Natural Language Processing (CS5803)

Lecture 11
(Discourse)

Coherence

- John took a train from Paris to Istanbul. He likes spinach

Jane took a train from Paris to Istanbul. She had to attend a conference.

Structured relations (e.g. reason) – Coherence relation

- John wanted to buy a piano for his living room.
- Jenny also wanted to buy a piano.
- He went to the piano store.
- It was nearby.
- The living room was on the second floor.
- She didn’t find anything she liked.
- The piano he bought was hard to get up to that floor

- John wanted to buy a piano for his living room.
- He went to the piano store.
- It was nearby.
- The living room was on the second floor.
- The piano he bought was hard to get up to that floor
- Jenny also wanted to buy a piano.
- She didn’t find anything she liked.

Focuses on salient entities – does not change focus back and forth between multiple entities

Attention region in NLP

- Inside a sentence
  - NER, POS
  - Parse tree
- Across the sentences?
  - Coherence
- What is coherence?

Discourse

- What all aspects are covered in the study of discourse?
- Consists of collocated, structured, coherent groups of sentences
  - What makes something a discourse as opposed to a set of unrelated sentences?
  - How can text be structured (related)?

- Coherence – central theme - models the logical flow of the discourse

Intrinsic features of discourse

Position, order, adjacency and **context**

- **Position**: opening sentence , Ending sentence.
- **Order**: different orders lead to various events/meaning
    - I said the magic words, and a genie appeared.

      vs.
    - A genie appeared, and I said the magic words.
- **Adjacency**: attributed material and contrasts are visible through sentences nearby
- **Context** : intended meaning can only be conveyed when understood in context.

Coherence

- Coherence as the main characteristic of discourse
- How do you recognize discourse?
  - It makes sense!
  - It is relevant!
  - It ‘hangs together’
  - "It is coherent!! !
- Coherence is a property of a well-written text, coherence detection plays a part in any task that requires measuring the quality of a text
  - Essay grading, summarization, detecting mental health

Coherence

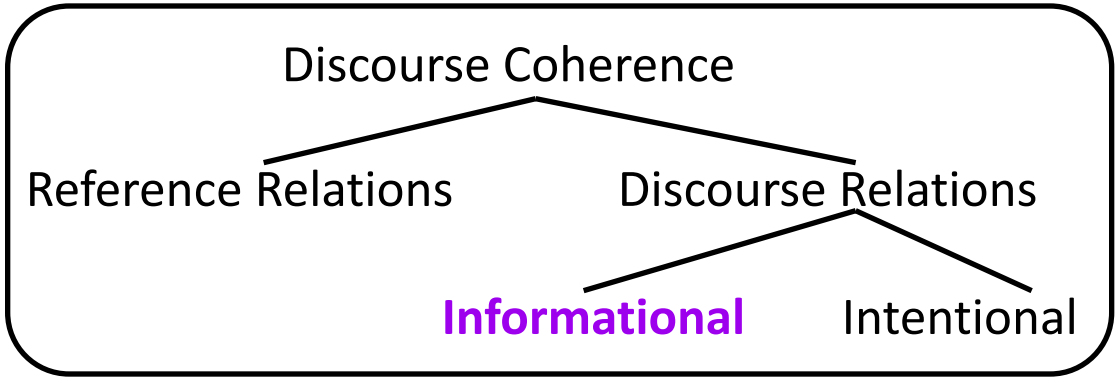
- Coherence is a property humans use to **evaluate text quality.**
- A coherent discourse must have **meaningful connections (i.e. coherence relations)** between its utterances.

Ram got caught in the rain.

He fell ill.

Coherence

- The meaning and coherence of a discourse results partly from how its constituents relate to each other.
  - Reference relations – co-reference or Anaphora Resolution - determining which entity a referring expression refers to
  - Discourse relations



Discourse Relations and Connectives

Jewelry displays in department stores were often cluttered and uninspired. And the merchandise was, well, fake. **As a result, marketers of faux gems steadily lost space in department stores to more fashionable rivals—cosmetics makers.**

In July, the Environmental Protection Agency imposed a gradual ban on virtually all uses of asbestos. (implicit=as a result) **By 1997, almost all remaining uses of cancer-causing asbestos will be outlawed.**

Discourse Relations and Connectives

Class	Type	Example
TEMPORAL	SYNCHRONOUS	The parishioners of St. Michael and All Angels stop to chat at the church door, as members here always have. (Implicit while) In the tower, five men and women pull rhythmically on ropes attached to the same five bells that first sounded here in 1614.
CONTINGENCY	REASON	Also unlike Mr. Ruder, Mr. Breeden appears to be in a position to get somewhere with his agenda. (implicit=because) **As a former White House aide who worked closely with Congress, he is savvy in the ways of Washington.**
COMPARISON	CONTRAST	The U.S. wants the removal of what it perceives as barriers to investment; Japan denies there are real barriers.
EXPANSION	CONJUNCTION	Not only do the actors stand outside their characters and make it clear they are at odds with them, but they often literally stand on their heads.

Discourse relations

- Explicit vs. implicit
  - ✔ I stayed home because I had a headache
    - Signaled by a discourse connective
    - Inferred without the presence of a connective
  - ✔ I took my umbrella. [Because] The forecast was for rain in the afternoon.

Coherence

- Coherent text is characterized by the presence of various types of cohesive links that facilitate text comprehension

- Reference and lexical reiteration
  - Pronouns, definite descriptions, semantically related words
- Discourse relations (conjunction)

✔ I closed the window because it started raining.

Referential coherence

- Centering theory
  - tracking focus of attention across adjacent sentences
- Syntactic form of references
  - Particularly first and subsequent mention
- Lexical chains
  - Identifying and tracking topics within a text

Lexical chains

- Often discussed as cohesion indicator, implemented systems, but not used in text quality tasks
  - Find all words that refer to the same topic
  - Find the correct sense of the words
- “Before winter I built a chimney, and shingled the sides of my house. I have thus a tight shingled and plastered house. with a garret and a closet, a large window on each side.”
- LC = {}?

- May need to use knowledge structure on top of common sense knowledge

Centering theory ingredients

(Grosz et al, 1995)

- Deals with local coherence
  - What happens to the flow from sentence to sentence
  - Does not deal with global structuring of the text (paragraphs/segments)

- Defines coherence as an estimate of the processing load required to "understand" the text

Nouns Tend to Repeat

Discourse, the study of how the meaning of a document is built out the meaning's of its sentences, is the inter-sentential analogue of semantics. In this talk we consider the following abstract problem in discourse. Given a document, randomly permute the order of the sentences and then attempt to distinguish the original from the permuted version. We present a sequence of generative models that can handle the problem with increasing accuracy. Each model accounts for some aspect of the document, and assigns a probability to the document's contents. In the standard generative way the subsequent models simply multiply individual probabilities to get their results. We also discuss the linkage of this abstract task to more realistic ones such as essay grading, document summarization and document generation.

A (Partial) Entity Grid

Discourse	S	X	-	-	-	-	-	
Meaning	X	-	-	-	-	-	-	
Document		X	-	X	-	X	-	-
Sentences		X	-	X	-	-	-	-
Talk	-	X	-	-	-	-	-	
Problem	-	O	-	O	-	-	-	
Order		-	-	O	-	-	-	-
Original	-	-	X	-	-	-	-	-
Version	-	-	X	-	-	-	-	-
Models		-	-	-	X	-	S	-

Transitions are mostly “smooth”.

The Grid for the Randomized Document

Discourse	-	-	-	-	X	-	S
Meaning-	-	-	-	-	-	X	
Document	-	X	X	-	-	-	X
Sentences	-	-	X	-	-	-	X
Talk	-	-	-	-	X	-	-
Problem	O	-	-	-	O	-	-
Order		-	-	O	-	-	-	-
Original		-	-	X	-	-	-	-
Version		-	-	X	-	-	-	-
Models		X	-	-	S	-	.	-

Transitions are “abrupt”.

Entity Grid: Another example

1 [The Justice Department]$_S$ is conducting an [anti-trust trial]$_O$ against [Microsoft Corp.]$_X$ with [evidence]$_X$ that [the company]$_S$ is increasingly attempting to crush [competitors]$_O$.
2 [Microsoft]$_O$ is accused of trying to forcefully buy into [markets]$_X$ where [its own products]$_S$ are not competitive enough to unseat [established brands]$_O$.
3 [The case]$_S$ revolves around [evidence]$_O$ of [Microsoft]$_S$ aggressively pressuring [Netscape]$_O$ into merging [browser software]$_O$.
4 [Microsoft]$_S$ claims [its tactics]$_S$ are commonplace and good economically.
5 [The government]$_S$ may file [a civil suit]$_O$ ruling that [conspiracy]$_S$ to curb [competition]$_O$ through [collusion]$_X$ is [a violation of the Sherman Act]$_O$.
6 [Microsoft]$_S$ continues to show [increased earnings]$_O$ despite [the trial]$_X$.

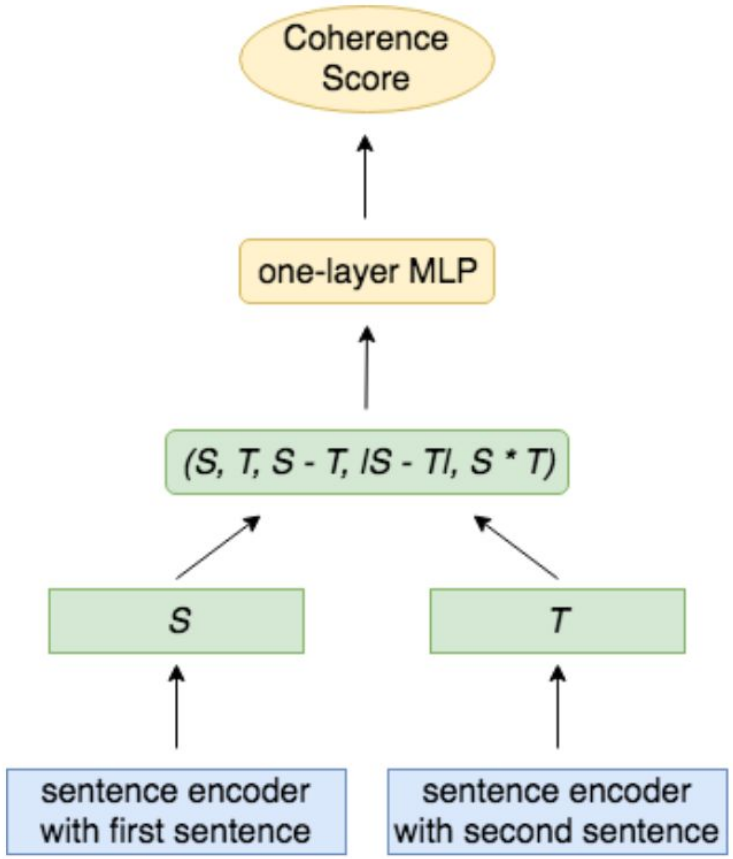
	Department	Trial	Microsoft	Evidence	Competitors	Markets	Products	Brands	Case	Netscape	Software	Tactics	Government	Suit	Earnings	
1	S	O	S	X	O	–	–	–	–	–	–	–	–	–	–	1
2	–	–	O	–	–	X	S	O	–	–	–	–	–	–	–	2
3	–	–	S	O	–	–	–	–	S	O	O	–	–	–	–	3
4	–	–	S	–	–	–	–	–	–	–	–	S	–	–	–	4
5	–	–	–	–	–	–	–	–	–	–	–	–	S	O	–	5
6	–	X	S	–	–	–	–	–	–	–	–	–	–	–	O	6

How can it help in measuring/detecting coherence in an automated manner?

Measuring/Detecting Local Coherence - Embedding based method

- Represent a sentence using sum of the LSA vectors for its words
- Take two sentences s_1, s_2, and get their vectors v_1, v_2
- Their similarity $sim(s_1, s_2)$ is defined by $\cos(v_1, v_2)$
- Coherence of the text is the average similarity over consecutive sentence pair, i.e. $c(T) = \frac{1}{n-1} \sum_{i=1}^{n-1} sim(s_i, s_{i+1})$
- Extensions/ modifications?

Local Coherence Discriminator



$$\mathcal{L}(\theta) = \sum_{d \in \mathcal{C}} \sum_{s_i \in d} \mathop{\mathbb{E}}_{p(s'|s_i)} \left[L(f_\theta(s_i, s_{i+1}), f_\theta(s_i, s')) \right]$$

- Loss in contrastive in nature
- Negative examples sampled uniformly
- Evaluated for discrimination, insertion and reconstruction tasks
- Simple yet effective model

Coherence evaluation

- Manually annotated corpus
- Self supervision
  - Sentence order discrimination
  - Sentence insertion
  - Sentence order reconstruction

Components

- Units of discourse
  - Texts can be segmented into minimal units, or spans
- Nuclearity
  - Some spans are more central to the text's purpose (nuclei), whereas others are secondary (satellites)
  - Based on hypotactic and paratactic relations in language
- Relations among spans
  - Spans are joined into discourse relations
- Hierarchy/recursion
  - Spans that are in a discourse relation may enter into new relations

Relations in discourse

Reason: The nucleus is an action carried out by an animate agent and the satellite is the reason for the nucleus.

(22.7) [NUC Jane took a train from Paris to Istanbul.] [SAT She had to attend a conference.]

Elaboration: The satellite gives additional information or detail about the situation presented in the nucleus.

(22.8) [NUC Dorothy was from Kansas.] [SAT She lived in the midst of the great Kansas prairies.]

Evidence: The satellite gives additional information or detail about the situation presented in the nucleus. The information is presented with the goal of convince the reader to accept the information presented in the nucleus.

(22.9) [NUC Kevin must be here.] [SAT His car is parked outside.]

Relations in discourse

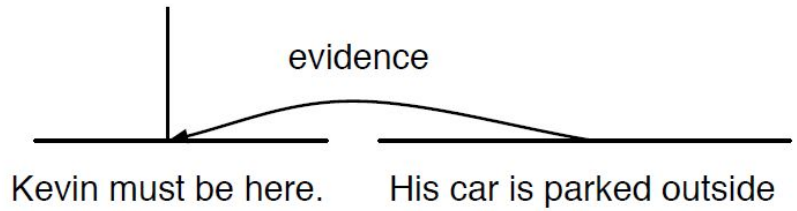
Attribution: The satellite gives the source of attribution for an instance of reported speech in the nucleus.

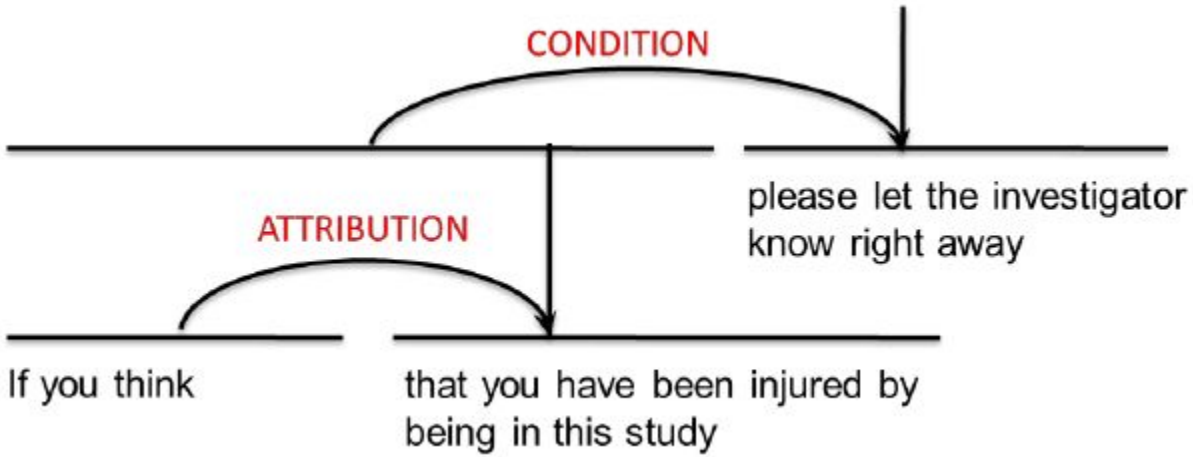
(22.10) [$_{\text{SAT}}$ Analysts estimated] [$_{\text{NUC}}$ that sales at U.S. stores declined in the quarter, too]

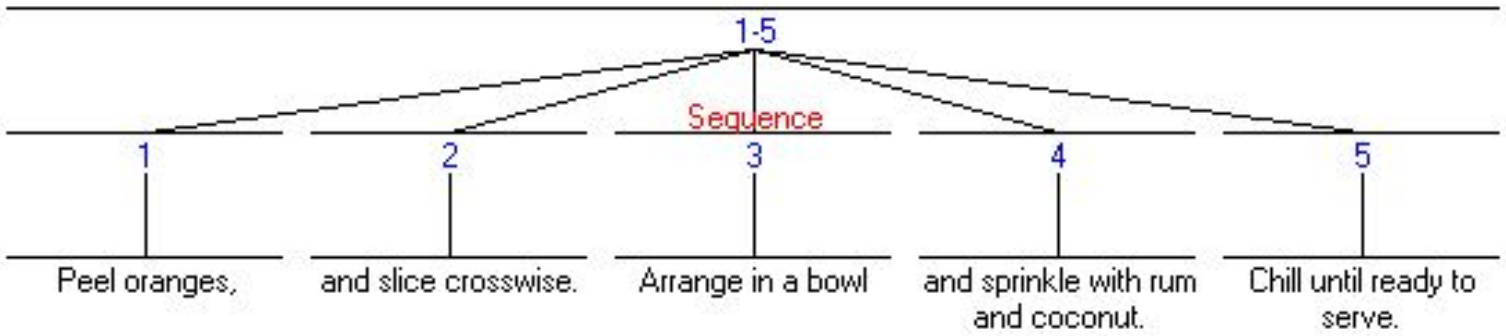
List: In this multinuclear relation, a series of nuclei is given, without contrast or explicit comparison:

(22.11) [$_{\text{NUC}}$ Billy Bones was the mate;] [$_{\text{NUC}}$ Long John, he was quartermaster]

Relations in discourse and hierarchy







RST

The text:

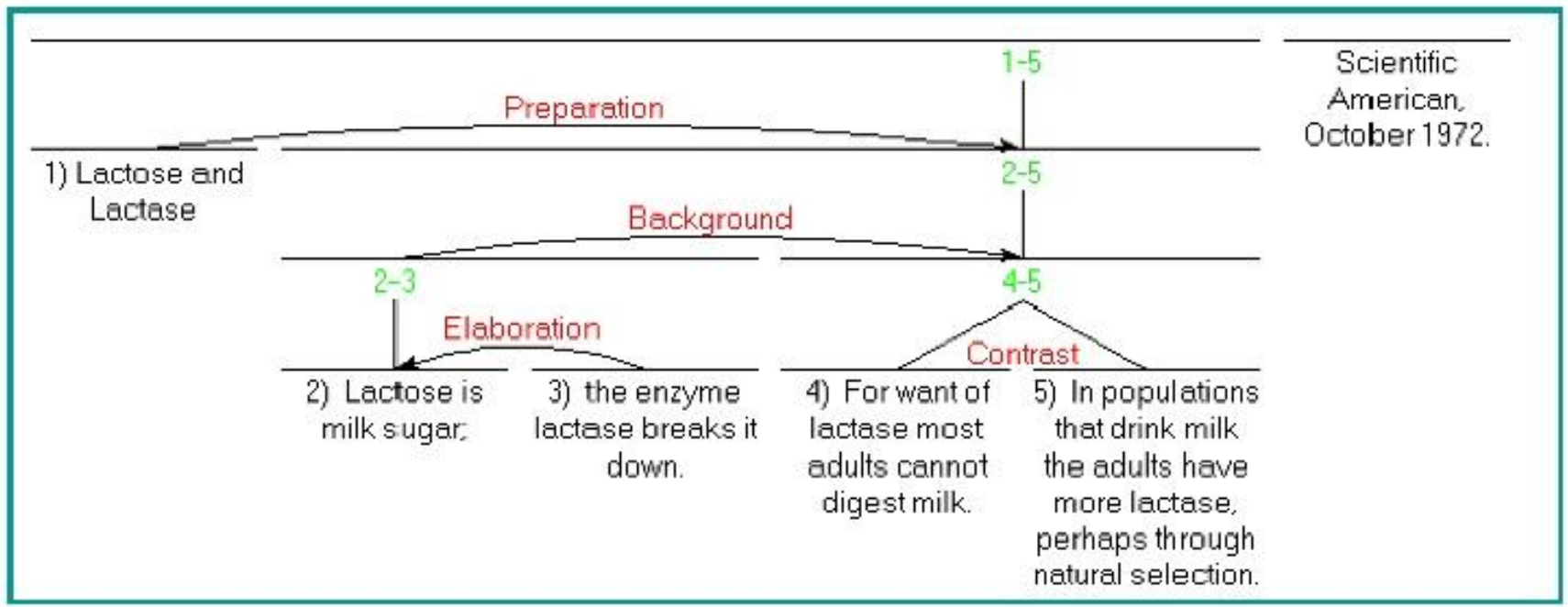
1) Lactose and Lactase

2) Lactose is milk sugar; 3) the enzyme lactase breaks it down.

4) For want of lactase most adults cannot digest milk.

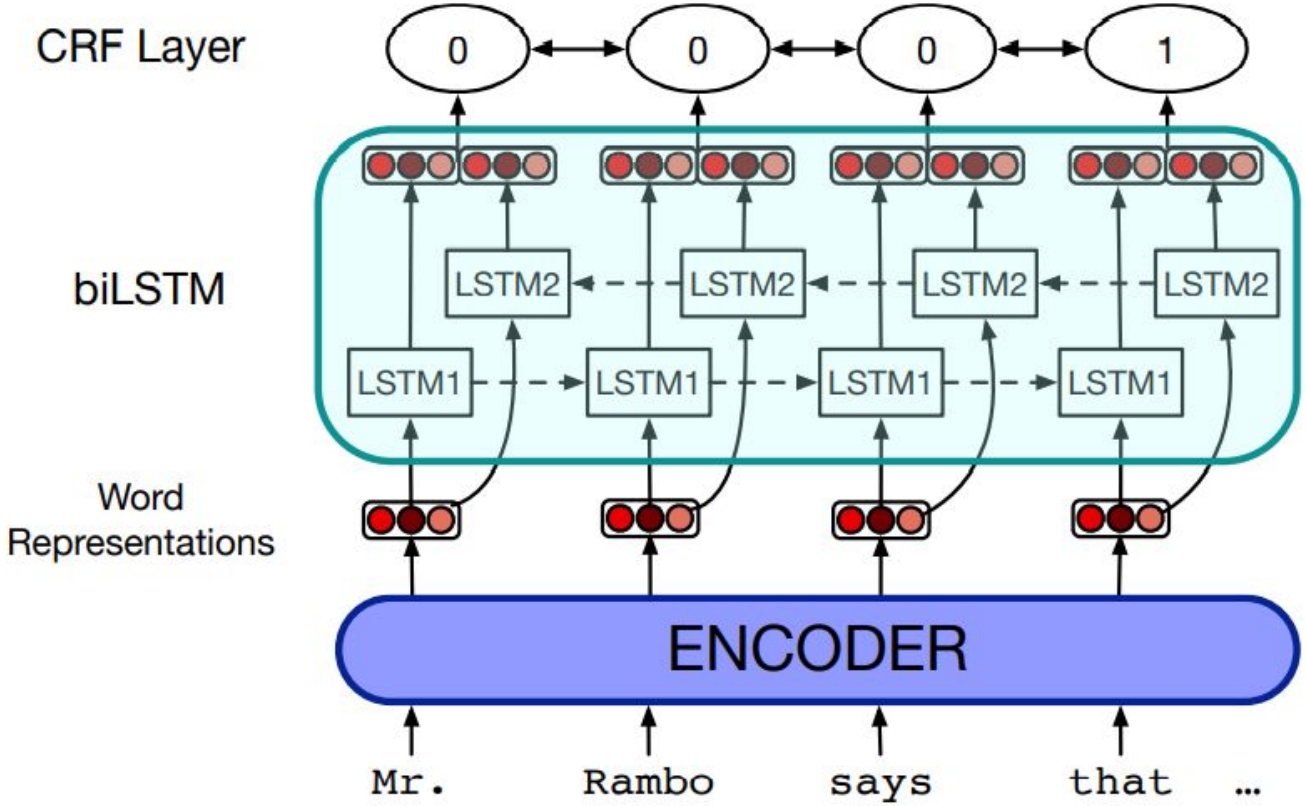
5) In populations that drink milk the adults have more lactase, perhaps through natural selection.

6) Norman Kretchmer, Scientific American, page 70, October 1972.



Elementary Discourse Units

- [Mr. Rambo says]e1 [that a 3.2-acre property]e2 [overlooking the San Fernando Valley]e3 [is priced at $4 million]e4 [because the late actor Erroll Flynn once lived there.]e5



RST Parsing

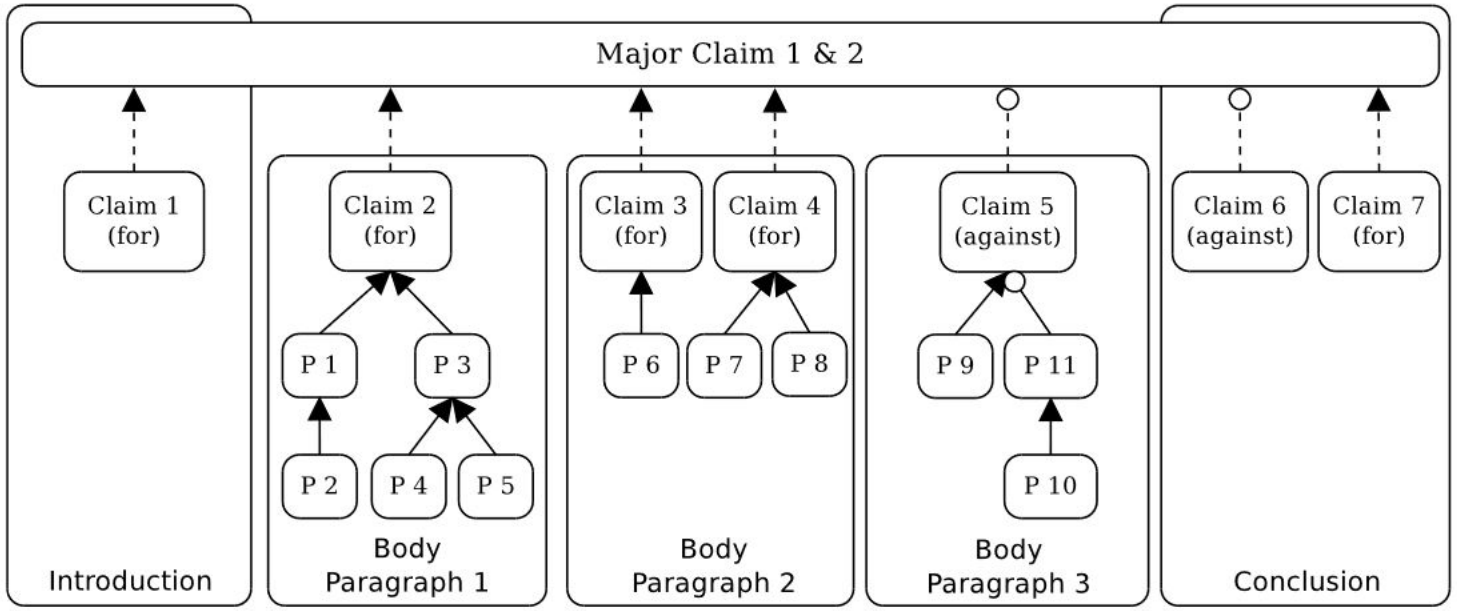
- Create the RST from the input text
- Can be done through Shift Reduce parsing
- Maintains a stack
- Items in stack get connected using labels, determined by actions
- Identifies EDUs, and represents them using hierarchical BiLSTM
- An MLP determines the action by considering top three trees from the stack and the first EDU in the queue
- Ground-truth RST is converted to sequence of actions to train (and evaluate) the system

Global Coherence

- A discourse must adhere globally
- Consider: Story, scientific articles, …
- Possible routes: Global Coherence Structures
- **Argumentation structure:**
  - Multiple work on logical structure of the document
  - Identifies relations between sentences
  - Identifies claims, premises
  - Identifies argumentative relations: SUPPORT/ATTACK

Argumentation Mining: An example

"(1) **Museums and art galleries provide a better understanding about arts than Internet.** (2) In most museums and art galleries, detailed descriptions in terms of the background, history and author are provided. (3) Seeing an artwork online is not the same as watching it with our own eyes, as (4) the picture online does not show the texture or three-dimensional structure of the art, which is important to study."



- Figure from Stab, C. and Gurevych, I. (2017). [Parsing argumentation structures in persuasive essays. Computational Linguistics](), 43(3), 619–659

Further Reading

- Tanner Bohn, Yining Hu, Jinhang Zhang, and Charles Ling. 2019. Learning Sentence Embeddings for Coherence Modelling and Beyond. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*
- Logeswaran, L., Lee, H., & Radev, D. Sentence ordering and coherence modeling using recurrent neural networks. AAAI 2018.

- Naismith, B., Mulcaire, P., & Burstein, J. Automated evaluation of written discourse coherence using GPT-4. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)* (pp. 394-403).