Coherence in Discourse

Topic-wise Notes: Definition | Characteristics | Architecture | Step-by-Step Example | Applications

Topic	Title
1	Local vs Global Coherence
2	Coherence Relations
3	Rhetorical Structure Theory (RST)
4	Temporal and Causal Relations
5	Summary table and exam questions

Background: Coherence is the meaningful connectedness of a discourse: its sentences and larger parts relate to each other (by cause, contrast, elaboration, time order, shared entities) so that the text forms one understandable message. **Cohesion** is only the surface linking (pronouns, repeated words, conjunctions); coherence is deeper and relies on reader inference and world knowledge. Example: 'A: The phone is ringing. B: I am in the shower.' has no shared words yet is coherent.

Topic 1: Local vs Global Coherence

1.1 Definition

Local coherence: the connectedness between adjacent or nearby sentences/utterances, i.e., whether sentence *i* follows smoothly from sentence *i-1* (through entity continuity, pronouns, connectives, lexical overlap).

Global coherence: the organisation of the whole discourse into a unified, purposeful structure, i.e., whether topics, sections and rhetorical moves add up to one main point (topic progression, discourse structure, genre schema).

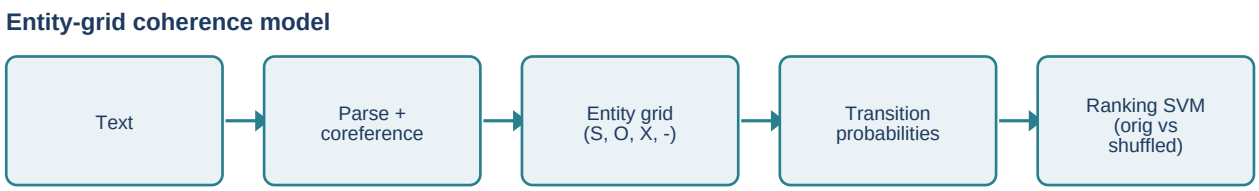
1.2 Characteristics

Feature	Local coherence	Global coherence
Scope	Adjacent sentences / short windows	Whole paragraph, section or document
Main devices	Anaphora, entity continuity, connectives, lexical chains, adjacency pairs	Topic structure, rhetorical organisation, genre schema (intro-body-conclusion), communicative goal
Theories	Centering Theory (Grosz, Joshi, Weinstein 1995); Entity Grid (Barzilay and Lapata 2008)	RST (Mann and Thompson 1988); Grosz and Sidner intentional structure; topic segmentation
Typical failure	Dangling pronoun, abrupt jump	Topic drift, missing conclusion, jumbled sections
Granularity	Sentence pairs	Segments, sections, trees
Cost	Cheap, window based	Heavier, hierarchical or tree models
Evaluation	Local sentence-order discrimination, insertion task	Document-level ranking, essay organisation scores

- Coherence is **graded** (more or less), **reader dependent**, and **hierarchical** (local ties build global structure).
- A text can be locally coherent but globally incoherent (topic drift), and rarely the reverse.
- **Centering transitions** (best to worst): Continue > Retain > Smooth-shift > Rough-shift.

1.3 Architecture

(a) Entity-grid model (local):



(b) Neural local + global model:

Neural hierarchical coherence model



Training objectives: sentence-order prediction, pairwise ranking of original vs shuffled documents with a margin loss, next-sentence prediction. Local scores from the window model are aggregated by the document encoder into a global score.

1.4 Example, step by step

Part A - Local coherence (Entity grid and Centering). Text: S1: Ravi bought a laptop. S2: It was expensive. S3: He returned it the next day.

Step 1 - Resolve coreference: It (S2) = laptop; He (S3) = Ravi; it (S3) = laptop. **Step 2 - List entities:** Ravi, laptop, day. **Step 3 - Fill grid** (S = subject, O = object, X = other, - = absent):

Entity	S1	S2	S3
Ravi	S	-	S
laptop	O	S	O
day	-	-	X

Step 4 - Count transitions (length 2, 6 in total): S-, -S, OS, SO, --, -X, each 1/6 = 0.167. The key entities Ravi and laptop keep returning in subject/object positions, a sign of good local coherence. **Step 5 - Centering:**

Utterance	Cf (ranked)	Cb	Cp	Transition
U1: Ravi bought a laptop.	Ravi, laptop	undefined	Ravi	- (first)
U2: It was expensive.	laptop	laptop	laptop	Continue
U3: He returned it the next day.	Ravi, laptop, day	laptop	Ravi	Retain

Step 6 - Interpret: discourse stays about the laptop (Continue then Retain). If U3 were 'The weather in Pune was hot', Cb would be undefined (a rough break) and the score drops.

Part B - Global coherence. Text: 'I love cooking. Cooking needs heat. Heat waves are rising. Rising seas threaten cities. Cities need metros.'

- **Step 1 - Local scores:** every adjacent pair shares a word, so local coherence is high.
- **Step 2 - Topic trajectory:** cooking, heat, climate, sea level, cities, transport: the topic drifts.
- **Step 3 - Global test:** compare first and last sentences (little similarity) and look for a main point or nucleus (none); an RST tree would have no clear central span.
- **Step 4 - Verdict:** locally coherent but **globally incoherent**.

1.5 Applications

- **Automated essay scoring:** organisation and flow of student essays.
- **Evaluating generated text:** detect incoherent or drifting output from chatbots and LLMs.
- **Sentence ordering and text repair:** re-arrange jumbled sentences; choose position for an inserted sentence.
- **Summarisation:** ensure the extracted sentences read smoothly.
- **Machine translation:** keep pronoun and entity continuity across sentences.
- **Readability and writing assistants:** flag abrupt shifts and dangling pronouns.

Topic 2: Coherence Relations

2.1 Definition

Coherence relation (discourse or rhetorical relation): a semantic or pragmatic link between two discourse units (clauses, sentences or larger spans) that explains *why the second unit follows the first*, e.g., Cause, Explanation, Elaboration, Contrast, Condition, Sequence.

2.2 Characteristics

- **Binary:** relate two arguments (Arg1 and Arg2).
- **Directional or symmetric:** Explanation is directional; Contrast and Parallel are symmetric.
- **Explicit or implicit:** signalled by a connective ('because') or inferred without one.
- **Semantic vs pragmatic:** semantic links hold between facts ('It rained so the road is wet'); pragmatic between speech acts ('It is cold, so close the window').
- **Ambiguous:** one connective may signal several relations ('since' = temporal or causal); a pair may carry two relations.
- **Theory dependent:** Hobbs (inference-based), RST (nucleus/satellite), PDTB (connective-anchored), Kehler (Cause-Effect, Resemblance, Contiguity).

Relation	Meaning	Connectives	Example
Result / Cause	Second is effect of first	so, therefore, as a result	'It rained, so the match was cancelled.'
Explanation	Second explains first	because, since	'He was late because the bus broke down.'
Sequence / Occasion	Events in order	then, after that	'She wrote the code, then tested it.'
Elaboration	Adds detail	in particular, specifically	'The phone is fast. It has an 8-core chip.'
Contrast	Differences	but, however, while	'Ram likes tea, but Sita prefers coffee.'
Concession	Expected result denied	although, yet	'Although it rained, we played.'
Condition	Enables if true	if, unless	'If you study, you will pass.'
Parallel	Similar structure	similarly, likewise	'Tom runs daily. Meera swims daily.'
Background	Gives setting	(usually implicit)	'It was winter. The lake had frozen.'
Evidence	Supports a claim	after all, for example	'He is guilty. His fingerprints were found.'

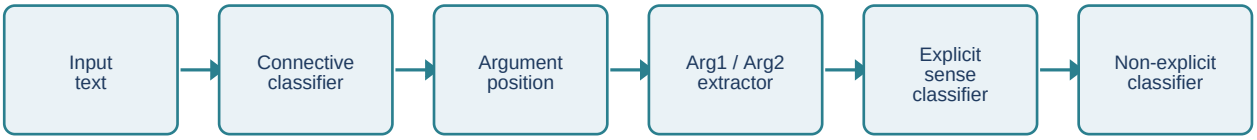
PDTB top-level senses: Temporal (synchronous, asynchronous), **Contingency** (cause, condition, purpose), **Comparison** (contrast, concession), **Expansion** (conjunction, instantiation, restatement, alternative, list).

	Explicit	Implicit
Signal	Connective present	No connective; inferred
Example	'He failed **because** he did not study.'	'He failed. He did not study.' (Explanation)
Automatic accuracy	High (above 90 percent given the connective)	Low (about 40-55 percent); the main bottleneck

2.3 Architecture

A PDTB-style end-to-end **discourse relation parser** (e.g., Lin et al. 2014) runs a pipeline. Explicit relations are found first; remaining adjacent sentence pairs are classified as implicit, entity-based (EntRel) or no relation.

End-to-end relation parser



Stage	Description
Connective classifier	Is 'since', 'and', 'when' used as a discourse connective or not? (about 95 percent accuracy)
Argument position / extraction	Decide whether Arg1 is in the same or a previous sentence, then find the spans of Arg1 and Arg2
Explicit sense classifier	Assign sense (Contrast, Cause...) from the connective, its context and syntax
Implicit relation classifier	For adjacent sentence pairs without a connective: encode (Arg1, Arg2) with BiLSTM / BERT and classify into senses; some models predict a hidden connective first
Output	List of (Arg1, Arg2, connective, sense) tuples

Implicit relation model:

Neural implicit relation classifier



2.4 Example, step by step

```
[1] Priya studied hard. [2] She passed the exam. [3] However, she did not get the scholarship
because the cutoff was very high. [4] For example, only 3 students qualified.
```

Step	Action	Result
1	Split into units	EDUs: [1], [2], [3a] she did not get the scholarship, [3b] because the cutoff was very high, [4]
2	Find connectives	'However' (start of 3), 'because' (inside 3), 'For example' (start of 4); none between 1 and 2
3	Explicit senses	However: Comparison.Concession/Contrast (2 vs 3); because: Contingency.Cause.Reason (3b explains 3a); For example: Expansion.Instantiation
4	Implicit relation (1 and 2)	No connective; world knowledge: studying leads to passing, so Contingency.Cause (Result). Could insert 'so'
5	Mark arguments	Arg1 = [2], Arg2 = [3] for 'However'; Arg1 = [3a], Arg2 = [3b] for 'because'
6	Output	[1] --Result--> [2] --Contrast--> [3a] <--Explanation-- [3b]; [3] --Instantiation--> [4]

2.5 Applications

- **Question answering:** why-questions need Cause/Explanation; comparison questions need Contrast.
- **Summarisation:** drop Elaboration/Example material, keep main claims.
- **Argument mining:** find claims, evidence, concessions and rebuttals.
- **Sentiment analysis:** contrast and concession flip polarity ('good camera but poor battery').
- **Machine translation:** translate connectives and implied relations correctly.
- **Text generation:** choose suitable connectives and plan paragraph logic.
- **Information extraction:** link events and explanations in legal, medical and news text.

Topic 3: Rhetorical Structure Theory (RST)

3.1 Definition

Rhetorical Structure Theory (Mann and Thompson, 1988) is a theory of text organisation that represents a text as a hierarchical tree. Leaves are Elementary Discourse Units (EDUs); internal nodes are spans connected by rhetorical relations, typically between a more central **nucleus** and a supporting **satellite**.

3.2 Characteristics

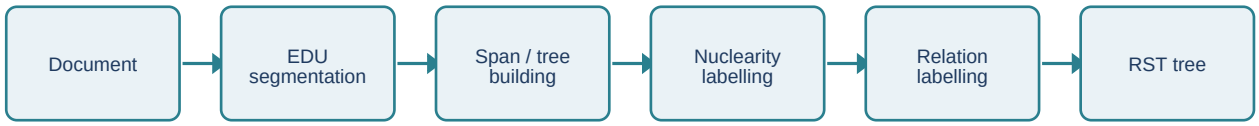
Concept	Explanation
EDU	Minimal clause-like unit; the leaves
Span	Contiguous text made of one or more EDUs
Nucleus / Satellite	Nucleus = essential content; satellite = supporting content that depends on it. Removing satellites leaves a readable core (nuclearity principle)
Mononuclear relations	One N and one S: Elaboration, Cause, Evidence, Background, Concession, Condition, Purpose, Evaluation, Summary
Multinuclear relations	Equal nuclei: Sequence, Contrast, Joint (List), Same-unit
Subject-matter vs presentational	Informational (Cause, Circumstance, Elaboration) vs reader-affecting (Evidence, Justify, Motivation, Antithesis)
Relation definition	Each relation has: constraints on N, constraints on S, constraints on N+S, and an intended effect on the reader
Four principles	Completeness (root spans text), Connectedness (every span is part of another), Uniqueness (no duplicate nodes), Adjacency (spans are contiguous)
Corpus	RST-DT: 385 WSJ articles, 78 relations in 18 classes, binarised trees

Example relation definition - **Evidence**: N = a claim the reader may not believe enough; S = information the reader accepts; N+S = S increases belief in N; Effect = reader's belief in N rises.

Strengths: captures global hierarchy and importance, supports summarisation. **Limits:** subjective annotation, strict tree (no crossing or double relations), small corpora, hard implicit relation labelling.

3.3 Architecture

RST discourse parser pipeline



Stage	Description
EDU segmentation	Token tagging (BiLSTM-CRF / BERT): is this token the start of a new EDU? About 93-96 F1
Tree building	Bottom-up (shift-reduce transition system, CKY chart parsing with dynamic programming) or top-down (recursively split a span at the best point)
Nuclearity labelling	Classify each merge as N-S, S-N or N-N
Relation labelling	Classify the relation using connectives, lexical cues, embeddings; hardest step

Stage	Description
Output / evaluation	Binary tree; scored by Span, Nuclearity, Relation and Full F1 (relation F1 much lower than span F1)

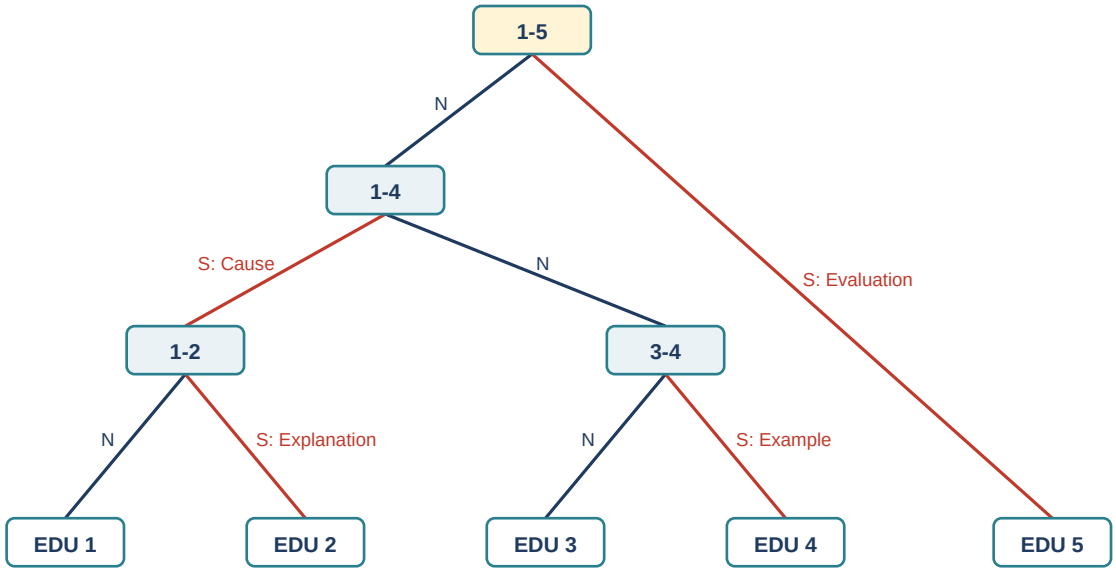
Parser families: SVM-based HILDA, two-stage CRF CODRA, shift-reduce DPLP, neural transition-based and top-down parsers, and recent sequence-to-sequence / LLM-based parsers. **Shift-reduce actions:** SHIFT (push next EDU), REDUCE-NS, REDUCE-SN, REDUCE-NN with a relation label.

3.4 Example, step by step

```
[1] The company's profits fell sharply in 2025.
[2] Demand for its phones dropped.
[3] The company will therefore cut 500 jobs.
[4] For instance, the Pune factory will close.
[5] Analysts expect a recovery by 2027.
```

Step	Action	Result
1	Segment into EDUs	Five EDUs, one per sentence
2	Join [1] and [2]	[2] explains [1]: relation Explanation, N = 1, S = 2, span (1-2)
3	Join [3] and [4]	[4] is an instance of [3]: Example, N = 3, S = 4, span (3-4)
4	Join (1-2) and (3-4)	Profit loss and weak demand are the reason for cuts: Cause, N = (3-4) (main message), S = (1-2)
5	Attach [5]	Analysts' view assesses the situation: Evaluation, N = (1-4), S = 5
6	Check principles	Root spans 1-5; all spans adjacent; every span has one parent
7	Summarise by nuclei	Root -> (1-4) -> (3-4) -> EDU 3: 'The company will cut 500 jobs.'

Fig: RST tree for the paragraph (black = nucleus link, red = satellite link with relation label)



3.5 Applications

- **Text summarisation:** keep nuclei, drop satellites (strong baseline for extractive summaries).
- **Discourse-aware sentiment analysis:** weight nucleus opinions more than satellites.
- **Machine translation and text generation:** plan document structure and relations.
- **Question answering:** locate Cause / Evidence spans for why and how questions.
- **Essay scoring and writing support:** assess argument structure and organisation.
- **Text simplification and compression:** remove satellite elaborations.
- **Information retrieval / argumentation:** rank passages by rhetorical importance.

Topic 4: Temporal and Causal Relations

4.1 Definition

Temporal relation: specifies the ordering or overlap in time between two events, or between an event and a time expression (before, after, during, simultaneous...).

Causal relation: holds when one event or situation (the cause) brings about, enables or prevents another (the effect).

4.2 Characteristics

Temporal relations (TimeML elements: EVENT, TIMEX3, SIGNAL, TLINK)

Allen relation (and inverse)	Meaning for events A and B	Example
before / after	A ends before B starts	'He ate, then left.'
meets / met-by	A ends exactly as B starts	'Class ended; the exam began.'
overlaps / overlapped-by	A starts first and partly overlaps B	'It rained as the match was starting.'
starts / started-by	Same start, A shorter	'The song began with lights dimming.'
during / contains	A entirely inside B	'He slept during the flight.'
finishes / finished-by	Same end, A shorter	'The call ended as the meeting finished.'
equals	Same interval	'She sang and played together.'

- **Cues:** connectives (before, after, when, while, until, then), tense and aspect (past perfect = earlier event), time expressions ('the day before'), narrative convention (text order is usually time order), world knowledge.
- **Textual order is not always temporal order** (flashbacks, explanations).

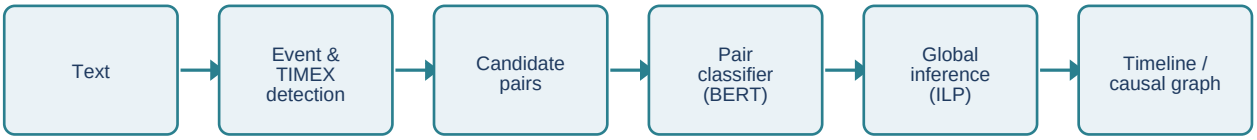
Causal relations

Type	Meaning	Example
Cause	Directly brings about effect	'Heavy rain flooded the road.'
Enable	Makes effect possible	'The key let him open the door.'
Prevent	Stops effect occurring	'The helmet prevented a head injury.'
Facilitate / Inhibit	Raises or lowers likelihood	'Good soil helps crops grow.'

- **Explicit markers:** because, since, so, therefore, hence, thus, as a result, due to, led to, caused. **Implicit causality** depends on world knowledge.
- **Tests:** counterfactual (without A, would B still occur?), temporal precedence (cause not after effect), plausible mechanism.
- **Link between the two:** Cause implies Before (or overlap at start); but Before does not imply Cause. Explanation reverses narrative order ('She was late. The bus broke down.').
- **Resources:** TimeBank, TempEval, TB-Dense, MATRES (temporal); Causal-TimeBank, EventStoryLine, BECauSE, PDTB (causal).

4.3 Architecture

Temporal and causal relation extraction pipeline



Stage	Description
Event and time detection	Tag event verbs/nominals; normalise time expressions ('yesterday' to a date using the document creation time)
Pair generation	Pair events in the same or adjacent sentences, and each event with the document creation time
Pair classification	Predict TLINK type (BEFORE, AFTER, INCLUDES...) and CAUSE / NO-CAUSE using features (tense, aspect, connectives, dependency path) or transformers with entity markers
Global inference	Enforce transitivity (A<B, B<C implies A<C) and consistency (Cause implies Before) with Integer Linear Programming or constrained learning
Output	Partial-order timeline and causal chains / graph

4.4 Example, step by step

Part A - Temporal.

```
John entered the restaurant. He ordered pasta, and after eating he left.
He had booked a table the day before.
```

Step	Action	Result
1	Detect events	e1 entered, e2 ordered, e3 eating, e4 left, e5 booked
2	Detect time expressions	t1 = 'the day before'
3	Collect cues	'and' (sequence), 'after' (e4 after e3), past perfect 'had booked' and 'day before' (earlier)
4	Classify pairs	e1 BEFORE e2; e2 BEFORE e3; e3 BEFORE e4; e5 BEFORE e1
5	Transitive closure	e5 < e1 < e2 < e3 < e4
6	Timeline	booked, entered, ordered, eating, left (text order differs: flashback)

Part B - Causal.

```
The heavy rain flooded the road, so the bus was late, which made Anita miss her exam.
```

Step	Action	Result
1	Detect events	E1 rain flooded road, E2 bus was late, E3 Anita missed exam
2	Find markers	'so' (E1-E2), 'which made' (E2-E3)
3	Direction	E1 causes E2; E2 causes E3
4	Counterfactual test	No flood, bus likely not late: valid. No delay, exam not missed: valid
5	Temporal consistency	E1 BEFORE E2 BEFORE E3: consistent
6	Output	Causal chain: flood, delay, missed exam

Implicit variant: 'Anita missed her exam. The bus was late.' has no marker but is read as Explanation: the later sentence states the earlier cause.

4.5 Applications

- **Event timelines:** news, clinical records (patient history), legal case chronologies.
- **Question answering:** when-questions (temporal) and why-questions (causal).
- **Causal knowledge graphs:** risk analysis, finance, healthcare, supply chains.
- **Summarisation of narratives and news:** order events correctly and keep cause-effect chains.
- **Story understanding and generation:** plot consistency via time order and causes.
- **Forecasting and planning:** predict likely effects of events.
- **Dialogue and assistants:** reason about schedules, reminders and explanations.

Topic 5: Summary and Exam Questions

5.1 Comparison of the four topics

	Local vs Global	Coherence relations	RST	Temporal / Causal
Core idea	Scope of connectedness	Link between two units	Hierarchical tree of relations	Time order and cause-effect
Key units	Sentences, whole text	Arg1, Arg2	EDU, span, N/S	Event, TIMEX, TLINK
Key models	Centering, entity grid, neural	PDTB parser, Hobbs	HILDA, CODRA, DPLP	TimeML, Allen, ILP
Hardest part	Global modelling	Implicit relations	Relation labelling	Implicit time and causality
Top use	Essay and generation scoring	QA, argument mining	Summarisation	Timelines, QA, causal graphs

5.2 Likely exam questions

1. Define local and global coherence. Compare them and explain Centering Theory and the entity grid with an example.
2. What are coherence relations? Explain explicit vs implicit relations and describe the PDTB relation parser architecture.
3. Explain Rhetorical Structure Theory: EDU, nucleus, satellite, relations, principles. Build an RST tree for a short paragraph.
4. Describe the architecture of an RST parser and how it is evaluated.
5. Explain temporal relations with Allen's interval algebra and give examples.
6. Explain causal relations, their markers and the counterfactual test. How are they related to temporal relations?
7. Describe a pipeline for extracting temporal and causal relations from text.
8. List applications of coherence modelling, RST and temporal/causal reasoning.

Note on figures: corpus sizes (e.g., RST-DT: 385 documents, 78 relations, 18 classes) and accuracy ranges quoted here come from the standard literature as I recall it; check them against your textbook or papers before quoting in an exam.