

Discourse Processing



Presented By: Sabita Rajbanshi
(MSIISE16)

September 18, 2022

Outline

- Motivation
- Introduction
- Cohesion
- Co-reference
- Coherence
- Centering Theory
- Entity Based Coherence
- Global Coherence
- Applications
- Future Challenges
- Reference

Motivation

Why Discourse Processing?

- To develop tools to automatically model language phenomena of discourse.
- The analysis and interpretability beyond the sentence-level has become more flexible due to neural network.
- Handle many complex applications such as machine translation, text categorization, sentiment analysis.

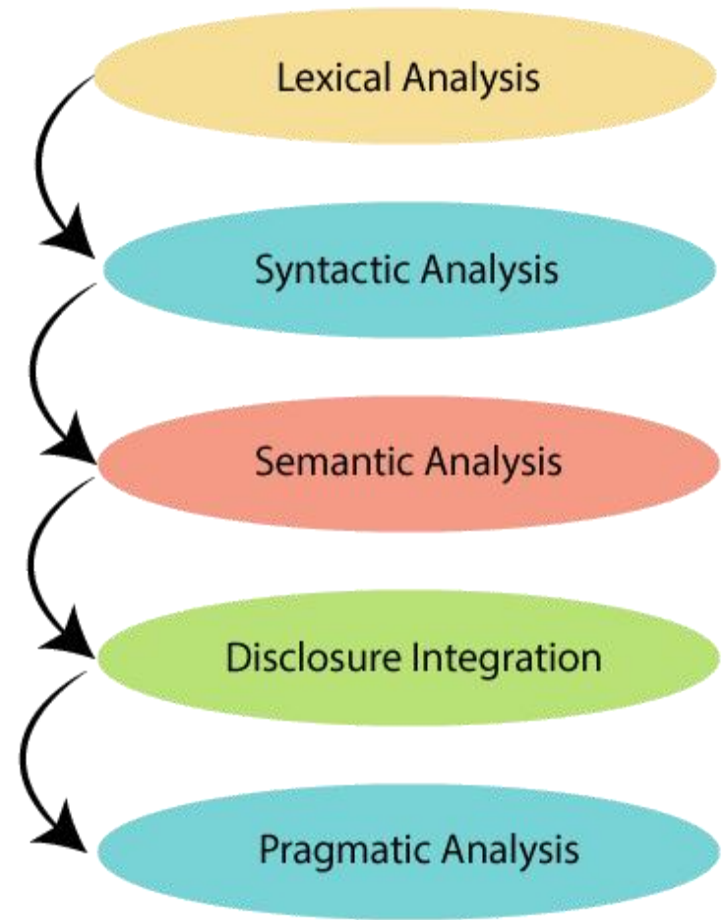


Fig.1: Stages of NLP

Introduction [1]

- The word discourse comes from latin discursus which denotes conversation or speech.
- In linguistics, collection of interrelated sentences is called discourse.
- For a set of sentences to make sense, it must consists of sentences that are related to each other.
- Discourse processing is the nlp tasks that deals with cohesion structure, co-reference structure, and the coherence structure.
- Discourse structure parsing : Given a sequence of text, automatically determine the coherence relations between spans within it.

Introduction [2]

Features of discourse :

- **Position:** opening sentence , Ending sentence.
- **Order:** different orders lead to various events/meaning

Example:

A : I said the magic words, and a genie appeared.

B : 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.

Introduction [3]

Discourse Processing :

- Investigates the structures, patterns, mental representations and processes that underlie written and spoken discourse.
- Mission of the field is to improve the production of discourse in textbooks, tutoring sessions, classroom, computer based training, etc.
- Five levels of Discourse Processing :
 - i. Surface code
 - ii. Textbase
 - iii. Situation Model
 - iv. Pragmatic Communication
 - v. Discourse genre

Discourse Cohesion [1]

- Grammatical relationship between parts of a sentence essential for its interpretation.
- Structural integrity of a text that means links and ties that exists within the text.
- Logical connections between the words and sentence of any text through various connectors.
- **Two types of cohesion:**
 - A. Grammatical cohesion : Expressed through the grammatical system of a language such as references, substitution, ellipsis, and conjunction.
 - B. Lexical cohesion : The semantically related vocabulary such as repetition, synonym, or general word.

Discourse Cohesion [2]

Example:

Grammatical Cohesion:

- i. Wow, how beautiful flower vessel! How much does it cost? - [reference]
- ii. He passed the exam. However, he did not obtain A plus.
- [conjunction]

Lexical Cohesion:

- i. I have a puppy. The puppy is black. - [Repetition]
- ii. I have a puppy. The pup is black. - [Synonym]

Discourse Coreference

- The coreference resolution is to identify all noun phrases (mentions) that refer to the same entity.
- Applications :
 - Text Summarization
 - Machine Translation
 - Information Extraction
 - Information Retrieval, Chatbots
- Ways to build a reference resolution system either by rule based or supervised algorithms.
- Example :
 - Barack Obama, the U.S former president tweeted his prayers for Queen Elizabeth II.
 - Despite her difficulty, Ana went ahead to help him.

Discourse Coherence [1]

- Coherence (connectedness) in linguistics is a logical connection in the meaning of a text.
- The order of statements relates one another by sense.
- **Examples:**

A text with coherence

A : Did you bring the car?

B : Yes, I brought it yesterday.

[Note: Question and answers are logically connected by it]

A text with no coherence

A : Where did you go last week?

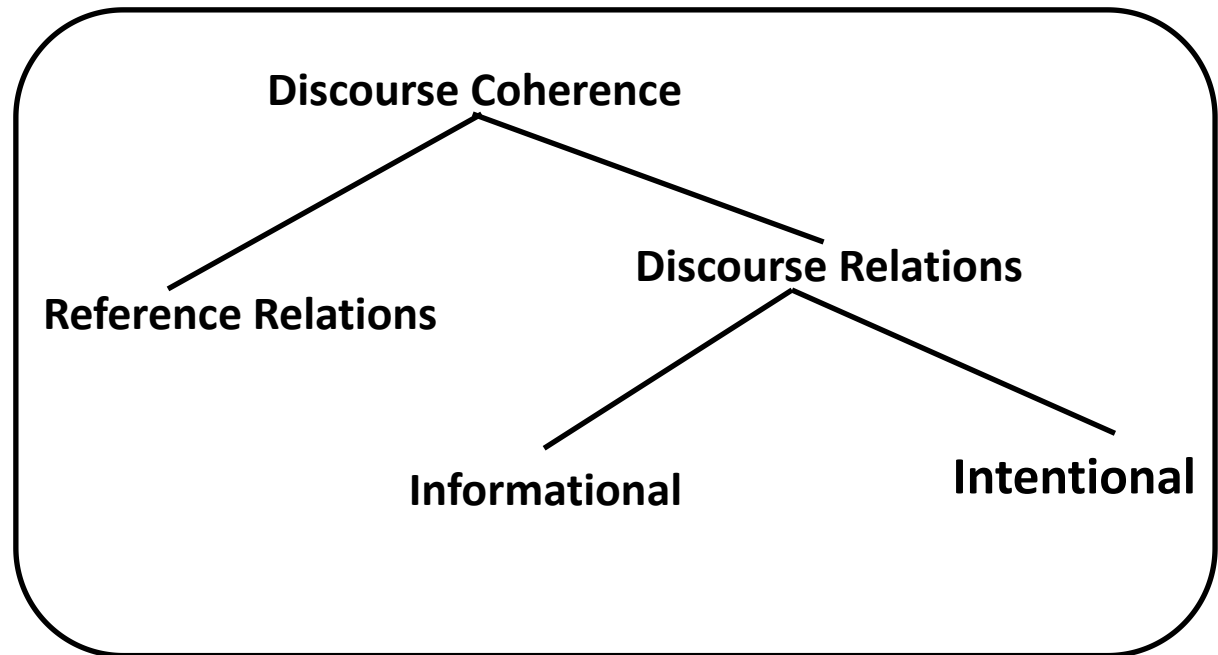
B : That sounds good. My sister paints it.

[Note: There is no meaningful connections between Ques & Ans]

Discourse Coherence [2]

- **Reference relations :**

- Reference is a means to link a referring expression to another referring expression in the surrounding text
- Example: Riya brought a printer. It costs her Rs. 20,000.



Discourse Coherence [3]

- Discourse relations :
 - **Informational** (or semantic) discourse relations holds the facts, beliefs, events etc.
 - To find the best explanation of why it would be true.
 - **Intentional** discourse relations specify the relevant purposes of discourse as well as relationships among them .
 - To find the best explanation of why it was said
- Example:

Which boutique designs with the low price ?
How many custom designs do you want ?

Centering Theory-[1]

- The most influential theory of entity-based coherence.
- A theory of both discourse salience and discourse coherence.
- Centering means that at any given point in the discourse one of the entities in the discourse model is salient : it is being centered.
- Discourses in which adjacent sentences continue to maintain the same salient entity are more coherent than those which shift back and forth between multiple entities.

Centering Theory-[2]

The two discourses have same propositional content but different saliences :

Discourse 1 :

- Anila was an assistant professor at TU.
- She taught a class there called Natural Language Processing.
- She enjoyed teaching the class, because she liked NLP a lot.
- She was planning to teach the class once per year.

Discourse 2 :

- Anila was an assistant professor at TU.
- TU had a class that she taught called Natural Language Processing.
- She enjoyed teaching the class, because she liked NLP a lot.
- The plan was that the class would be taught by Anila once per year.

Centering Theory-[3]

- Maintains two representations for each utterance (U_n)
 - $C_b(U_n)$: Backward-looking center of U_n
 - Salient entity being focused on in the discourse after U_n is interpreted
 - $C_f(U_n)$: Forward-looking centers of U_n
 - Set of potential future salient entities (potential $C_b(U_{n+1})$)
- Set of $C_f(U_n)$ are ranked based on a variety of factors (i.e; grammatical role)
- Highest-ranked $C_f(U_n)$ is the preferred center C_p

Centering Theory-[4]

Four possible relationships between U_n and U_{n+1}

Relationships depend on $C_b(U_{n+1})$, $C_b(U_n)$, and $C_p(U_{n+1})$

	$C_b(U_{n+1}) = C_b(U_n)$ (or undefined)	$C_b(U_{n+1}) \neq C_b(U_n)$
$C_b(U_{n+1}) = C_p(U_{n+1})$	Continue	Smooth-Shift
$C_b(U_{n+1}) \neq C_p(U_{n+1})$	Retain	Rough-Shift

Figure 2: Centering Matrix

Centering Theory-[5]

Based on these relationships, we can define two rules:

Rule 1: The centered entities should be pronouns

Rule 2: Transition states are ordered such that Continue > Retain > Smooth-Shift > Rough-Shift.

Centering Theory-[6]

Discourse 1 :

- **Anila** was an assistant professor at **TU**.
- **She** taught a **class** there called Natural Language Processing.

$Cf(U1) : \{Anila, TU\}$

$Cp(U1) : Anila$

$Cb(U1) : Undefined$

$Cf(U2) : \{Anila, TU, class\}$

$Cp(U2) : Anila$

$Cb(U2) : Anila$

Discourse 2 :

- **Anila** was an assistant professor at **TU**.
- **TU** had a **class** that **she** taught called Natural Language Processing.

$Cf(U1) : \{Anila, TU\}$

$Cp(U1) : Anila$

$Cb(U1) : Undefined$

$Cf(U2) : \{TU, class, Anila\}$

$Cp(U2) : TU$

$Cb(U2) : Anila$

Entity Based Coherence [1]

- The entity grid model is a way to capture entity based coherence.
- Uses machine learning to induce patterns of entity mentioning that make a discourse more coherent.
- Based on an entity grid
 - 2d array representing the distribution of entity mentions across sentences
 - Rows = sentences
 - Columns = discourse entities
 - Values in cells = Whether the entity appears in the sentence, and its grammatical role (subject, object, neither, absent)

Entity Based Coherence [2]

Discourse: contains four sentences

- [Anila]s was an assistant professor at [TU]x.
- [Anila]s taught a [class]o at [TU]x called Natural Language Processing.
- [Anila]s enjoyed teaching the [class]x because [Anila]s liked [NLP]o a lot.
- [Anila]s was planning to teach the [class]x once per year.

Entity Based Coherence [3]

	Anila	TU	class	NLP
S1	S	X	-	-
S2	S	X	O	-
S3	S	-	X	O
S4	S	-	X	-

Figure 3: Entity grid model

Entity Based Coherence [4]

- Dense columns indicate entities mentioned often (ie; Anila)
- Sparse columns indicate entities mentioned rarely (ie; NLP)
- Coherence is thus measured by patterns of local entity transition
- Each transition ends up with a probability

Probability Transition :

- col2 : $p([x, x, -, -]) = 1/4$
- col3 & col4 : $p([- , o]) = 2/12 = 1/6$

Entity Based Coherence [5]

These transitions and their probabilities can be used as features for a machine learning model, trained to predict coherence scores.

These models can be trained in a self-supervised manner:

- Learns to distinguish the natural order of sentences in a discourse (expected to be coherent) from a modified order.

Global Coherence -[1]

Coherence relations and entity salience focus on local coherence.

However, discourse must be globally coherent as well.

For example;

- Good stories should have overall narrative structure

- Essays should follow specific argument structure

- Scientific papers are characterized by a structure common across research publications

An area of global coherence that has particularly received strong attention is argumentation structure.

Specifically, an active research problem is argumentation mining

Global Coherence -[2]

At high level these rhetorical elements correspond to claims and premises.

Persuasive arguments generally contain well-defined argumentative components:

Claim : The central, controversial, component of the argument

Premise : A persuasive support or attack of the claim or another premise

Global Coherence -[3]

Example : Argumentation Structure

MSIISE is the best program at Thapathali campus. It covers a very exciting subject : natural language processing. It also offers both the structure of a lecture based class and the flexibility of a seminar course. This mix is nice because you can learn fundamental principles but also get up to speed on contemporary research.

❑ How can we detect argumentation structure?

Classifiers to identify claims, premises, and non-argumentation

Methods to detect specific argumentation schemes:

Argument from example

Argument from cause to effect

Argument from consequences

Applications of Discourse

- Summarization
- Generation
- Sentiment Analysis
- Machine Translation

Future Challenges

1. Learning from limited annotated data
2. Language and domain transfer
3. Discourse generation
4. New emerging applications

Reference

- [1] Barzilay, R. and M. Lapata. 2005. Modeling local coherence: An entity-based approach. ACL.
- [2] Barzilay, R. and M. Lapata. 2008. Modeling local coherence: An entity-based approach. Computational Linguistics, 34(1):1–34.
- [3] Habernal, I. and I. Gurevych. 2017. Argumentation mining in user-generated web discourse. Computational Linguistics, 43(1):125–179.
- [4] Kehler, A. 1994. Temporal relations: Reference or discourse coherence? ACL.
- [5] Kehler, A. 2000. Coherence, Reference, and the Theory of Grammar. CSLI Publications.

Thank You for Listening.