

Computational Discourse

Jon Park

DEPARTMENT OF COMPUTER SCIENCE
CORNELL UNIVERSITY

Overview

Discourse Segmentation

Discourse Parsing

~~Reference Resolution~~

TOPICS

Overview

Discourse Segmentation

Discourse Parsing

Reference Resolution

TOPICS

Jack took out a match. He lit a candle.

What is a discourse?

Jack took out a match. He lit a candle.

Jack took out a match. The sun set.

What is a discourse?

Jack took out a match. He lit a candle.

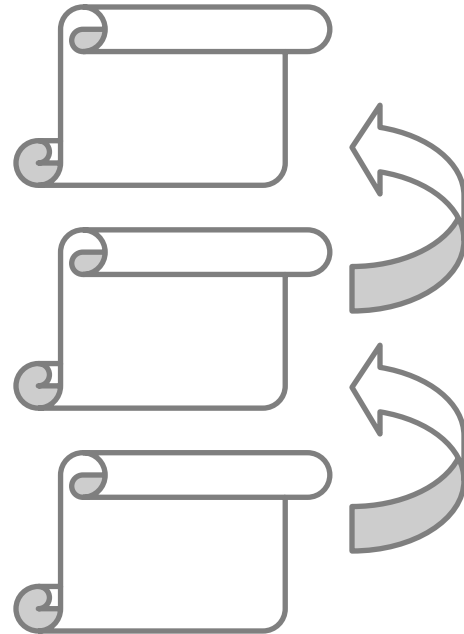
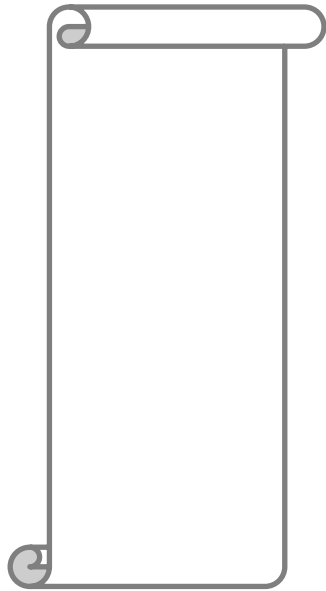
Jack took out a match. The sun set.

Jack took out a match. He likes
tangerines.

What is a discourse?

A discourse consists of
collocated, structured,
coherent groups of sentences,
such as monologues & dialogues.

What is a discourse?



Goal: Discourse Parsing

Overview

Discourse Segmentation

Discourse Parsing

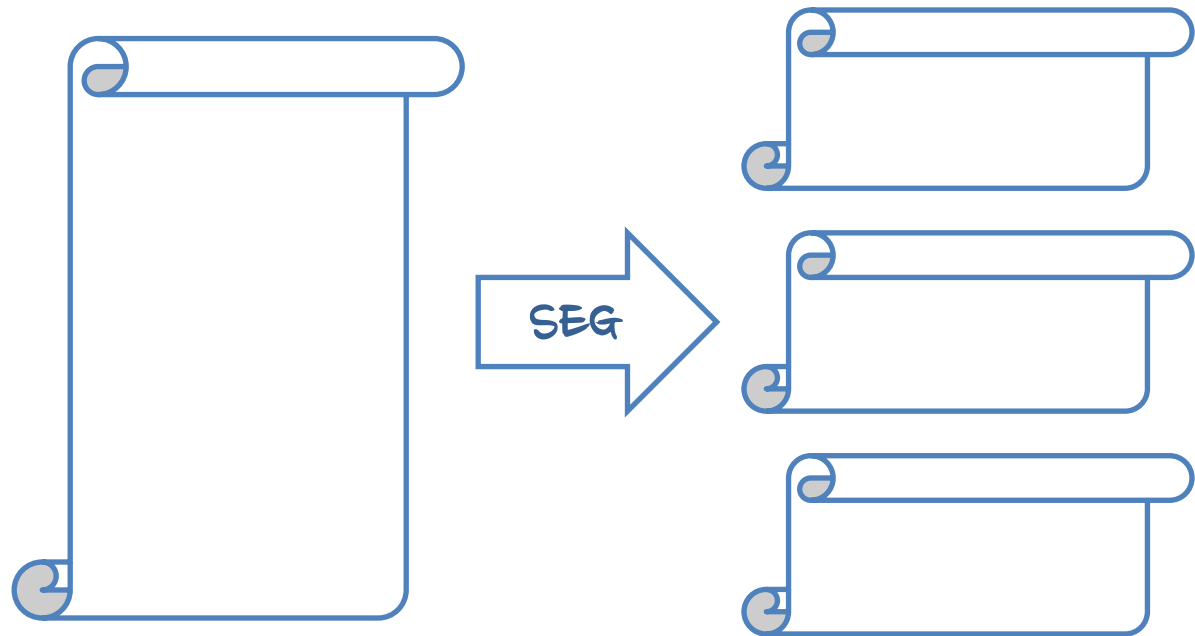
Reference Resolution

TOPICS

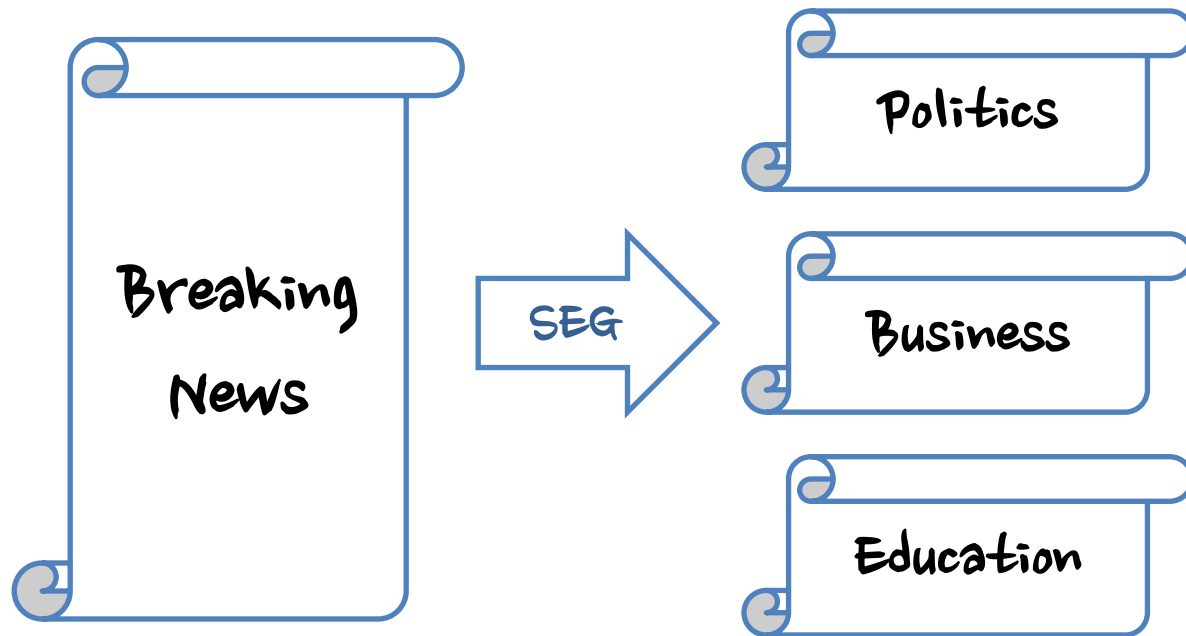
Discourse Segmentation:

Segmenting a document into a
linear sequence of subtopics

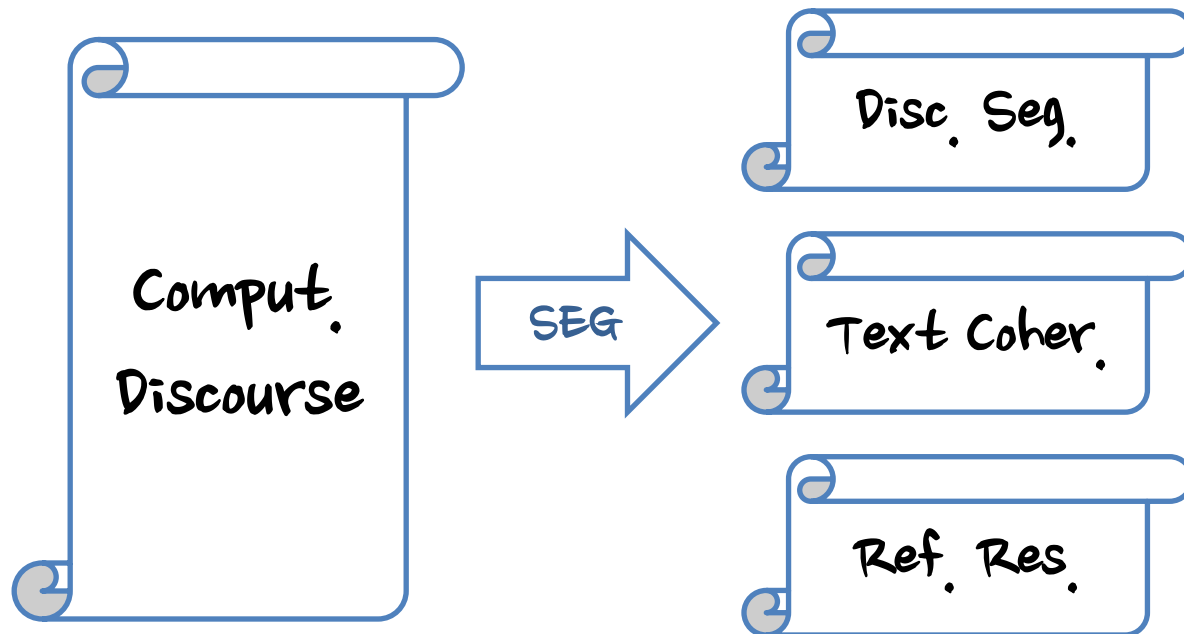
What?



What?



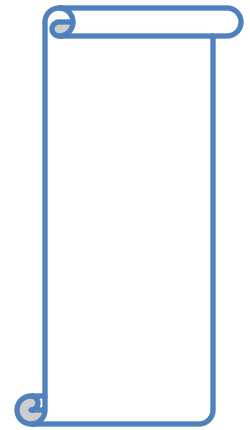
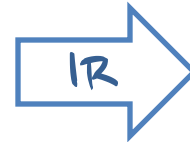
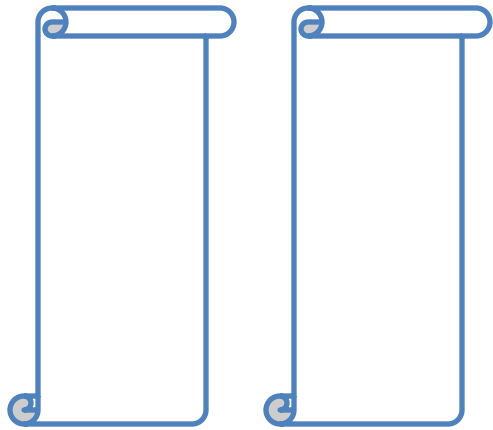
Example



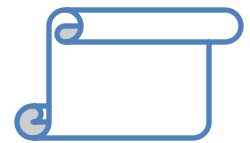
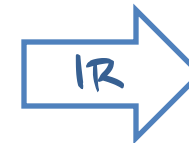
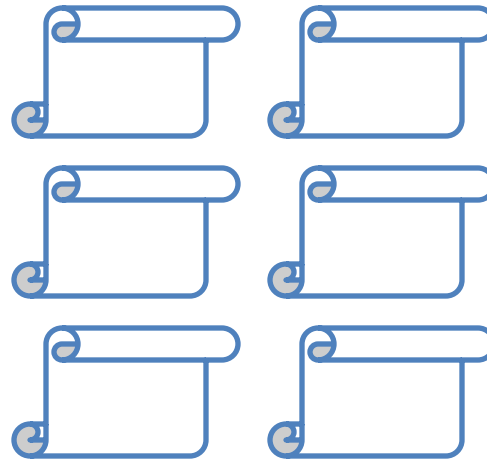
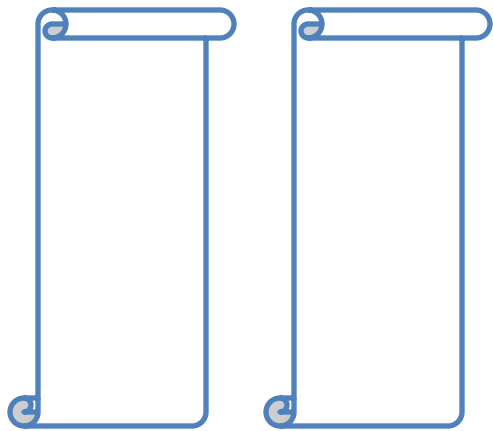
Example

Is discourse segmentation
relevant for something
other than discourse
parsing?

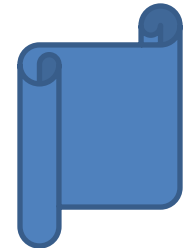
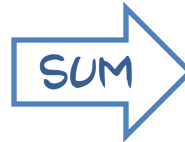
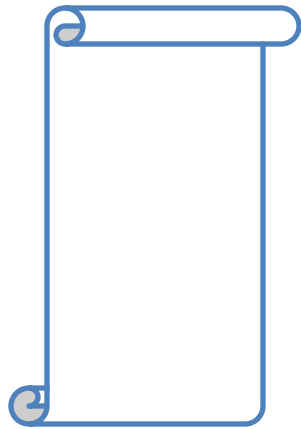
Why?



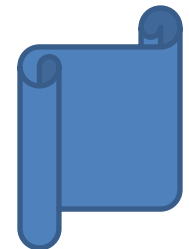
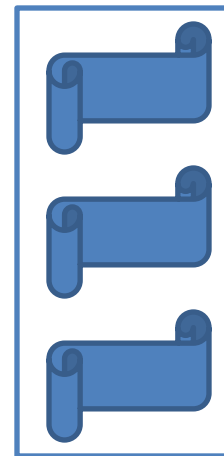
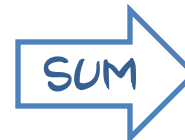
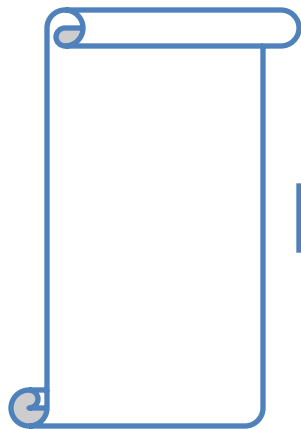
VS



Information Retrieval



VS



Text Summarization

2 major approaches:

- Unsupervised
- Supervised

How?

2 major approaches:

- Unsupervised

- Supervised

How?

Cohesive textual units
can be group by meaning
(via linguistic devices)

Cohesion

Identical words:

I built a chimney, and shingled the sides of my house.

I have thus a tight shingled house.

Cohesion

Hyponym/hypernym:

Peel, core and slice the pears and
the apples.

Add the fruit to the skillet.

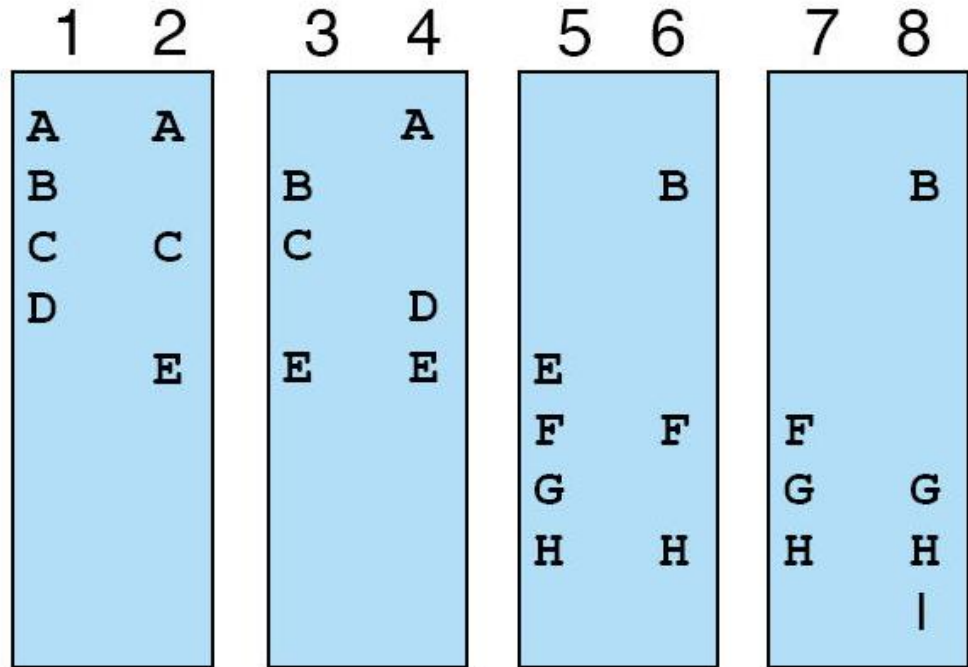
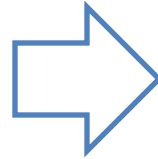
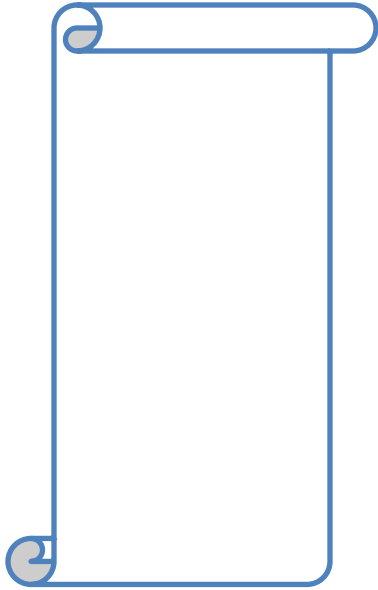
Cohesion

Anaphora:

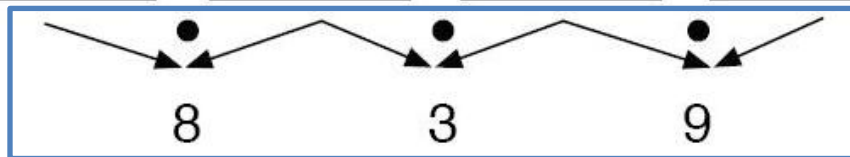
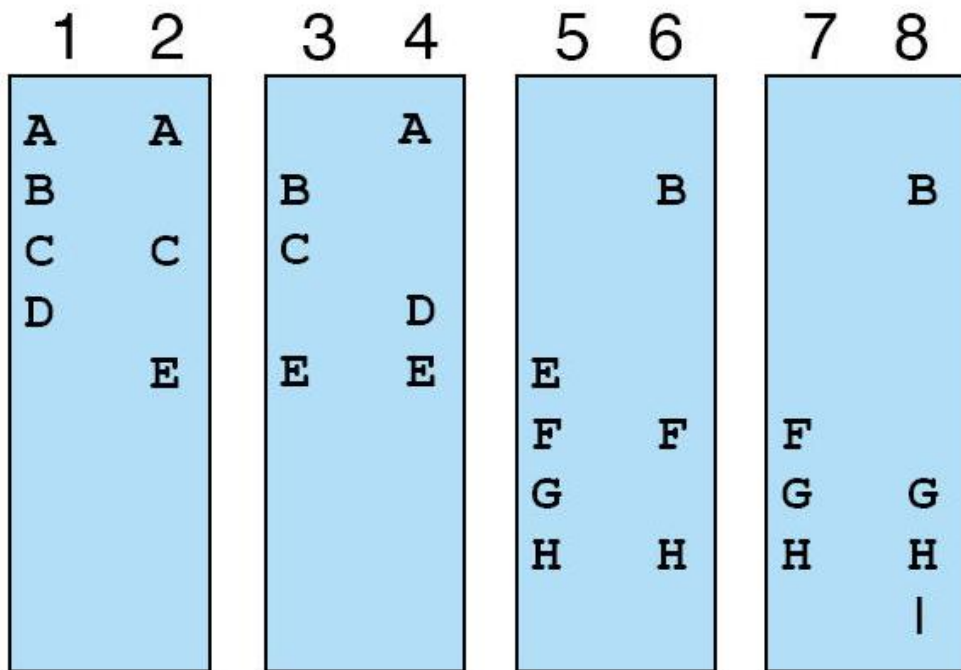
The woodhouses were first in
consequence there.

All looked up to them.

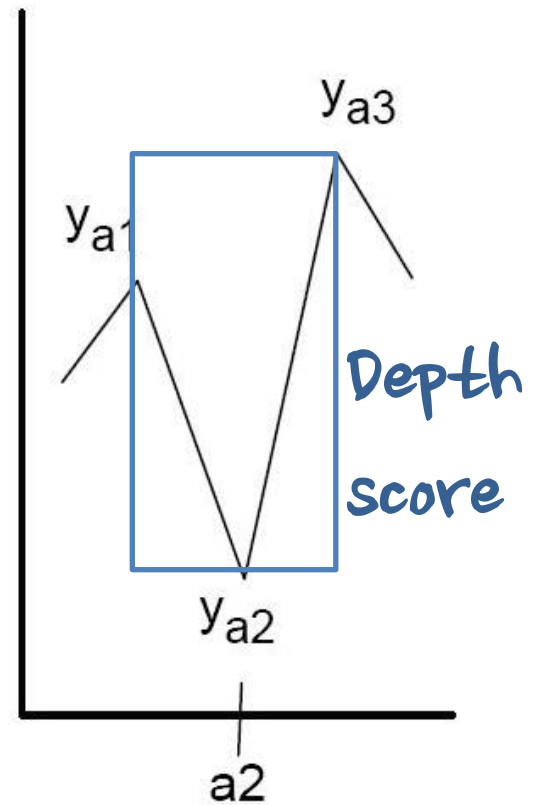
Cohesion



TextTiling



Cohesion scores



TextTiling

2 major approaches:

- Unsupervised

- Supervised

How?

Overview

Discourse Segmentation

Discourse Parsing

~~Reference Resolution~~

TOPICS

A discourse consists of
collocated, structured,
coherent groups of sentences,
such as monologues & dialogues.

What is a discourse?

Coherent textual units are
linked by meaning relations:
result, explanation, parallel,
elaboration, occasion, etc.

Coherence

Result:

The Tin Woodman was caught in
the rain.

His joints rusted.

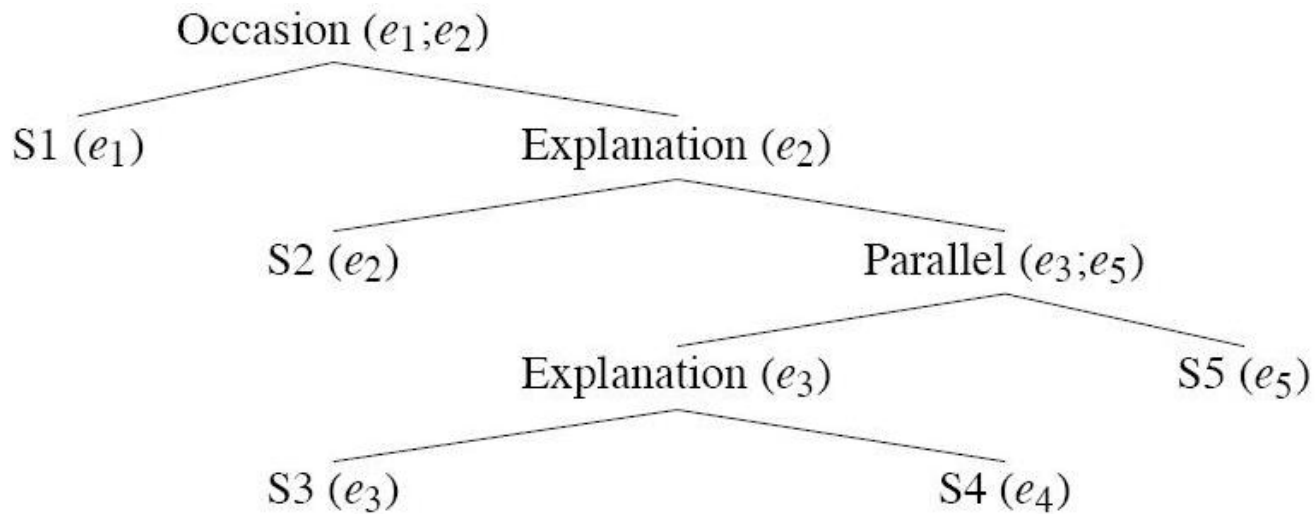
Coherence Relation

Elaboration:

Dorothy was from Kansas.

She lived in the midst of the
great Kansas Prairies.

Coherence Relation



Discourse Structure (Hobbs)



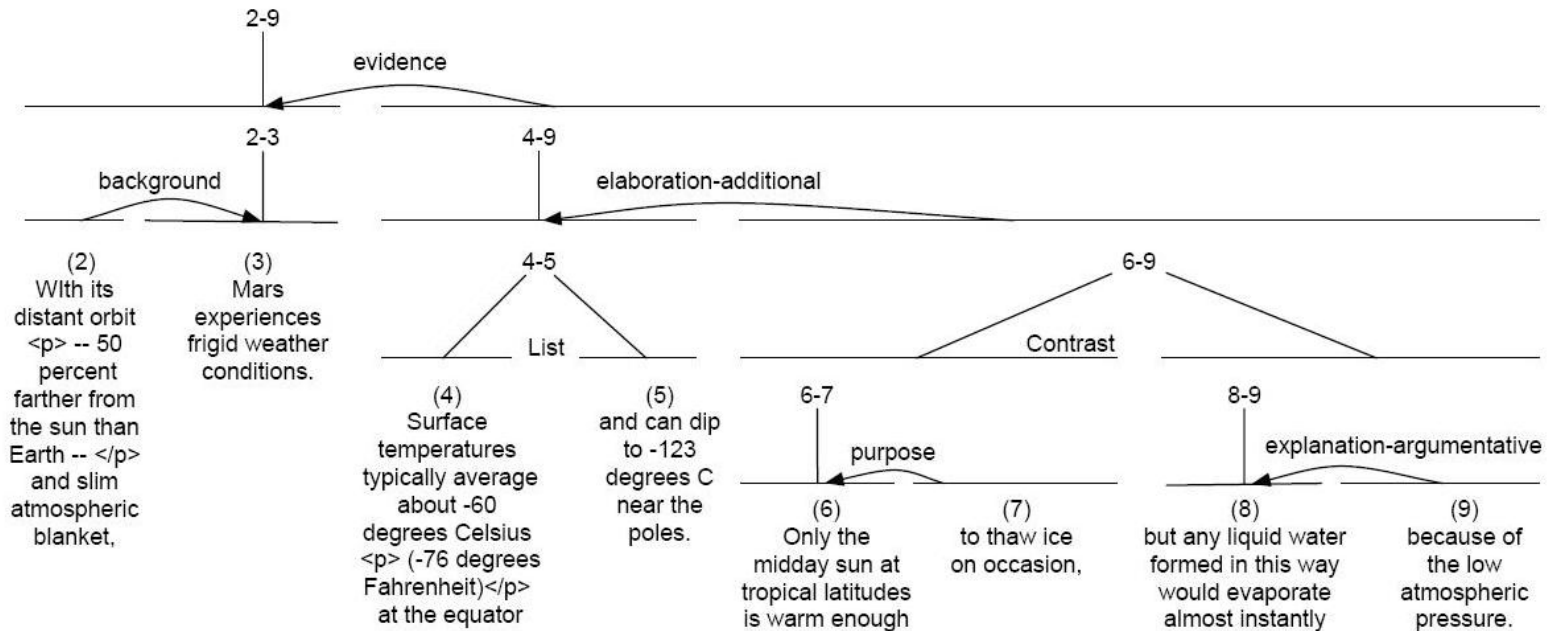
Kevin must be here.

His car is parked outside

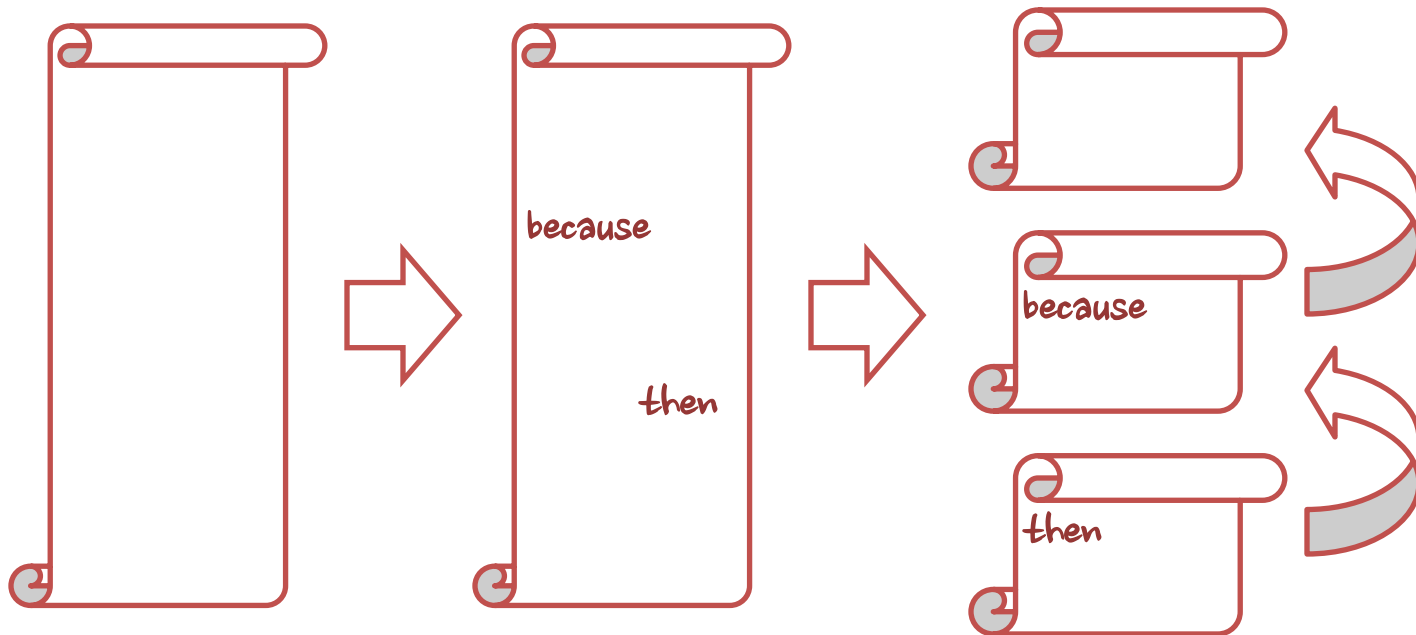
Relation Name:	Evidence
Constraints on N:	R might not believe N to a degree satisfactory to W
Constraints on S:	R believes S or will find it credible
Constraints on N+S:	R's comprehending S increases R's belief of N
Effects:	R's belief of N is increased

Rhetorical Structure Theory

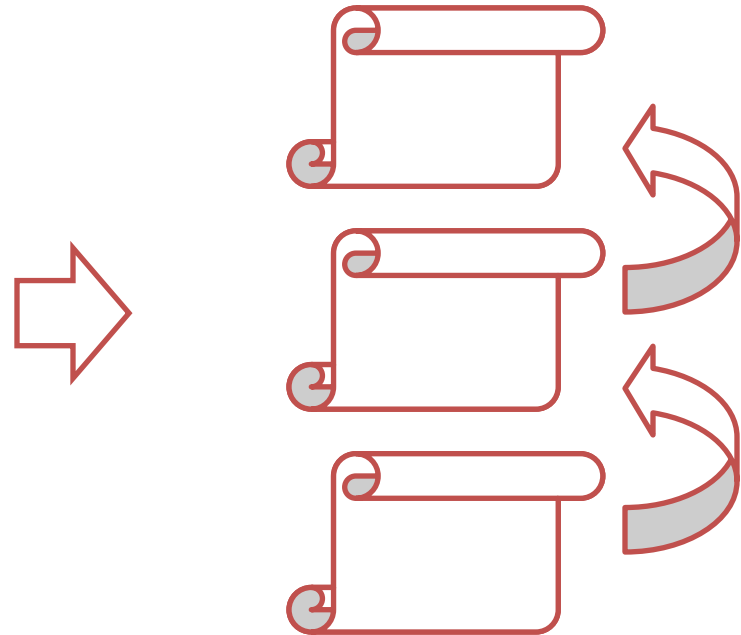
Title
(1)
Mars



Rhetorical Structure Theory



Cue-Phrase-based



(Typical) Discourse Parsing

Approach: SVM, Naïve-Bayes, ...

Features: Cue phrases, word-pairs, ...

Explicit : **Easier** (MacroF1 ~80%)

Implicit : **Harder** (MacroF1 ~50%)

(Typical) Discourse Parsing

Overview

Discourse Segmentation

Discourse Parsing

~~Reference Resolution~~

TOPICS

Cohesion vs Coherence?

Equivalent in meaning?

VS

Collectively form a discourse?

Challenge