Computational Discourse

Yejin Choi
University of Washington

Many slides from Aravind Joshi, Rashmi Prasad, Bonnie Weber,
Regina Barzilay, Edward Loper, Luke Zettlemoyer

Plan

1. Textual Coherence
2. Rhetorical Structure Theory (RST)
3. Penn Discourse Tree Bank (PDTB)
4. Coreference Resolution / Entity-grid Model

Textual Coherence

- John hid Bill’s car keys. He was drunk.
- John hid Bill’s car keys. He likes spinach.

Textual Coherence

- John went to his favorite music store to buy a piano.
- He had frequented the store for many years.
- He was excited that he could finally buy a piano.
- He arrived just as the store was closing for the day.

- John went to his favorite music store to buy a piano.
- It was a store John had frequented for many years.
- He was excited that he could finally buy a piano.
- It was closing just as John arrived.

Why Model Coherence

“How much wood could a woodchuck chuck if a woodchuck would chuck wood.”

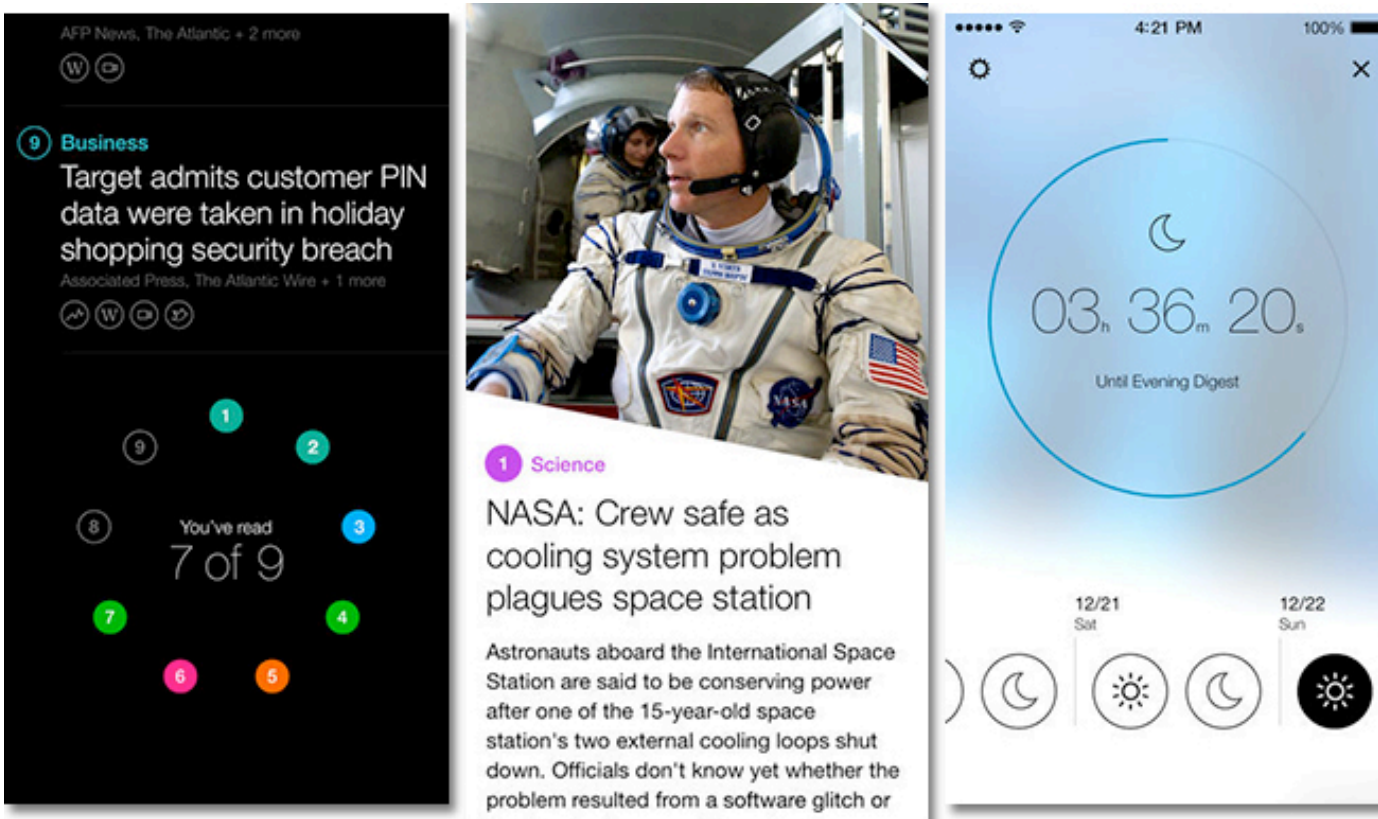
It depends on whether you are talking about African or European woodchucks.

“European woodchucks”

I found 8 European restaurants fairly close to you.

Long-term Coherent Conversation

News aggregation and summary app



Journalism: Robot or Human?

Despite an expected dip in profit, analysts are generally optimistic about **Steelcase** as it prepares to reports its third-quarter earnings on Monday, December 22, 2014. The consensus earnings per share estimate is 26 cents per share.

The consensus estimate remains unchanged over the past month, but it has decreased from three months ago when it was 27 cents. Analysts are expecting earnings of 85 cents per share for the fiscal year. Revenue is projected to be 5% above the year-earlier total of $784.8 million at $826.1 million for the quarter. For the year, revenue is projected to come in at $3.11 billion.

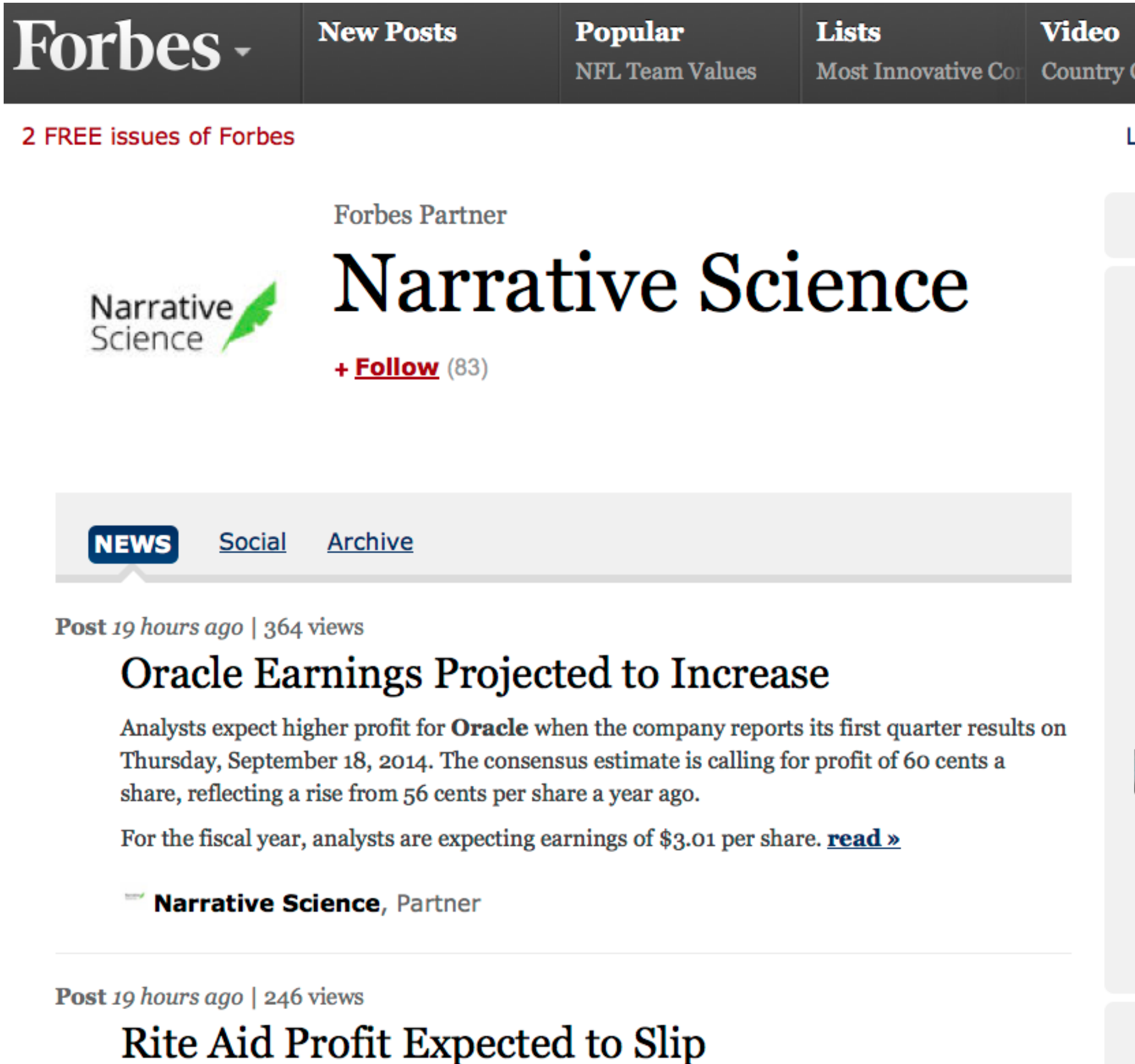
The company has seen revenue grow for three quarters straight. The less than a percent revenue increase brought the figure up to $786.7 million in the most recent quarter. Looking back further, revenue increased 8% in the first quarter from the year earlier and 8% in the fourth quarter.

The majority of analysts (100%) rate Steelcase as a buy. This compares favorably to the analyst ratings of three similar companies, which average 57% buys. Both analysts rate Steelcase as a buy.

[Forbes.com; Dec 19, 2014]

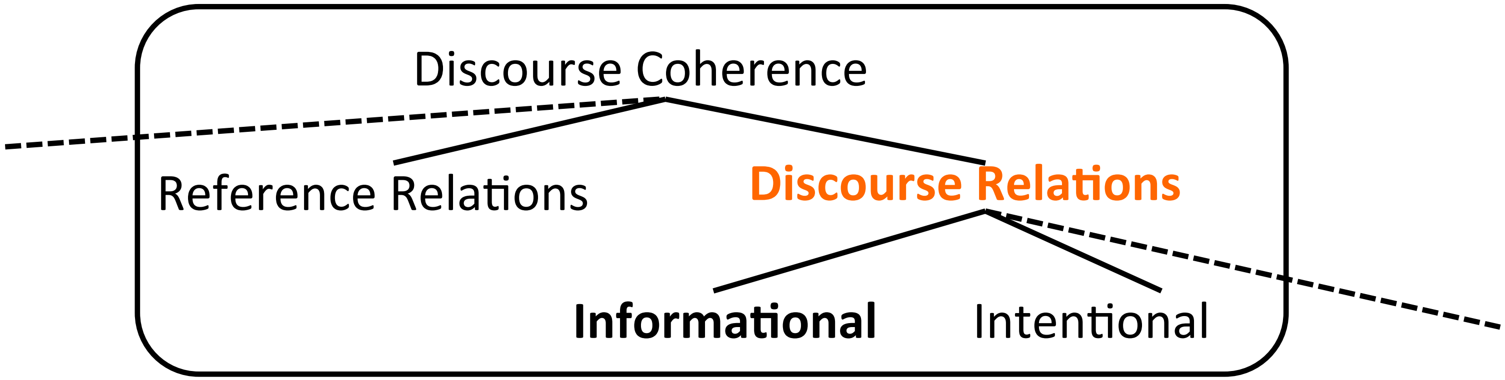
Writer-bots for earthquake & financial reports

While far from op-ed, some of the formulaic news articles are now written by computers.



What is “discourse”?

Discourse is a coherent structured group of textual units



Discourse “Relations”

- John hid Bill’s car keys. He was drunk.
  ➔ “Explanation” relation
- John hid Bill’s car keys. He likes spinach.
  ➔ ??? relation

Discourse "Relations"

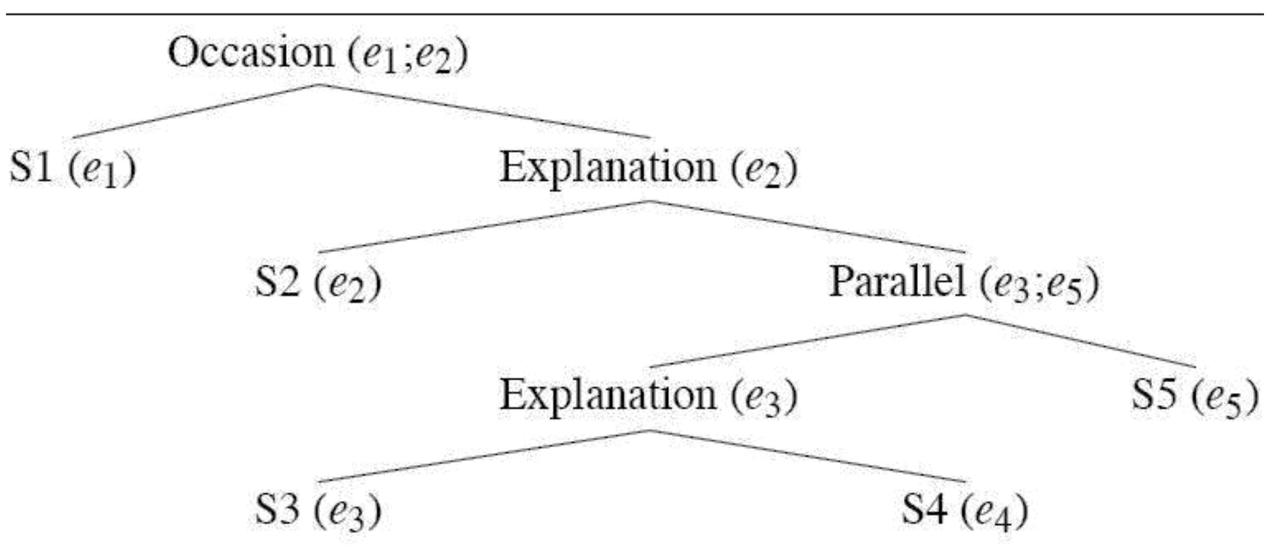
- Dorothy was from Kansas. She lived on the Kansas prairies.
- The tin woodman was caught in the rain. His joints rusted.
- The scarecrow wanted some brains. The tin woodsman wanted a heart.
- Dorothy picked up the oil-can. She oiled the Tin Woodman's joints.

- Result
  - *"as a result …"*
- Occasion
  - *"and then …"*
- Elaboration
  - *"more specifically …"*
- Parallel

Discourse Parsing: Tree of Relations

- Explanation
- Elaboration
- Result
- Parallel
- Occasion

- John went to the bank to deposit the paycheck. (e1)
- He then took a train to Bill's car dealership. (e2)
- He needed to buy a car. (e3)
- The company he works for now isn't near any public transportation. (e4)
- He also wanted to talk to Bill about their softball league. (e5)

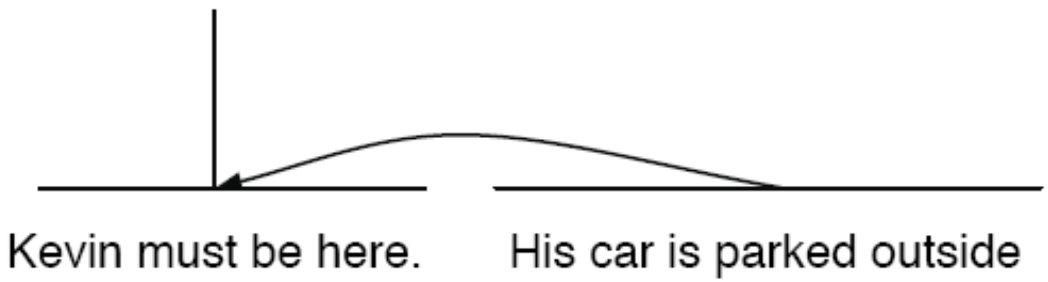


Plan

1. Textual Coherence
2. Theory: Rhetorical Structure Theory (RST)
3. Corpus: Penn Discourse Tree Bank (PDTB)
4. Coreference Resolution

Rhetorical structure theory (RST)

Mann and Thompson, 1987



- **Nucleus** – the central unit, interpretable independently.
- **S**atellite – interpretation depends on N

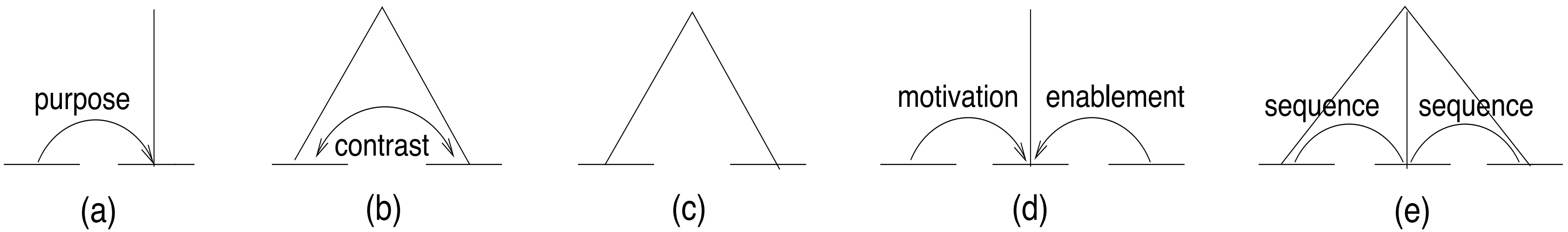
- RST relation --- a set of constraints on the nucleus and satellite, w.r.t. the goals/beliefs/effects of the writer (W) and the reader (R)

Relation Name:	Evidence
Constraints on N:	R might not believe N to a degree satisfactory to W
Constraints on S:	R believes S or will find it credible
Constraints on N+S:	R's comprehending S increases R's belief of N
Effects:	R's belief of N is increased

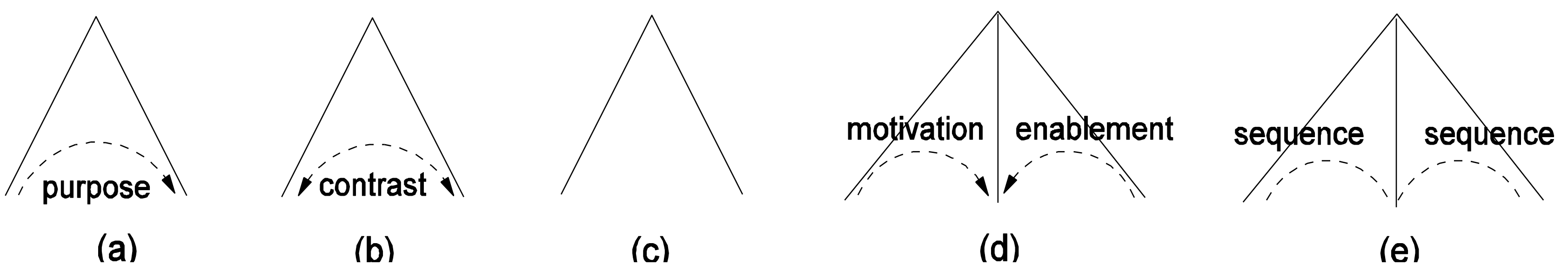
Types of Schemas in RST

RST **schemas** := context-free rules for discourse structure

- whether or not the schema has binary, ternary, or arbitrary branching.
- whether or not the RHS has a head (called a *nucleus*);
- what rhetorical relation, if any, hold between right-hand side (RHS) sisters;



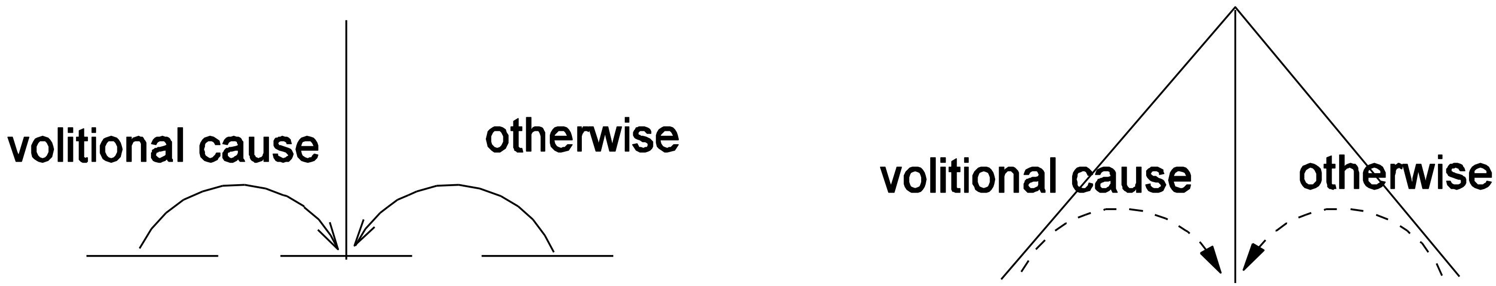
RST schema types in RST annotation



RST schema types in standard tree notation

RST Example

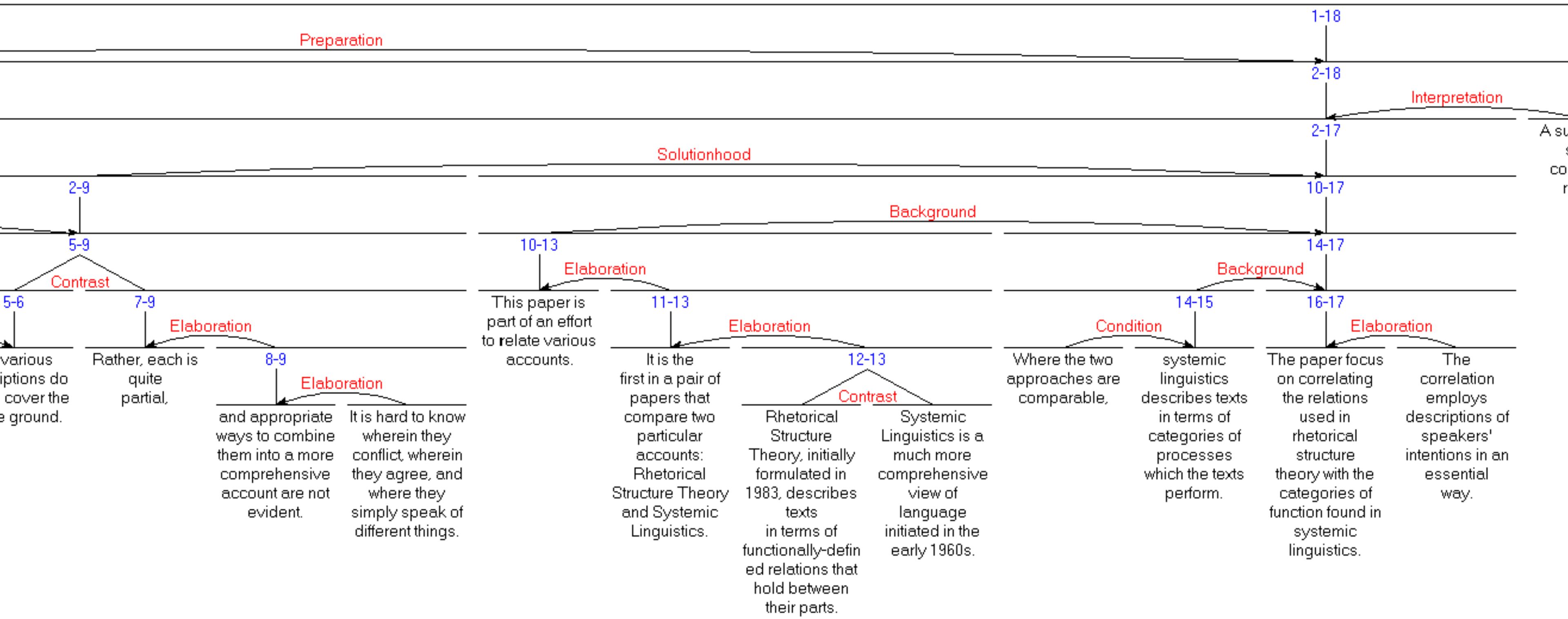
(1) George Bush supports big business.
(2) He's sure to veto House Bill 1711.
(3) Otherwise, big business won't support him.



Discourse structure as a **tree**:

- Leaf := an *elementary discourse unit* (a *continuous text span*)
- non-terminal := a contiguous, non-overlapping text span
- root := a complete, non-overlapping cover of the text

RST Example



1-18
Preparation
2-18
Interpretation
2-17
Solutionhood
2-9
10-17
Background
5-9
Contrast
10-13
Elaboration
14-17
Background
5-6
7-9
Elaboration
This paper is part of an effort to relate various accounts.
11-13
Elaboration
14-15
Condition
16-17
Elaboration
various iptions do cover the e ground.
Rather, each is quite partial,
8-9
Elaboration
and appropriate ways to combine them into a more comprehensive account are not evident.
It is hard to know wherein they conflict, wherein they agree, and where they simply speak of different things.
It is the first in a pair of papers that compare two particular accounts: Rhetorical Structure Theory and Systemic Linguistics.
12-13
Contrast
Rhetorical Structure Theory, initially formulated in 1983, describes texts in terms of functionally-defin ed relations that hold between their parts.
Systemic Linguistics is a much more comprehensive view of language initiated in the early 1960s.
Where the two approaches are comparable,
systemic linguistics describes texts in terms of categories of processes which the texts perform.
The paper focus on correlating the relations used in rhetorical structure theory with the categories of function found in systemic linguistics.
The correlation employs descriptions of speakers' intentions in an essential way.
A su
co

From Theory to TreeBank

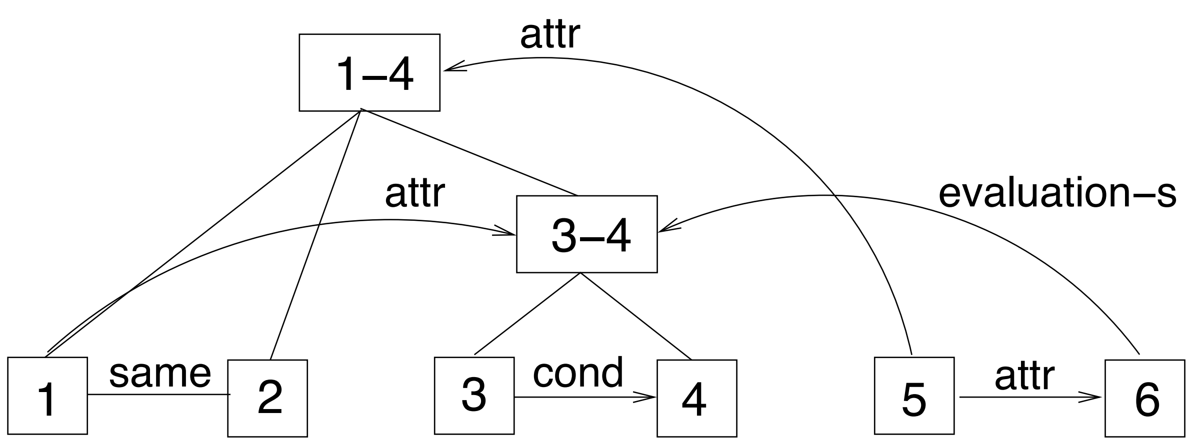
- Rhetorical Structure Theory: Mann and Thompson (1987)
- RST TreeBank: Carlson et al., (2001) defines 78 different RST relations, grouped into 16 classes.

Relation Name:	Evidence
Constraints on N:	R might not believe N to a degree satisfactory to W
Constraints on S:	R believes S or will find it credible
Constraints on N+S:	R's comprehending S increases R's belief of N
Effects:	R's belief of N is increased

Discourse GraphBank [Wolf & Gibson 2005]

Graph instead of a tree:

- (1) The administration should now state
  (2) that
  (3) if the February election is voided by the Sandinistas
  (4) they should call for military aid,
  (5) said former Assistant Secretary of State Elliot Abrams.
  (6) In these circumstances, I think they'd win.



Penn Discourse Treebank (PDTB)

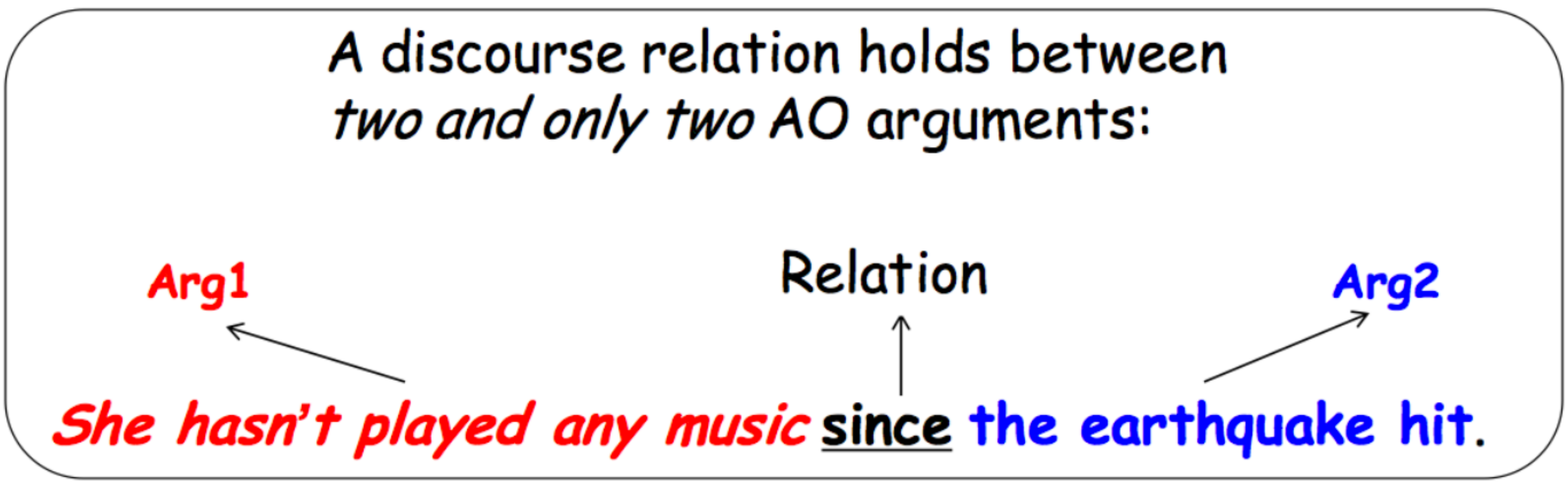
Discourse relations defined over "abstract objects"

Abstract Objects:

events, states, propositions (Asher, 1993)

Example of discourse relations:

Cause, temporal, contrast, condition



Explicit Connectives

Explicit connectives are the lexical items that trigger discourse relations.

- Subordinating conjunctions (e.g., *when*, *because*, *although*, etc.)
  - ➢ The federal government suspended sales of U.S. savings bonds because Congress hasn't lifted the ceiling on government debt.

- Coordinating conjunctions (e.g., *and*, *or*, *so*, *nor*, etc.)
  - ➢ The subject will be written into the prime-time shows, and viewers will be given a 900 number to call.

- Discourse adverbials (e.g., *then*, *however*, *as a result*, etc.)
  - ➢ In the past, the socialist policies of the government strictly limited the profits businessmen could make. As a result, industry operated out of highly inefficient industrial units.

▪ Arg2: the argument with which connective is syntactically associated
▪ *Arg1*: the other argument

Explicit Connectives

Explicit connectives are the lexical items that trigger discourse relations.

- Subordinating conjunctions (e.g., *when*, *because*, *although*, etc.)
  - *The federal government suspended sales of U.S. savings bonds* because Congress hasn't lifted the ceiling on government debt.

- Coordinating conjunctions (e.g., *and*, *or*, *so*, *nor*, etc.)
  - The subject will be written into the prime-time shows, and viewers will be given a 900 number to call.

- Discourse adverbials (e.g., *then*, *however*, *as a result*, etc.)
  - In the past, the socialist policies of the government strictly limited the profits businessmen could make. As a result, industry operated out of highly inefficient industrial units.

- Arg2: the argument with which connective is syntactically associated
- *Arg1*: the other argument

Explicit Connectives

Explicit connectives are the lexical items that trigger discourse relations.

- Subordinating conjunctions (e.g., *when*, *because*, *although*, etc.)
  - *The federal government suspended sales of U.S. savings bonds* because Congress hasn't lifted the ceiling on government debt.
- Coordinating conjunctions (e.g., *and*, *or*, *so*, *nor*, etc.)
  - *The subject will be written into the prime-time shows*, and viewers will be given a 900 number to call.
- Discourse adverbials (e.g., *then*, *however*, *as a result*, etc.)
  - In the past, the socialist policies of the government strictly limited the profits businessmen could make. As a result, industry operated out of highly inefficient industrial units.

- Arg2: the argument with which connective is syntactically associated
- *Arg1*: the other argument

Explicit Connectives

Explicit connectives are the lexical items that trigger discourse relations.

- Subordinating conjunctions (e.g., *when*, *because*, *although*, etc.)
  - *The federal government suspended sales of U.S. savings bonds* because Congress hasn't lifted the ceiling on government debt.

- Coordinating conjunctions (e.g., *and*, *or*, *so*, *nor*, etc.)
  - *The subject will be written into the prime-time shows*, and viewers will be given a 900 number to call.

- Discourse adverbials (e.g., *then*, *however*, *as a result*, etc.)
  - *In the past, the socialist policies of the government strictly limited the profits businessmen could make*. As a result, industry operated out of highly inefficient industrial units.

- Arg2: the argument with which connective is syntactically associated
- *Arg1*: the other argument

Argument Labels and Order

- Arg2 is the argument with which connective is syntactically associated.
- *Arg1* is the other argument.

- Most oil companies, <u>when</u> they set exploration and production budgets for this year, forecast revenue of $15 for each barrel of crude produced.
- The chief culprits, he says, are big companies and business groups that buy huge amounts of land "not for their corporate use, but for resale at huge profit." … The Ministry of Finance, <u>as a result</u>, has proposed a series of measures that would restrict business investment in real estate even more tightly than restrictions aimed at individuals.

Argument Labels and Order

- Arg2 is the argument with which connective is syntactically associated.
- *Arg1* is the other argument.

➢ *Most oil companies*, <u>when</u> they set exploration and production budgets for this year, *forecast revenue of $15 for each barrel of crude produced*.

➢ The chief culprits, he says, are big companies and business groups that buy huge amounts of land "not for their corporate use, but for resale at huge profit." … The Ministry of Finance, <u>as a result</u>, has proposed a series of measures that would restrict business investment in real estate even more tightly than restrictions aimed at individuals.

Argument Labels and Order

- Arg2 is the argument with which connective is syntactically associated.
- *Arg1* is the other argument.

➢ *Most oil companies*, when they set exploration and production budgets for this year, *forecast revenue of $15 for each barrel of crude produced*.

➢ *The chief culprits*, he says, *are big companies and business groups that buy huge amounts of land "not for their corporate use, but for resale at huge profit*." … The Ministry of Finance, as a result, has proposed a series of measures that would restrict business investment in real estate even more tightly than restrictions aimed at individuals.

Relative location of Arg1?

Finding Arg 1: Preliminary Experiment

- where to we find Arg 1 the most often?

CONN	Same	Previous	Multiple Previous	Distant
nevertheless				
otherwise				
as a result				
therefore				
instead				

Finding Arg 1: Preliminary Experiment

- where to we find Arg 1 the most often?
- which connective has highest % of "distant" arg-1 ?

CONN	Same	Previous	Multiple Previous	Distant
nevertheless	9.7%	54.8%		
otherwise	11.1%	**77.8%**		
as a result	**4.8%**	69.8%		
therefore	**55%**	**35%**		
instead	22.7%	63.9%		

Finding Arg 1: Preliminary Experiment

- where to we find Arg 1 the most often?
- which connective has highest % of "distant" arg-1 ?

CONN	Same	Previous	Multiple Previous	Distant
nevertheless	9.7%	54.8%	9.7%	**25.8%**
otherwise	11.1%	**77.8%**	5.6%	5.6%
as a result	**4.8%**	69.8%	7.9%	19%
therefore	**55%**	**35%**	5%	**5%**
instead	22.7%	63.9%	2.1%	11.3%

Hierarchy of PDTB Discourse Relations

CONTINGENCY
- Cause
  - Reason
  - Result
- Condition
  - Hypothetical
  - ...
- ...

COMPARISON
- Contrast
  - Juxtaposition
  - Opposition
- Concession
  - Expectation
  - Contra-expectation
- ...

TEMPORAL
- Asynchronous
- Synchronous
  - Precedence
  - Succession

Operating revenue rose 69% to A$8.48 billion from A $5.01 billion.
But the net interest bill jumped 85% to A$686.7 million from A$371.1 million.

The Texas oilman has acquired a 26.2% stake valued at more than $1.2 billion in an automotive lighting company, Koito Manufacturing Co.
But he has failed to gain any influence at the company.

EXPANSION
- Conjunction
- Instantiation
- Restatement
  - Specification
  - Equivalence
  - Generalization
- ...
- Exception
- List

Annotation Overview (PDTB 1.0): Explicit Connectives

- All WSJ sections (25 sections; 2304 texts)

- 100 distinct types

  - Subordinating conjunctions – 31 types
  - Coordinating conjunctions – 7 types
  - Discourse Adverbials – 62 types

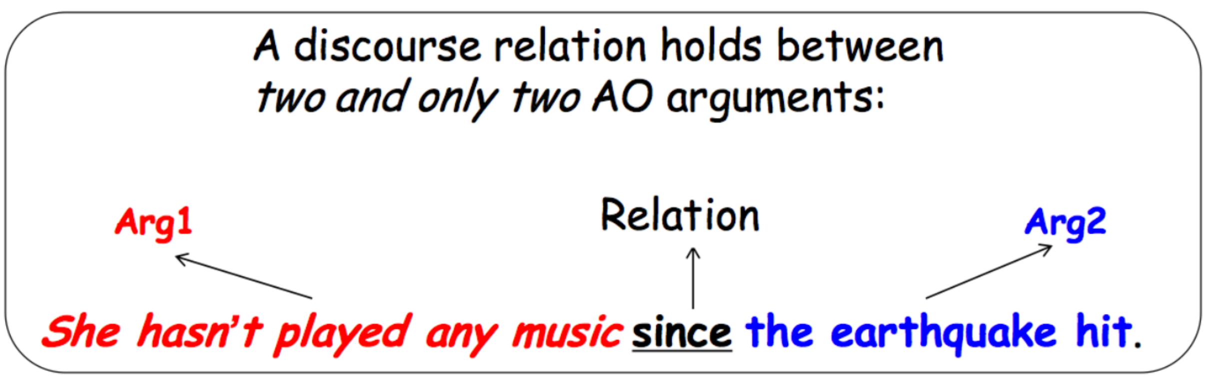
  Some additional types will be annotated for PDTB-2.0.

- 18505 distinct tokens

Natural Language Generation: Sentence Planning

Discourse analysis can help enhancing NLG. How?

- the relative linear order of component semantic units
- whether or not to explicitly realize discourse relations (occurrence), and if so, how to realize them (lexical selection and placement)



NLG: Preliminary Experiment 2

Question: Given a subordinating conjunction and its arguments, in what relative order should the arguments be realized? Arg1-Arg2? Arg2-Arg1?

☞ Different patterns for different connectives

- When almost equally distributed:
  54% (Arg1-Arg2) and **46%** (Arg2-Arg1)

- Although and **(even) though** have opposite patterns:
  Although: **37%** (Arg1-Arg2) and **63%** (Arg2-Arg1)
  (Even) though: **72%** (Arg1-Arg2) and **28%** (Arg2-Arg1)

NLG: Preliminary Experiment 2

Question: What constrains the lexical choice of a connective for a given discourse relation? (Prasad et al., 2005)

- Testing a prediction for lexical choice rule for CAUSAL <u>because</u> and <u>since</u> (Elhadad and McKeown,1990):
  - Assumption: *New* information tends to be placed at the end and *given* information at the beginning.
  - Claim: <u>Because</u> presents *new* information, and <u>since</u> presents *given* information
  - Lexical choice rule: Use <u>because</u> when subordinate clause is postposed (Arg1-Arg2); use <u>since</u> when subordinate clause is preposed (Arg2-Arg1)

☞ <u>Because</u> does tend to appear with Arg1-Arg2 order (90%), but CAUSAL <u>since</u> is equally distributed as Arg1-Arg2 and Arg2-Arg1.

Sense Disambiguation of Connectives

Some discourse connectives are polysemous, e.g.,

- While: comparative, oppositive, concessive
- Since: temporal, causal, temporal/causal
- When: temporal/causal, conditional

Sense disambiguation is required for many applications:

- **Discourse parsing:** identification of arguments
- **NLG:** relative order of arguments
- **MT:** choice of connective in target language

Sense Disambiguation: Preliminary Experiment

- Features (from raw text and PTB):
  - Form of auxiliary *have* - *Has*, *Have*, *Had* or *Not Found*.
  - Form of auxiliary *be* – *Present* (am, is, are), *Past* (was, were), *Been*, or *Not Found*.
  - Form of the head - *Present* (part-of-speech VBP or VBZ), *Past* (VBD), *Past Participial* (VBN), *Present Participial* (VBG).
  - Presence of a modal - *Found* or *Not Found*.
  - Relative position of Arg1 and Arg2: preposed, postposed
  - If the same verb was used in both arguments
  - If the adverb "not" was present in the head verb phrase of a single argument

- MaxEnt classifier (McCallum, 2002)
- Baseline: most frequent sense (**CAUSAL**)
- 10-fold cross-validation

Experiment	Accuracy	Baseline
(T,C,T/C)	75.5%	53.6%
({T,T/C}, C)	90.1%	53.6%
(T,{C,T/C})	74.2%	65.6%
(T,C)	89.5%	60.9%

T=temporal, C=causal, T/C=temporal/causal

☞**15-20% improvement over baseline across the board, with state of the art.**

Robot or Human?

Despite an expected dip in profit, analysts are generally optimistic about Steelcase as it prepares to reports its third-quarter earnings on Monday, December 22, 2014. The consensus earnings per share estimate is 26 cents per share.

The consensus estimate remains unchanged over the past month, but it has decreased from three months ago when it was 27 cents. Analysts are expecting earnings of 85 cents per share for the fiscal year. Revenue is projected to be 5% above the year-earlier total of $784.8 million at $826.1 million for the quarter. For the year, revenue is projected to come in at $3.11 billion.

The company has seen revenue grow for three quarters straight. The less than a percent revenue increase brought the figure up to $786.7 million in the most recent quarter. Looking back further, revenue increased 8% in the first quarter from the year earlier and 8% in the fourth quarter.

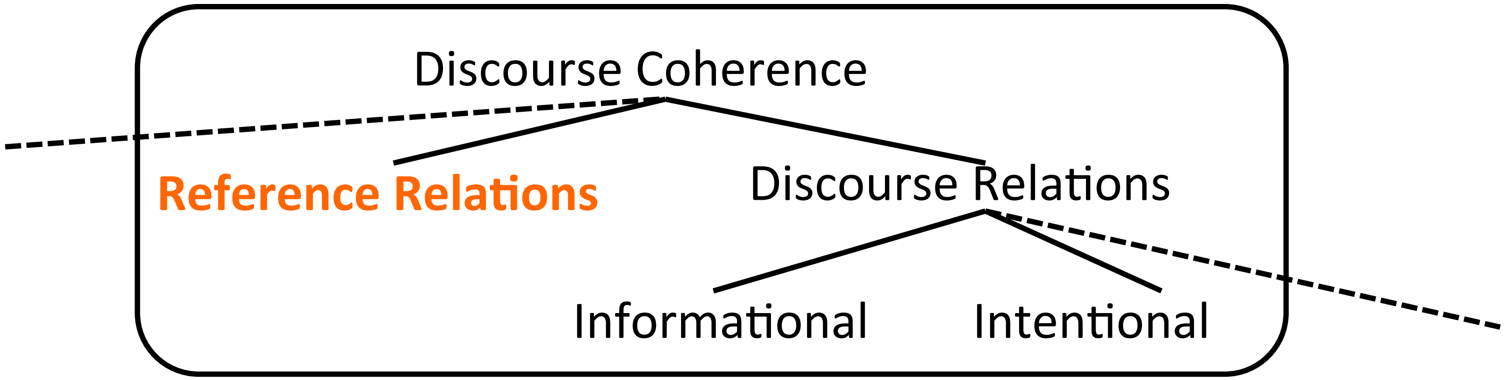
The majority of analysts (100%) rate Steelcase as a buy. This compares favorably to the analyst ratings of three similar companies, which average 57% buys. Both analysts rate Steelcase as a buy.

Plan

1. Textual Coherence
2. Rhetorical Structure Theory (RST)
3. Penn Discourse Tree Bank (PDTB)
4. Coreference Resolution

Discourse Coherence

Discourse is a coherent structured group of textual units



The Problem: Find and Cluster Mentions

Victoria Chen, Chief Financial Officer of Megabucks banking corp since 2004, saw her pay jump 20%, to $1.3 million, as the 37 year old also became the Denver-based financial services company's president. It has been ten years since she came to Megabucks from rival Lotsabucks.

Mention Detection

[Victoria Chen], [Chief Financial Officer of [Megabucks banking corp] since 2004], saw [[her] pay] jump 20%, to $1.3 million, as [the 37 year old] also became the [[Denver-based financial services company]'s president]. It has been ten years since [she] came to [Megabucks] from rival [Lotsabucks].

The Problem: Find and Cluster Mentions

[Victoria Chen], [Chief Financial Officer of [Megabucks banking corp] since 2004], saw [[her] pay] jump 20%, to $1.3 million, as [the 37 year old] also became the [[Denver-based financial services company]'s president]. It has been ten years since [she] came to [Megabucks] from rival [Lotsabucks].

Mention Clustering

Co-reference chains:

1 {Victoria Chen, Chief Financial Officer...since 2004, her, the 37-year-old, the Denver-based financial services company's president}

2 {Megabucks Banking Corp, Denver-based financial services company, Megabucks}

3 {her pay}

4 {rival Lotsabucks}

Types of Coreference (I)

- Types of coreferent phrase:
  - Referential ("semantically definite") NPs

    The author of the book walked in.

    His name was John Smith.

    Mr. Smith said…
  - Anaphors

    Mr. Smith walked in.

    He talked about his car.
  - Descriptive NPs

    The stock price fell from $4.02 to $3.85.

Types of Coreference (II)

- Types of antecedent:
  - Non-generic referring NPs
    - *Mr. Smith likes his car.*
  - Generic referring NPs
    - *People like their cars.*
  - Non-referring NPs
    - *No one talked about their car.*
  - Clauses
    - *Driving fast isn't safe, but it's fun.*

Coreference as Clustering

The coreference problem can be solved by assigning all NPs in the text to equivalence classes, i.e., by clustering. [Cardie and Wagstaff, 1999]

We need:

- a *representation* of NPs (as a set of features)
- a *distance metric*
- a clustering *algorithm*.

Representing Mentions

Each NP is represented as a set of features:

- head noun: last word of the NP;
- position in the document;
- pronoun type: nominative, accusative, possessive, ambiguous;
- article: indefinite, definite, none;
- appositive: based on heuristics (commas, etc.)
- number: plural, singular;
- proper name: based on heuristics (capitalization, etc.);
- semantic class: based on Wordnet;
- gender: masculine, feminine, either, neuter;
- animacy: based on semantic class.

Example Mentions

Words, Head Noun (in **bold**)	Position	Pronoun Type	Article	Appositive	Number	Proper Name	Semantic Class	Gender	Animacy
John **Simon**	1	NONE	NONE	NO	SING	YES	HUMAN	MASC	ANIM
Chief Financial **Officer**	2	NONE	NONE	NO	SING	NO	HUMAN	EITHER	ANIM
Prime **Corp.**	3	NONE	NONE	NO	SING	NO	COMPANY	NEUTER	INANIM
1986	4	NONE	NONE	NO	PLURAL	NO	NUMBER	NEUTER	INANIM
his	5	POSS	NONE	NO	SING	NO	HUMAN	MASC	ANIM
pay	6	NONE	NONE	NO	SING	NO	PAYMENT	NEUTER	INANIM
20%	7	NONE	NONE	NO	PLURAL	NO	PERCENT	NEUTER	INANIM
$1.3 **million**	8	NONE	NONE	NO	PLURAL	NO	MONEY	NEUTER	INANIM
the **37-year-old**	9	NONE	DEF	NO	SING	NO	HUMAN	EITHER	ANIM
the financial-services **company**	10	NONE	DEF	NO	SING	NO	COMPANY	NEUTER	INANIM
president	11	NONE	NONE	NO	SING	NO	HUMAN	EITHER	ANIM

Clustering

Distance Metric $$dist(NP_1, NP_2) = \sum_{f \in F} w_f \cdot incompatibility_f(NP_1, NP_2)$$

Feature f	Weight	Incompatibility function
Words	10.0	(# of mismatching words[a]) / (# of words in the longer NP)
Head Noun	1.0	1 if the head nouns differ; else 0
Position	5.0	(difference in position) / (maximum difference in document)
Pronoun	r	1 if NP_i is a pronoun and NP_j is not; else 0
Article	r	1 if NP_j is indefinite and not appositive; else 0
Words–Substring	$-\infty$	1 if NP_i subsumes (entirely includes as a substring) NP_j;
Appositive	$-\infty$	1 if NP_j is appositive and NP_i is its immediate predecessor; else 0
Number	∞	1 if they do not match in number; else 0
Proper Name	∞	1 if both are proper names, but mismatch on every word; else 0
Semantic Class	∞	1 if they do not match in class; else 0
Gender	∞	1 if they do not match in gender (allows EITHER to match MASC or FEM); else 0
Animacy	∞	1 if they do not match in animacy; else 0

compatible classes, compute transitive closure

Pairwise Model: Features matter! [Bengston & Roth, 2008]

Category	Feature	Source
Mention Types	Mention Type Pair	Annotation and tokens
String Relations	Head Match	Tokens
	Extent Match	Tokens
	Substring	Tokens
	Modifiers Match	Tokens
	Alias	Tokens and lists
Semantic	Gender Match	WordNet and lists
	Number Match	WordNet and lists
	Synonyms	WordNet
	Antonyms	WordNet
	Hypernyms	WordNet
	Both Speak	Context
Relative Location	Apposition	Positions and context
	Relative Pronoun	Positions and tokens
	Distances	Positions
Learned	Anaphoricity	Learned
	Name Modifiers Predicted Match	Learned
Aligned Modifiers	Aligned Modifiers Relation	WordNet and lists
Memorization	Last Words	Tokens
Predicted Entity Types	Entity Types Match	Annotation and tokens
	Entity Type Pair	WordNet and tokens

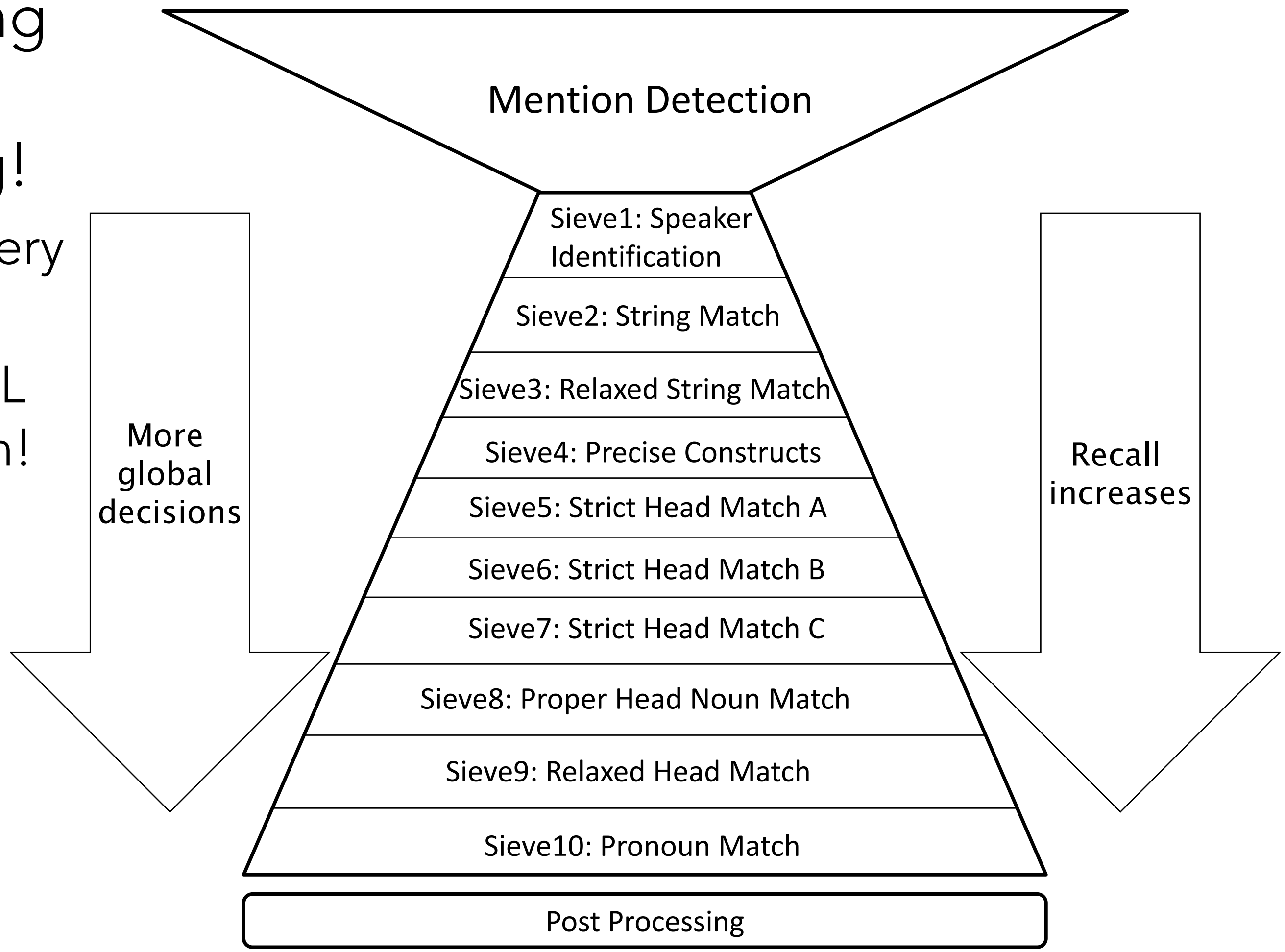
Two Recent Supervised Learners

- Linear Model
  - [Bengston & Roth 2008]
  - Pairwise classification
  - Careful experimental setup with tons of features!
  - 80.8 B^3 F1

- FOL-based approach
  - [Culotta et al. 2007]
  - Includes global constraints on clusters
  - 79.3 B^3 F1

Multi-pass Sieve

- Basically, a ranking model with no machine learning!
  - 10 sieves, each very simple
  - Winner of CONLL 2011 competition!



A Carefully Constructed Example

Input:	John is a musician. He played a new song. A girl was listening to the song. "It is my favorite," John said to her.
Mention Detection:	[John]$^1_1$ is [a musician]$^2_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^6_6$. "[It]$^7_7$ is [[my]$^9_9$ favorite]$^8_8$," [John]$^{10}_{10}$ said to [her]$^{11}_{11}$.
Speaker Sieve:	[John]$^1_1$ is [a musician]$^2_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^6_6$. "[It]$^7_7$ is [**[my]**$^9_9$ favorite]$^8_8$," **[John]**$^9_{10}$ said to [her]$^{11}_{11}$.
String Match:	**[John]**$^1_1$ is [a musician]$^2_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^6_6$. "[It]$^7_7$ is [[my]$^1_9$ favorite]$^8_8$," **[John]**$^1_{10}$ said to [her]$^{11}_{11}$.
Relaxed String Match:	[John]$^1_1$ is [a musician]$^2_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^6_6$. "[It]$^7_7$ is [[my]$^1_9$ favorite]$^8_8$," [John]$^1_{10}$ said to [her]$^{11}_{11}$.
Precise Constructs:	**[John]**$^1_1$ is **[a musician]**$^1_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^6_6$. "**[It]**$^7_7$ is **[[my]**$^1_9$ **favorite]**$^7_8$," [John]$^1_{10}$ said to [her]$^{11}_{11}$.
Strict Head Match A:	[John]$^1_1$ is [a musician]$^1_2$. [He]$^3_3$ played **[a new song]**$^4_4$. [A girl]$^5_5$ was listening to **[the song]**$^4_6$. "[It]$^7_7$ is [[my]$^1_9$ favorite]$^7_8$," [John]$^1_{10}$ said to [her]$^{11}_{11}$.
Strict Head Match B,C:	[John]$^1_1$ is [a musician]$^1_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^4_6$. "[It]$^7_7$ is [[my]$^1_9$ favorite]$^7_8$," [John]$^1_{10}$ said to [her]$^{11}_{11}$.
Proper Head Noun Match:	[John]$^1_1$ is [a musician]$^1_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^4_6$. "[It]$^7_7$ is [[my]$^1_9$ favorite]$^7_8$," [John]$^1_{10}$ said to [her]$^{11}_{11}$.
Relaxed Head Match:	[John]$^1_1$ is [a musician]$^1_2$. [He]$^3_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^4_6$. "[It]$^7_7$ is [[my]$^1_9$ favorite]$^7_8$," [John]$^1_{10}$ said to [her]$^{11}_{11}$.
Pronoun Match:	**[John]**$^1_1$ is [a musician]$^1_2$. **[He]**$^1_3$ played [a new song]$^4_4$. **[A girl]**$^5_5$ was listening to **[the song]**$^4_6$. "**[It]**$^4_7$ is [[my]$^1_9$ favorite]$^4_8$," [John]$^1_{10}$ said to **[her]**$^5_{11}$.
Post Processing:	[John]$^1_1$ is **a musician**. [He]$^1_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^4_6$. "[It]$^4_7$ is **[my]**$^1_9$ **favorite**," [John]$^1_{10}$ said to [her]$^5_{11}$.
Final Output:	[John]$^1_1$ is a musician. [He]$^1_3$ played [a new song]$^4_4$. [A girl]$^5_5$ was listening to [the song]$^4_6$. "[It]$^4_7$ is [my]$^1_9$ favorite," [John]$^1_{10}$ said to [her]$^5_{11}$.

Table 1

The Most Useful Sieves

- **2: Exact string match --** e.g., *[the Shahab 3 ground-ground missile]* and *[the Shahab 3 ground-ground missile]*. Precision is over 90% B3 **[+16 F1]**
- **5: Entity head match** – The mention head word matches *any* head word of mentions in the antecedent entity. Also, looks ar modifiers, e.g. to separate *Harvard University* and *Yale University*. **[+3 F1]**
- **10: Pronominal Coreference Resolution** – observe constraints on number, gender, person, animacy, and NER types. Link to closest, with a maximum distance. **[+10 F1]**
- Most others get between 0-2 points improvement, but are cumulative

Some Results

System	MUC			B^3		
	R	P	F1	R	P	F1
ACE2004-Culotta-Test						
This paper	70.2	82.7	75.9	74.5	88.7	81.0
Haghighi and Klein (2009)	77.7	74.8	79.6	78.5	79.6	79.0
Culotta et al. (2007)	–	–	–	73.2	86.7	79.3
Bengston and Roth (2008)	69.9	82.7	75.8	74.5	88.3	80.8
ACE2004-nwire						
This paper	75.1	84.6	79.6	74.1	87.3	80.2
Haghighi and Klein (2009)	75.9	77.0	76.5	74.5	79.4	76.9
Poon and Domingos (2008)	70.5	71.3	70.9	–	–	–
Finkel and Manning (2008)	58.5	78.7	67.1	65.2	86.8	74.5
MUC6-Test						
This paper	69.1	90.6	78.4	63.1	90.6	74.4
Haghighi and Klein (2009)	77.3	87.2	81.9	67.3	84.7	75.0
Poon and Domingos (2008)	75.8	83.0	79.2	–	–	–
Finkel and Manning (2008)	55.1	89.7	68.3	49.7	90.9	64.3

Table 5
Comparison of our system with the other reported results on the ACE and MUC corpora. All these systems use gold mention boundaries.

[Lee et al, 2013]

Back to … Textual Coherence

- John went to his favorite music store to buy a piano.
- He had frequented the store for many years.
- He was excited that he could finally buy a piano.
- He arrived just as the store was closing for the day.

- John went to his favorite music store to buy a piano.
- It was a store John had frequented for many years.
- He was excited that he could finally buy a piano.
- It was closing just as John arrived.

➔ Same content, different realization through different syntactic choices

Centering Theory (Grosz et al.,1983)

- John went to his favorite music store to buy a piano.
- He had frequented the store for many years.
- He was excited that he could finally buy a piano.
- He arrived just as the store was closing for the day.

❑ Focus is the most salient entity in a discourse segment

❑ Constraints on linguistic realization of focus

– Focus is more likely to be realized as subject or object

– Focus is more likely to be referred to with anaphoric expression

❑ Constraints on the entity distribution in a coherent text

– Transition between adjacent sentences is characterized in terms of focus switch

Entity-grid Model

1. [Former Chilean dictator Augusto Pinochet]$_{\mathbf{S}}$, was arrested in [London]$_{\mathbf{X}}$ on [October 14th]$_{\mathbf{X}}$ 1998.
2. [Pinochet]$_{\mathbf{S}}$, 82, was recovering from [surgery]$_{\mathbf{X}}$.
3. [The arrest]$_{\mathbf{S}}$ was in [response]$_{\mathbf{X}}$ to [an extradition warrant]$_{\mathbf{X}}$ served by [a Spanish judge]$_{\mathbf{S}}$.
4. [Pinochet]$_{\mathbf{S}}$ was charged with murdering [thousands]$_{\mathbf{O}}$, including many [Spaniards]$_{\mathbf{O}}$.
5. [He]$_{\mathbf{S}}$ is awaiting [a hearing]$_{\mathbf{O}}$, [his fate]$_{\mathbf{X}}$ in [the balance]$_{\mathbf{X}}$.
6. [American scholars]$_{\mathbf{S}}$ applauded the [arrest]$_{\mathbf{O}}$.

Notation: **S**=subjects, **O**=object, **X**=other

Entity-grid Model

1. [Former Chilean dictator Augusto Pinochet]$_{\mathbf{S}}$, was arrested in [London]$_{\mathbf{X}}$ on [October 14]$_{\mathbf{X}}$ 1998.
2. [Pinochet]$_{\mathbf{S}}$, 82, was recovering from [surgery]$_{\mathbf{X}}$.
3. [The arrest]$_{\mathbf{S}}$ was in [response]$_{\mathbf{X}}$ to [an extradition warrant]$_{\mathbf{X}}$ served by [a Spanish judge]$_{\mathbf{S}}$.
4. [Pinochet]$_{\mathbf{S}}$ was charged with murdering [thousands]$_{\mathbf{O}}$, including many [Spaniards]$_{\mathbf{O}}$.
5. [He]$_{\mathbf{S}}$ is awaiting [a hearing]$_{\mathbf{O}}$, [his fate]$_{\mathbf{X}}$ in [the balance]$_{\mathbf{X}}$.
6. [American scholars]$_{\mathbf{S}}$ applauded the [arrest]$_{\mathbf{O}}$.

Entity-grid Model

	Pinochet	London	October	Surgery	Arrest	Extradition	Warrant	Judge	Thousands	Spaniards	Hearing	Fate	Balance	Scholars	
1	s	x	x	–	–	–	–	–	–	–	–	–	–	–	1
2	s	–	–	x	–	–	–	–	–	–	–	–	–	–	2
3	–	–	–	–	s	x	x	s	–	–	–	–	–	–	3
4	s	–	–	–	–	–	–	–	o	o	–	–	–	–	4
5	s	–	–	–	–	–	–	–	–	–	o	x	x	–	5
6	–	–	–	–	o	–	–	–	–	–	–	–	–	s	6

Comparing Grids

S	S	S	X	X	–	–	–	–	–	–	–	–	–	–
–	–	S	–	–	X	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	S	X	X	O	–	–	–	–	–
–	–	S	–	–	–	–	–	–	–	O	O	–	–	–
–	–	S	–	–	–	–	–	–	–	–	–	O	X	X
–	–	–	–	–	–	O	–	–	–	–	–	–	–	–

S	S	X	X	X	–	–	–	–	–	–	–	–	–	X
–	–	X	–	–	X	–	–	–	–	–	–	–	–	X
–	–	X	–	–	–	–	X	X	O	–	–	–	–	X
–	–	X	–	–	–	–	–	–	–	O	O	–	–	X
–	–	X	–	–	–	–	–	–	–	–	–	O	X	X
–	–	X	–	–	–	O	–	–	–	–	–	–	–	X

Quantifying Textual Coherence

- Text is encoded as a distribution over entity transition types
- Entity transition type — $\{\text{s}, \text{o}, \text{x}, –\}^n$

	S S	S O	S X	S –	O S	O O	O X	O –	X S	X O	X X	X –	– S	– O	– X	– –
d_{i1}	0	0	0	.03	0	0	0	.02	.07	0	0	.12	.02	.02	.05	.25
d_{i2}	.02	0	0	.03	0	0	0	.06	0	0	0	.05	.03	.07	.07	.29

How to select relevant transition types?:

- Use all the unigrams, bigrams, . . . over $\{\text{s}, \text{o}, \text{x}, –\}$
- Do feature selection

Evaluation / Applications

Goal: recover the most coherent sentence ordering

Basic set-up:
– Input: a pair of a source document and a permutation of its sentences
– Task: find a source document via coherence ranking

Data: Training 4000 pairs, Testing 4000 pairs (Natural disasters and Transportation Safety Reports)

Conclusion

- Computational modeling of discourse coherence
- Theories: Rhetorical structure theory / Centering theory
- Corpus: Penn Discourse Tree Bank
- Applications:
  - better summarization
  - automatic ESL grading
  - better QA (sharp et al., NAACL 2010)
  - better machine translation (this workshop!)
- Coreference Relations
- Entity-grid Models