

Experiment: For each of the following connectives, identify the minimal text unit in the preceding discourse containing the source of its "left-hand" argument:

- Resultatives: *as a result, so, therefore*
- Additives: *also, in addition, moreover*
- Concessives: *nevertheless, yet, whereas*

Data from Brown corpus, WSJ corpus, Switchboard corpus, and 58 transcribed oral histories of online Social Security Administration (SSA) Oral History Archives (http://www.ssa.gov/history/orallist.html).

Initially, 75 tokens of each connective.

Annotation Scheme

Demarcate beginning and end of left and right arguments

Features of left argument

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

Features of right argument

- Combines with: punctuation alone, conjunctions, other adverbial connectives
- Position of connective: initial, medial or clause-final

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

Sample annotation: *therefore*

<ARG REF=27 TYPE=MAIN>Philip Lee was the Chancellor of the campus at San Francisco</ARG><CONN REF=27 COMB=AND POSITION=MED>and he was therefore the person who hired me for the post as Director of the Medical Center.</CONN>

Results of Initial Annotation: Summary

Arguments almost always found locally, in the immediately preceding sequence of sentences.

Wide variety in distribution patterns across connectives

- *So* always takes a sentence or sequence of sentences as its left argument. Apears only in initial position ⇒ structural connective.
- *Nevertheless* often takes XP arguments.
- *Therefore* often takes its left-hand argument from a subordinate clause.

For 3 connectives (*as a result, in addition, nevertheless*), 25 more tokens of each were later added. ⇒ Relative percentages of tags remained stable, indicating that the patterns may be systematic enough to allow automatic detection.

Inter-annotator consistency

A second study was done with 4 annotators and the added 25 tokens of each 3 connectives (*as a result, in addition, nevertheless*).

Also annotated more specifically the location of the left-argument:

- SS = same sentence
- PS = immediately preceding sentence
- PP = multiple sentences immediately preceding the right arg
- NC = sentence(s) non-contiguous with right argument

Some results

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

3-way agreement >50% in all cases

Systematic sources of disagreement – e.g., size of left arg depends on structure assigned to previous context

(43) Lee won the lottery. So he was less stressed about money. *As a result*, his blood pressure went down.

Study 2: Towards Feature-based Anaphor Resolution

With respect to procedures for resolving anaphoric discourse connectives, we are at a much earlier stage than with respect to procedures for resolving pronouns or even definite NPs.

The problem is to identify both

- the antecedent of their anaphoric argument;
- the argument to derive from that antecedent – i.e., the abstract object (AO) that makes sense as an alternative in the current context.

(44) NBC is contemplating getting out of the cartoon business. *Instead*, it may "counter-program" with shows for an audience that is virtually ignored in that time period: adults.
antecedent: getting out of the cartoon business
argument: being/remaining in the cartoon business

Approach here is empirical, involving (for each of a set of examples):

Annotating the minimal text span containing the antecedent of the anaphoric argument and verifying it through inter-annotator agreement;

Developing a set of features that are potentially relevant to identifying the antecedent of the anaphoric argument;

Annotating those features with respect to the agreed-upon antecedents;

Annotating those features with respect to potential competitors (i.e., text spans whose proximity alone suggests their potential as antecedents);

(*optional*: comparing patterns of distribution);

Inducing a decision procedure based on the features (e.g., a decision tree, Naive Bayes, Maximum Entropy, etc.) and computing its accuracy;

Refining the feature set, and repeating from step 3.

"Instead"

The first four steps have been carried out for the discourse adverbial "instead".

Background: "Instead" comes in two forms:

- as a bare adverbial;
- with an "of" PP modifier
  (45) John ate an apple *instead of* a pear.
  (46) John spent the afternoon at the zoo *instead of* at the museum.

With an "of" PP, both args of "instead" derive *structurally*: the first, from the modified phrase (e.g., "an apple"), the second, from the object of the "of" PP (e.g., "a pear").

That second argument is a salient but unchosen *alternative* to the first, with respect to the given predication.

This is basic to the interpretation of "instead" in both its modified and bare forms.

bare adverbial, "instead" gets its second argument *anaphorically*, from the urse context.

ver, not every text span in the previous discourse appears able to provide atives:

a. John ate an apple. #*Instead* he decided to eat a pear.
. John decided to eat a pear. *Instead* he ate an apple.
. John won't eat fruit. *Instead*, he eats only candy bars and potato chips.

o one has established what kind of phrases/clauses suggest alternatives and don't.

resumably, the ability to suggest alternatives is only *one* of several factors nt to identifying the intended antecedent of the anaphoric argument of ad".

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

res were chosen that were observed in examples of "instead" that we had ted earlier:

lausal negation
48) John *couldn't* sleep. Instead, he wrote code. (**Verbal neg**)
49) No one could slept. Instead, everyone wrote code. (**Subj neg**)
50) John ate none of his spinach. Instead, he fed it to his frog. (**Obj neg**)

resence of a monotone-decreasing quantifier (**MDQ**)
51) *Few* students like to do homework. Instead, they would rather party.
52) Students *seldom* sleep in class. Instead, they take notes assiduously.

resence of a modal auxiliary (**Modal**)
53) You *should* exercise more. Instead you sit like a couch potato.

- whether the antecedent is embedded in a higher clause (**Embed**)
  (54) *John wanted* to eat a pear. *Instead*, he ate an apple.
  (55) *Chrysler officials resisted* cutting output. *Instead*, they slapped $1000 cash rebates on vehicles.
  (56) *Paine Webber considered* recommending specific stocks. *Instead*, it just urged its clients to stay in the market.

Step 3: Annotate Features of Agreed Upon Antecedents

Features	YES (of 97)	NO (of 97)
Verbal neg.	37 (38%)	60 (62%)
Subj. neg.	5 (5%)	92 (95%)
Obj. neg.	10 (10%)	82 (85%)
MDQ	1 (1%)	96 (99%)
Modal	12 (12%)	85 (88%)
Condit.	1 (1%)	96 (99%)
Embed.	57 (59%)	40 (41%)

Antecedents could display $\geq$0 features – e.g., both **Subj neg** and **Modal**.

atures of potentially competing antecedents display the same pattern?

Step 4: Annotate potentially competing antecedents (PCAs)

PCAs are any finite or non-finite clause contained in the sentence that contains the antecedent of "instead" or that intervenes between the antecedent and the sentence containing "instead".

(Sentence = the clause containing the main verb and all its associated finite or non-finite clauses, including relative and adverbial clauses. Also here, a coordinated VP whose subject is omitted is considered a sentence.)

For the 97 tokens of "instead" on which annotators agreed, there were 169 potentially competing antecedents (PCAs).

Distribution of Features of PCAs

Features	YES (of 169)	No (of 169)
Verbal neg.	21 (12%)	148 (88%)
Subj. neg.	8 (5%)	161 (95%)
Obj. neg.	6 (4%)	139 (82%)
MDQ	0 (0%)	169 (100%)
Modal	17 (10%)	152 (90%)
Condit.	0 (0%)	169 (100%)
Emb.	14 (8%)	155 (91%)

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 ($\approx$ 53%) versus 35/169 times ($\approx$ 20%).
2. The antecedent of the anaphoric argument of "instead" is over 7 times more frequently embedded in a higher verb than is a potentially competing antecedent – 57/97 times ($\approx$ 59%) vs 14/169 times ($\approx$ 8%).

Other differences between antecedents and PCAs

he case of antecedents, the main verb of the embedding clause included *abandon, doubt, expect, tell, say, concede, want, be appropriate*, etc.

embedded PCAs were also dominated by these main verbs – e.g., *say, be tain, doubt*, etc. But they were also dominated by factive verbs like *know*.

e verbs presuppose the truth of their embedded clause:

John knows that Fred eats meat.

embedded clauses do not appear to give rise to alternatives:

John knows that Fred eats meat. *Instead Fred likes tofu.

y be useful to annotate identity of embedding verb, in order to exclude ilities.

2. Certain verbs themselves appear to suggest alternatives, independent of explicit negation, or a MDQ, or a modal or clausal embedding:

(59) John *doubted* Mary's resolve. *Instead*, he thought she would give up as soon as they left.

(60) NBC is contemplating *getting out of* the cartoon business. *Instead*, it may "counter-program" with shows for an audience that is virtually ignored in that time period: adults.

(61) Investors have *lost* their enthusiasm for the stock market. *Instead*, they are buying government bonds.

(62) But respectability still *eludes* Italy's politics. *Instead*, it has the phenomenon of Mr. Belusconi.

∴ May be useful to annotate main verbs, in order to admit additional possibilities.

her lexico-syntactic elements may trigger alternatives:

The tension was evident on Wednesday evening during Mr. Nixon's final anquet toast, normally an opportunity for reciting platitudes about eternal riendship. *Instead*, Mr. Nixon reminded his host, Chinese President Yang hangkun, that Americans haven't forgiven China's leaders for the military ssault of June 3-4 that killed hundreds, and perhaps thousands, of emonstrators.

ers "normally" and "opportunity" (either is sufficient):

Normally, we eat pasta on Tuesday. *Instead*, tonight we're having fish.

John had the opportunity to buy a cheap used car. *Instead*, he bought a cooter.

ed to broaden the range of features being considered: It does not correspond previously defined set!

Other features we should annotate:

- Location of the antecedent of the anaphoric argument of "instead" (cf. SS, PS, PP, NC), or alternatively, a measure of its *distance* from "instead", measured in e.g. clausal constituents.
- Syntactic type of the antecedent (cf. Main, Subordinate, XP, ...)

After additional annotaitons, we should try to

- Induce a decision procedure based on the set of features, and assess its accuracy and its dependency on those features
- Reassess the situation.

V. Comparison with other approaches

proaches to discourse structure and semantics address the obvious fact that urse conveys more than the meaning of its component clauses/sentences. In l approaches, such as

hetorical Structure Theory [Mann & Thompson, 1988]

inguistic Discourse Model [Polanyi, 1988; Polanyi & van den Berg, 1996]

tructured DRT [Lascarides and Asher, 1991, 1993]

st part of the additional meaning is associated with some type of *discourse* or *ical* relation holding between the adjacent units.

er approaches, additional meaning is associated with *speaker intentions*, but ntentions are grounded in syntax and semantics has not been of primary rn.

Computational approaches to discourse underpinned by the idea of meaning added by discourse/rhetorical relations include:

- Interpretation as Abduction [Hobbs, Stickel, et al, 1993]
- Implementations of RST
  - Text analysis: [Marcu, 1999; Marcu & Echihabi, 2002; Carlson et al, 2003]
  - Text generation: [Hovy, 1988], [Moore, 1995], [Scott & deSouza, 1990]

In trying to implement these approaches to discourse, practitioners have found it difficult to separate discourse and clause-level syntax. This has led to a tension between theory and practice.

xample, RST makes the following assumptions:

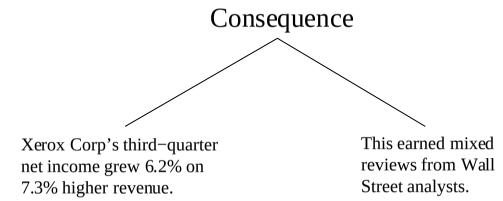
he terminal nodes of discourse structure are text spans that represent the inimal units of discourse.

chemas (abstract patterns) define how *adjacent, non-overlapping text spans* elate to one another. Non-terminal nodes come from schema applications.

Most schemas specify involve a *rhetorical relation* holding between $\geq$2 *on-overlapping, adjacent text spans.*

One schema application spans the entire text. No part of the text is uncovered.

ifficult to separate discourse and clause-level syntax because the same lying set of *rhetorical relations* can be expressed in a discourse, or in a sentence or even a single clause.



[Carlson et al., 2003]

(66) Xerox Corporation's third-quarter net income grew 6.2% on 7.3% higher revenue. This earned mixed reviews from Wall Street analysts.

(67) Xerox Corporation's third-quarter net income grew 6.2% on 7.3% higher revenue, which earned mixed reviews from Wall Street analysts.

(68) Xerox Corporation's third-quarter net income grew 6.2% on 7.3% higher revenue, earning mixed reviews from Wall Street analysts.

(69) The 6.2% growth of Xerox Corporation's third-quarter net income on 7.3% higher revenue earned mixed reviews from Wall Street analysts.

tation guidelines [Marcu, 1999; Carlson et al, 2003]:

Nominalised clauses are not treated as minimal discourse units. So (69) is nnotated as a single discourse unit, while (66)–(68), as two units.

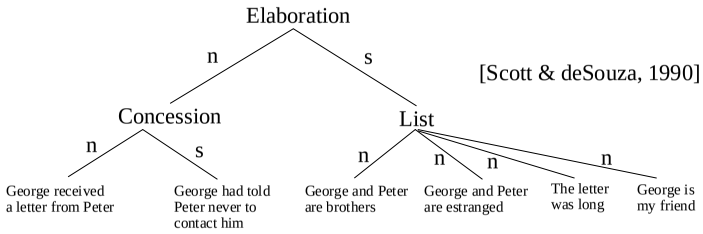
mbedded discourse units are introduced to allow relative clauses, nominal ost-modifiers, and verb complements to be treated as minimal discourse nits.

70) [The results underscore Sears' difficulties] [in implementing the "everyday low pricing" strategy].

71) [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].

72) [There is just so much going on] [that it's difficult to pick just one factor] [that's driving the market].

esults violate RST assumptions 2 and 3.



[Scott & deSouza, 1990]

[Scott & deSouza, 1990] note that all the following paraphrases capture this rhetorical structure – some, more fluently than others.

(73) My friend George received a long letter from his estranged brother Peter, even though he had told Peter never to contact him.

(74) My friend George received a long letter from his brother Peter, even though he had told Peter, from whom he is estranged, never to contact him.

(75) My friend George received a letter from his estranged brother Peter, even though he had told Peter never to contact him. The letter was long.

My friend George received a long letter from Peter, who is his brother and rom whom he is estranged, even though he had told Peter never to contact im.

paraphrases also violate assumption 2 and 3 of RST, including *crossing*, and cott & deSouza) treat them as arising from a post-process of *aggregation*.

eory, RST is attempting to set out a compositional discourse semantics, based rticular assumptions about discourse syntax. They are assumptions that tioners find difficult to adhere to, both in analysis and in generation.

We believe that it is simpler to recognise that

- Some, possibly all, the same range of meaning can be carried within as across sentences.
- Clause-level syntax and semantics and discourse-level syntax and semantics use the same mechanisms for carrying meaning: compositional semantics, anaphora, and inference.
- Clause-level syntax $\neq$ discourse-level syntax. In many languages, including English, clause-level syntax allows much more complex variation than discourse-level syntax.

The object should be to understand how complex meaning is conveyed within the clause/sentence, and then, how those meanings can be combined together. One can only go so far with a theory of discourse structure and relations with a hollow core of meaning at the sentence/clause level.

VI. DLTAG: A Lexicalised Grammar for Discourse

o incorporate this treatment of discourse relations into a computational
ach to discourse semantics and discourse structure?

ncorporate into a sentence-level grammar since, syntactically, both discourse
dverbials and structural connectives fall within the sentence.
Downside: Sentence-level grammars don't provide for forming the meaning
f multi-clausal units that cross sentence-level punctuation.

ake a different approach to discourse-level syntax and semantics than to
entence-level syntax and semantics, combining (for example) a Definite
lause Grammar with Rhetorical Structure Theory.
Downside: This requires discourse semantics reaching deeper and deeper into
entence-level syntax and semantics to handle relations between main and
mbedded clauses, and between embedded clauses themselves.

- Extend a sentence-level grammar and its associated semantic mechanisms to discourse, using either
  - a Phrase-structure Grammar, cf. [Polanyi and van der Berg, 1996]
  - a Lexicalized Grammar that pairs each lexical item with the syntactic structures that it heads (and also, possibly, relevant semantic, discourse, stylistic and pragmatic information), cf. Combinatory Categorial Grammar (CCG), Dependency Grammar, Lexicalized Tree Adjoining Grammar (LTAG)

  Downside: ????

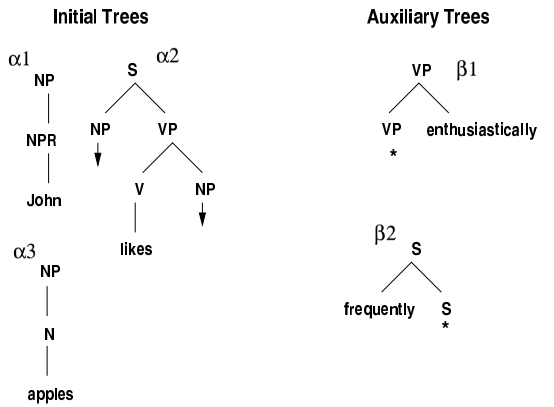
Lexicalized TAG

lexical anchor has ≥ 1 associated tree structures (*elementary trees*, one for
minimal syntactic construct in which it can appear.

are two kinds of elementary trees:

itial trees that reflect basic function-argument dependencies;

uxiliary trees that introduce recursion and also serve to extend or modify a
escription begun in the previous discourse.



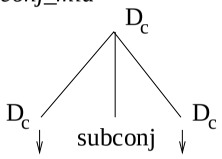
$\downarrow$ = site available for substitution

$*$ = the unique foot node in an auxiliary tree, used for adjoining.

Casting Discourse Syntax in LTAG: Initial Trees

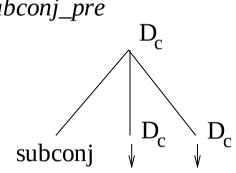
Unlike the wide variety of trees needed at the clause level, extending LTAG course has required only a few elementary tree structures, possibly because urse doesn't exploit structural variation in ways that clause-level syntax does.

α: *subconj_mid*



(a)

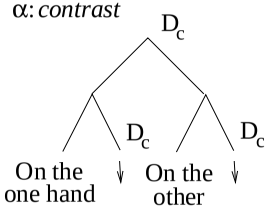
α: *subconj_pre*



(b)

amily of initial trees for main/subordinate constructions, all of which share me predicate-argument dependencies. "subconj" stands for the particular dinate conjunction that anchors the tree.

α: *contrast*



An initial tree for a parallel contrastive construction anchored by "on the one hand" and "on the other hand". Note the *pair* of anchors.

Other initial trees exist for parallel disjunctive constructions ("either . . . or . . . ") and parallel additive constructions ("not only . . . but also . . . ").

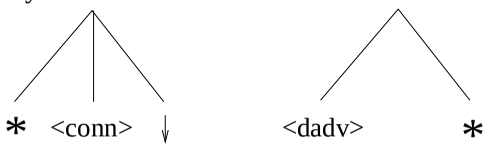
Casting Discourse Syntax in LTAG: Auxiliary Trees

iary tree (i) is anchored by an explicitly realised connective (77) or an ised connective (78) – i.e., simple adjacency:

f John is generous *and* he donates money to every charity, why don't you sk him for help.

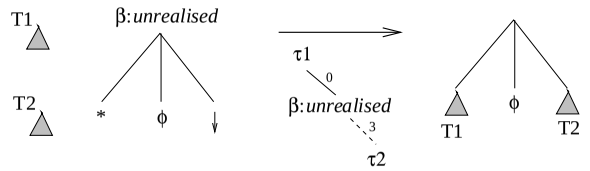
a. John went to the zoo. b. He took his cell phone with him.

iary tree (ii) is anchored by a discourse adverbial whose first argument is ed anaphorically.



(i) (ii)

Derivation of Example 78 using AuxTree (i)

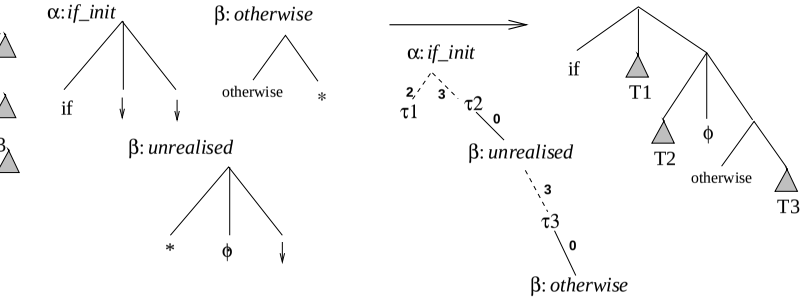


- T1 – the LTAG tree for clause 78a
- T2 – the LTAG tree for clause 78b
- β:*unrealised* – the instance of auxiliary tree (i) connecting adjacent clauses without an overt connective.

The derivation records that the foot node of β:*unrealised* is adjoined to the root of T1 (solid line) and its substitution site filled by T2 (dashed line).

Derivation of Example 17 using AuxTrees (i)-(ii)

17) If the light is red, stop. *Otherwise* you'll get a ticket.



T1 – LTAG tree for "the light is red"

T2 – LTAG tree for "stop"

T3 – LTAG tree for "you'll get a ticket"

Example Derivations

Our aim is to explain discourse semantics in terms of a product of the same three interpretive mechanisms that operate within clause-level semantics:

- compositional rules on syntactic structure (here, discourse structure)
- anaphor resolution
- inference triggered by adjacency and structural connection.

The example derivations here are meant to support the first mechanism. We gloss over the issue of inference for now. The work presented next in Part VII aims to address resolving anaphoric discourse adverbials and the interpretation that this contributes.

a. You shouldn't trust John because he never returns what he borrows.

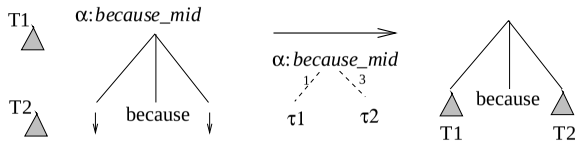
b. You shouldn't trust John. He never returns what he borrows.

c. You shouldn't trust John because, for example, he never returns what he borrows.

d. You shouldn't trust John. For example, he never returns what he borrows.

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

Derivation of Example 79a

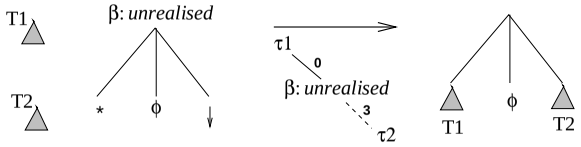


$T1$ stands for the LTAG parse tree for "you shouldn't trust John", τ1, its derivation tree, and interp($T1$), the entity associated with its AO interpretation.

$T2$ stands for the LTAG parse tree for "he never returns what he borrows", τ2, its derivation tree, and interp($T2$), the entity associated with its AO interpretation.

Compositional Interp: *explanation*(interp($T2$),interp($T1$))

Derivation of Example 79b

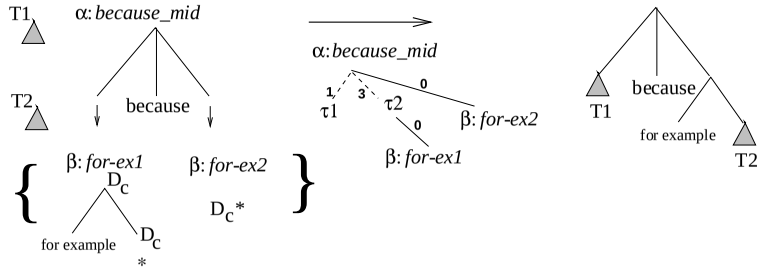


positional interpretation: *elaboration*(interp($T2$),interp($T1$))
sible inference ⇒ *explanation*(interp($T2$),interp($T1$))

defeasible? Because the inference can be contradicted.

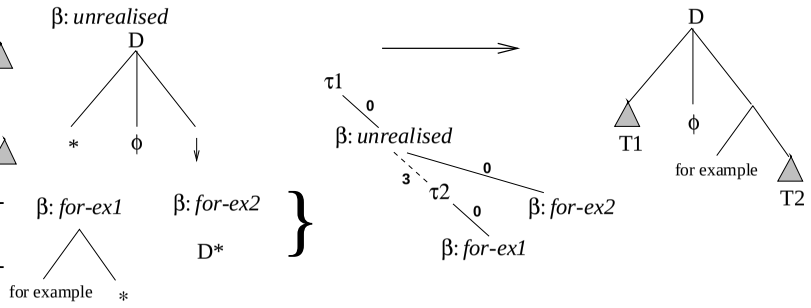
You shouldn't trust John. He never returns what he borrows. But that's not why you shouldn't trust him.

Derivation of Example 79c



Compositional Interp: *exemplify*(interp($T2$), λX . *explanation*(X, interp($T1$)))

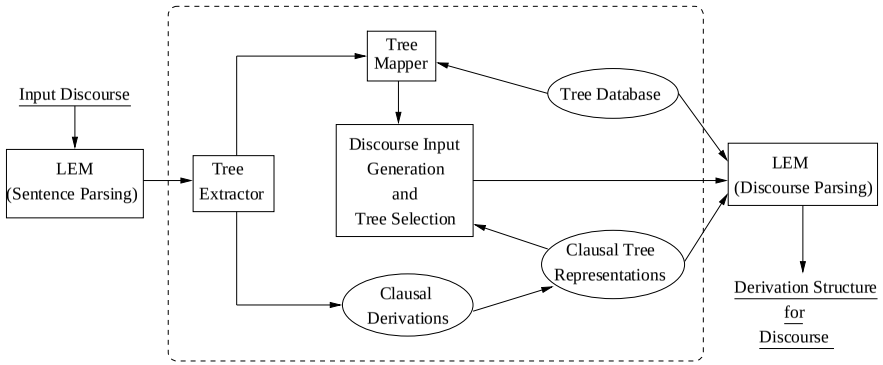
Derivation of Example 79d



positional interp: *exemplify*(interp($T2$), λX . *elaboration*(X, interp($T1$)))
sible inference ⇒ *exemplify*(interp($T2$), λX . *explanation*(X, interp($T1$)))

VII. Discourse Parsing with DLTAG [Forbes, 2001]

Initial two-pass implementation, each pass using the same chart-based left-corner LTAG parser, LEM [Sarkar, 2000]



LEM **Parser** [**Sarkar, 2000**]

s used to parse first the sequence of words and punctuation making up a nce, and later the sequence of clause units and connectives (realised and lised) making up a discourse.

e sentence-level, LEM is designed to produce a single parse according to the G English grammar [XTAG-group, 2001], using heuristics to

ecide which elementary tree to assign to each word;

ick the lowest attachment point between these trees.

e work will involve using a statistical version of LEM, based on training on enn TreeBank.

Tree Extractor

TE extracts from each sentence derivation (not derived tree) individual clausal derivations and any elementary trees anchored in discourse connectives.

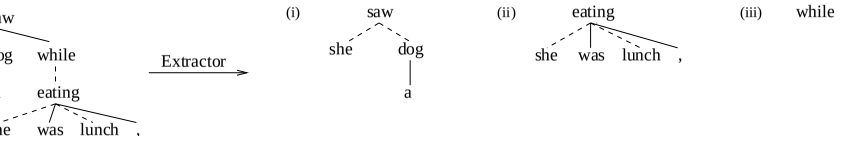
Pass 1 involves a top-down traversal of the derivation tree to identify the discourse connectives.

This requires both *lexical* and *structural* information.

- Lexical information alone is insufficient: Discourse connectives usually have multiple functions (e.g., "instead" as an NP post-modifier – "an apple instead of a pear"; "and" as an NP conjunction).
- Structural information alone is insufficient: Structurally, there is no difference between clausal adverbials like "frequently" and discourse adverbials like "otherwise".

2 detaches clausal derivations at their substitution and/or adjunction nodes.

While she was eating lunch, she saw a dog.

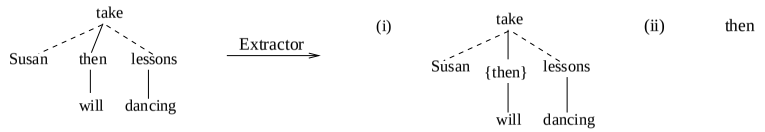


the **Tree Extractor** extracts two clausal derivations and one elementary tree red in a discourse connective.

With clause-medial discourse connectives, as in

(82) Susan will *then* take dancing lessons.

the **Tree Extractor** makes a *copy* of the derivation and replaces the discourse connective with an *index*, to retain its clause-internal position. This appears to be relevant to Information Structure [Steedman, 2001].



Here **TE** extracts a single clausal derivation and one elementary tree anchored in a discourse connective.

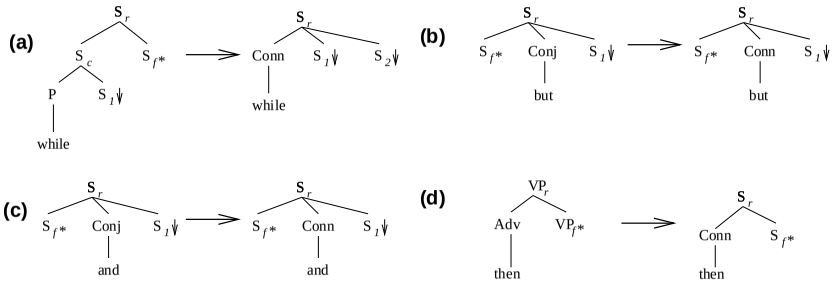
Example of Tree Extraction

While Mary was eating lunch, a dog approached her. It barked *and* she gave a sandwich. *Then* it barked again.

lementary tree: *while*
lausal derivation: Mary was eating lunch
lausal derivation: a dog approached her
lausal derivation: it barked
lementary tree: *and*
lausal derivation: she gave it a sandwich
lementary tree: *then*
lausal derivation: {then} it barked again.

Tree Mapper

The **Tree Mapper** maps *sentence-level* structural descriptors of connective elementary trees to their *discourse-level* structural descriptors.



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 constrained) mechanism.

Discourse Input Generation

to LEM is a sequence of lexicalised trees. The role of **DIG** is, essentially, to ce such a sequence at the discourse level.

this, **DIG** first converts clausal derivations into elementary tree entations (*clausal units*)

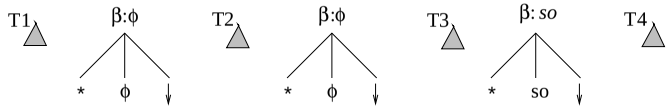
clausal units, along with the discourse connectives, make up the sequence of lised trees.

Where there is no structural connective between clausal units, **DIG** inserts an ary tree with an empty lexical anchor into the input sequence.

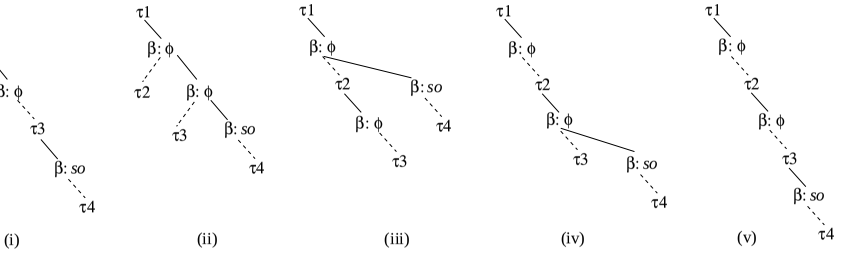
Ambiguity in Discourse-level Parsing

While we assume no *lexical ambiguity* associated with tree selection (but cf. *Open Problems*), there are *attachment ambiguities* associated with the auxiliary trees for structural connectives (including the empty connective, ϕ):

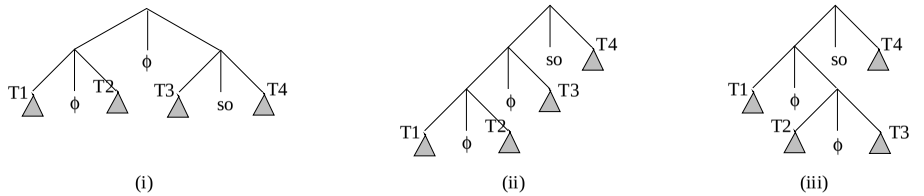
(84) John is stubborn. (T1)
His sister is stubborn. (T2)
His parents are stubborn. (T3)
So they are always arguing. (T4)



ple 84 has five possible derivations:



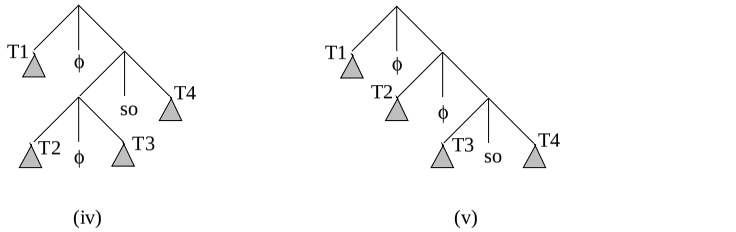
sponding to the following derived structures and (paraphrased)
retations:



i. John and his sister are stubborn. His parents are stubborn. So they [his parents] are always arguing.

ii. John and his sister and his parents are stubborn. So all of them are always arguing.

iii. John and his sister and his parents are stubborn. So all of them are always arguing.



hn is stubborn. His sister and his parents are stubborn. So they [his sister and is parents] are always arguing.

hn and his sister and his parents are stubborn. So all of them are always rguing.

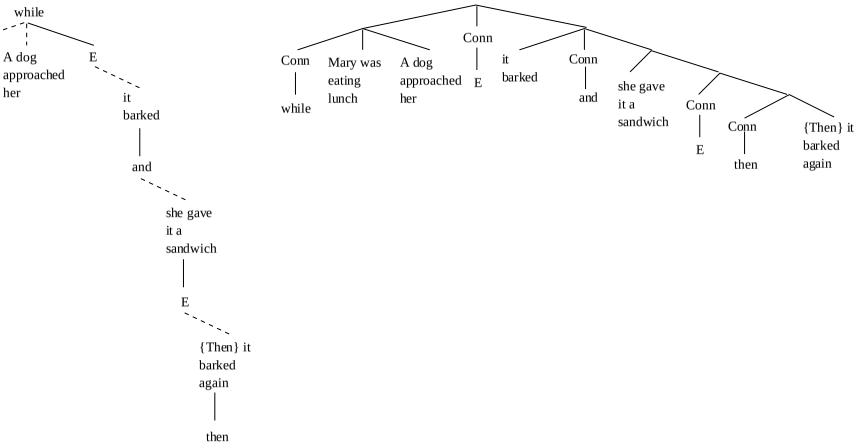
Currently, LEM only considers the unique derivation that satisfies the following criteria:

1. Adjunction in initial trees is only allowed at the root node.
2. For all other trees, only the lowest adjunction is allowed.

That means LEM only associates derivation (v) and derived tree (v) with Example 84, rather than any of the others.

This derivation tree is meant to serve as the input to discourse semantic interpretation [Forbes, 2003], similar to LTAG-based compositional semantics at the sentence-level [Joshi & Vijay-Shankar, 2001; Kallmeyer, Joshi & Romero, 2003].

Derived and Derivation Trees produced for Example 83



VIII. Open Problems

- Lexical ambiguity
- Treatment of true semantic embedding
- Integrated incremental processing

Open Problem: Lexical Ambiguity

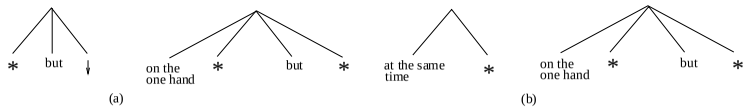
s discourse adverbials like “otherwise” may have non-discourse roles as well as adjectival modifiers), they may also have more than one discourse role.

der: Of the Brown Corpus’s eleven instances of the contrastive construction e initial anchor is “on the one hand”, four have something other than “on the (hand)” as their medial anchor:

On the one hand, the Public Health Service declared as recently as October 6 that present radiation levels resulting from the Soviet shots “do not warrant ndue public concern” . . . *But* the PHS conceded . . . (*cb21*)

(86) Brooklyn College students have an ambivalent attitude toward their school. *On the one hand*, there is a sense of not having moved beyond the ambiance of their high school. This is particularly acute for those who attended Midwood High School directly across the street from Brooklyn College. . . . *At the same time*, there is a good deal of self-congratulation at attending a good college . . . (*cf25*)

This suggests that both “but” and “at the same time” can serve as one anchor for the initial tree for parallel contrastive constructions (cf. Section VI). But “but” is also a structural connective, anchoring auxiliary tree (a), while “at the same time” is also an anaphoric connective, anchoring auxiliary tree (b).



when one of these discourse connectives appears, there may be ambiguity as ich connective tree they serve to anchor in the current discourse – a parallel tree or an auxiliary tree – cf.

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 ood college. *At the same time*, they know they're saving money by living at ome. *On the other hand*, there is a sense of not having moved beyond the mbiance of their high school.

Webber, Knott & Joshi, 1999] follows [Knott and Mellish, 1996] in treating ctives and anchors as features structures: Any connective whose features can with an anchor can realise that anchor.

Open Problem: Embedded Discourse

A discourse can occur within direct or indirect speech or a thought or belief or other attitude – for example,

(88) The pilots could play hardball by noting [that they are crucial to any sale or restructuring because they can refuse to fly the airplanes].

The pilots' ability to refuse is what makes them crucial to airline sale or restructuring.

Our current DLTAG parser misanalyzes this: It does not currently treat "by" as a (structural) discourse connective (easily fixed), and it does not treat the clausal object of "note" at something to be parsed at the discourse level.

esult is that "because" is taken to relate the clauses

he pilots could play hardball by noting that they are crucial to any sale or estructuring

ey can refuse to fly the airplanes

s, the pilots' ability to refuse is what allows them to play hardball.

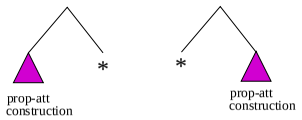
ple extraction gives the wrong results.

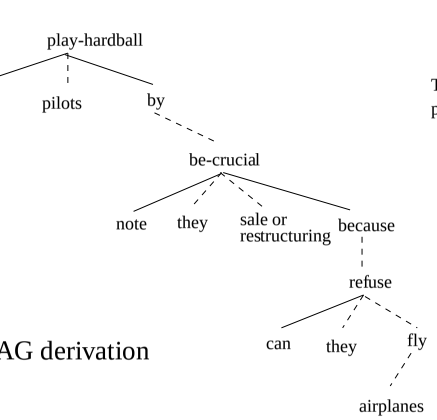
ndle (88) correctly, we need to treat "by" as a discourse connective, and to ish a recursive procedure that could apply to discourse found at any level n syntactic structure, including relative clauses.

Embedded Discourse: Possible Solution

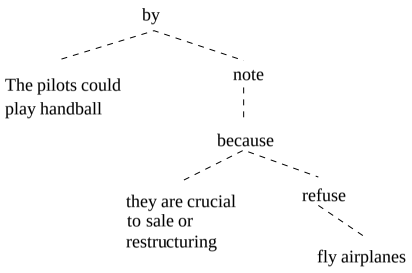
We can introduce *inital trees* at the discourse-level for propositional attitude constructions. (These are *auxiliary trees* at the clause-level.)

We can then restrict extraction of embedded connectives to **below** the level of the propositional attitude.





AG derivation



DLTAG derivation

Open Problem: Integrated Incremental Parsing

The holy grail of discourse parsing is a realistic model that is computed left-to-right, in parallel with incremental sentence-level parsing.

What would an integrated incremental method of sentence-discourse processing require? Possibly at minimum:

- A LR parser that would simultaneously compute increments to both sentence-level syntax/semantics, and discourse-level syntax/semantics.
- An incremental anaphor resolution mechanism, similar to that in [Strube, 1998], but extended both to deictic pronouns, as in [Eckert & Strube, 2001; Byron, 2002], and to the anaphoric argument of discourse adverbials.
- Defeasible and non-defeasible inference "on demand", to compute the relations underlying adjacency, both *intra-clausally* (for noun-noun compounds) and *inter-clausally*, for unmarked adjacent sentences.

Conclusion

Discourse adverbials make an anaphoric, rather than a structural connection to he previous discourse.

Viewing discourse adverbials as anaphoric allows one to investigate the ways n which the semantic contribution of the adverbial interacts with that of a tructural connective or of an adjacency inference.

mpirical studies have provided evidence for differences between discourse dverbials and structural connectives.

mpirical studies can also be used to provide a basis for resolution procedures or discourse adverbials. http://www.cis.upenn.edu/ dltag/

rom the perspective of semantics, there is little difference between the echanisms by which clauses get their interpretation and those by which iscourse does so: computational semantics, inference, anaphora.

- We can have a common grammatical framework for the two, though the range of structural variation in discourse is probably less.
- While there are still open problems, hope to have shown that the approach here has both elegance and merit.

Bibliography

olas Asher (1993). *Reference to Abstract Objects*. Dordrecht: Kluwer.

na Byron (2002). Resolving Pronominal Reference to Abstract Entities. *Proc. 40th l Meeting, Association for Computational Linguistics*, Philadelphia PA, pp. 80–87.

Carlson, Daniel Marcu and Mary Ellen Okurowski (2003). Building a rse-Tagged Corpus in the Framework of Rhetorical Structure Theory. In *Current ions in Discourse and Dialogue*, J. van Kuppevelt and R. Smith (eds), Kluwer.

andra Creswell, Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi onnie Webber (2002). The Discourse Anaphoric Properties of Connectives. *edings of DAARC-2002*, Lisbon.

am Eckert and Michael Strube (2001). Dialogue Acts, Synchronising Units and ora Resolution. *Journal of Semantics* 17, pp. 51–89.

• Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi and Bonnie Webber (2001). D-LTAG System: Discourse Parsing with a Lexicalized Tree Adjoining Grammar. *ESSLLI'2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. Helsinki, Finland. To appear in *Journal of Logic, Language and Information*, 2003.

• Katherine Forbes (2003). *Discourse Semantics of S-Modifying Adverbials.* PhD dissertation, Department of Linguistics, University of Pennsylvania.

• Jerry Hobbs, Mark Stickel, et al (1993). Interpretation as Abduction. *Artificial Intelligence* 63(1-2), pp. 69-142.

• Eduard Hovy (1988). Planning Coherent Multisentential Text. *Proc. 26th Annual Meeting, Association for Computational Linguistics*, Buffalo NY, pp. 163-169.

•Aravind Joshi and K. Vijay-Shankar (2001). Compositional Semantics with Lexicalised Tree-Adjoining Grammar (LTAG). In H. Bunt, R. Muskens and E. Thijsse (eds.), *Computing Meaning* volume 2, Kluwer, pp. 147-163.

a Kallmeyer, Aravind Joshi and Maribel Romero (2003). Flexible Composition in Quantifier Scope and Inverse Linking. *Proceedings of the 2003 International hop on Computational Semantics* IWCS-5, Tilburg, pp. 179-194.

air Knott and Chris Mellish (1996). A Feature-based Account of the Relations led by Sentence and Clause Connectives. *Language and Speech* 39(2-3), pp. 83.

Lascarides and Nicholas Asher (1991). Discourse Relations and Defeasible edge. *Proceedings of the 29th Annual Meeting of the Association for Computational stics*, Berkeley CA, pp. 55-63.

Lascarides and Nicholas Asher (1993). Temporal Interpretation, Discourse ons and Commonsense Entailment. *Linguistics and Philosophy* 16(5), pp. 437-493.

iam Mann & Susan Thompson (1988). Rhetorical Structure Theory: Toward a onal Theory of Text Organization. *Text* 8(3), pp. 243-281.

el Marcu (1999). *Instructions for Manually Annotating the Discourse Structure of* Available from http://www.isi.edu/~marcu.

• Daniel Marcu & Abdessamad Echihabi (2002). An Unsupervised Approach to Recognizing Discourse Relations. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Philadelphia PA, pp. 368-375.

• Eleni Miltsakaki, Cassandre Creswell, Katherine Forbes, Aravind Joshi and Bonnie Webber (2003). Anaphoric arguments of discourse connectives: Semantic properties of antecedents versus non-antecedents. *EACL Workshop on the Computational Treatment of Anaphora*, Budapest.

• Johanna Moore (1995). *Participating in Explanatory Dialogues*, Cambridge MA: MIT Press.

• Livia Polanyi (1988). A Formal Model of the Structure of Discourse. *Journal of Pragmatics* 12, pp. 601-638.

• Livia Polanyi and Martin van den Berg (1996). Discourse Structure and Discourse Interpretation. *Proceedings of the Tenth Amsterdam Colloquium*, University of Amsterdam, pp. 113-131.

op Sarkar (2000). Practical Experiments in Parsing Using Tree-Adjoining Grammars. *edings of the 5th TAG+ Workshop*, pp. 193-198.

op Sarkar (2001). Applying Co-training Methods to Statistical Parsing. *Proceedings 2nd NAACL*. Pittsburgh PA.

a Scott & Clarisse Sieckenius de Souza (1990). Getting the Message Across in ased Text Generation. In *Current Research in Natural Language Generation*, R. Dale eds), London: Academic Press.

k Steedman (2000). Information Structure and the Syntax-Phonology Interface. *stic Inquiry* 34, pp. 649-689.

hael Strube (1998). Never Look Back: An alternative to centering. *Proceedings, NG/ACL'98*, Montreal, pp. 1251–1257.

hie Webber (1991). Structure and Ostension in the Interpretation of Discourse Deixis. *l Language and Cognitive Processes* 6(2), pp. 107-135.

• Bonnie Webber, Alistair Knott and Aravind Joshi (1999). Multiple Discourse Connectives in a Lexicalized Grammar for Discourse. *Proc. Third International Workshop on Computational Semantics*, Tilburg, The Netherlands, pp. 309–325.

• Bonnie Webber, Matthew Stone, Aravind Joshi, and Alistair Knott (2003). Anaphora and Discourse Structure. *Computational Linguistics*, to appear.

• XTAG-Group (2001). A Lexicalized Tree-Adjoining Grammar for English. University of Pennsylvania, Technical Report IRCS-01-03.
ftp://ftp.cis.upenn.edu/pub/ircs/technical-reports/01-03.