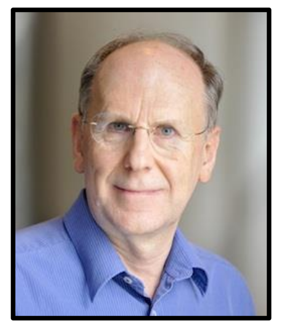
A Linear-Time Bottom-Up Discourse Parser with Constraints and Post-Editing

Vanessa Wei Feng and Graeme Hirst

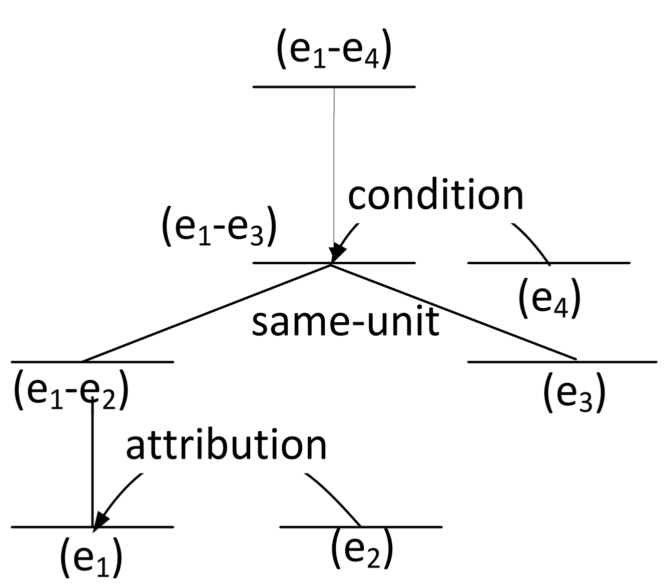
Department of Computer Science, University of Toronto, Canada



1. Rhetorical Structure Theory[1]

- A discourse tree representation of the full text.
  - Leaves: elementary discourse units (EDUs).
  - Internal nodes: concatenation of continuous EDUs, with discourse relations among them labeled.
- Example

[Catching up with commercial competitors in retail banking and financial services,]e_1 [they argue,]e_2 [will be difficult,]e_3 [particularly if market conditions turn sour.]e_4



2. Two-stage Discourse Parsing

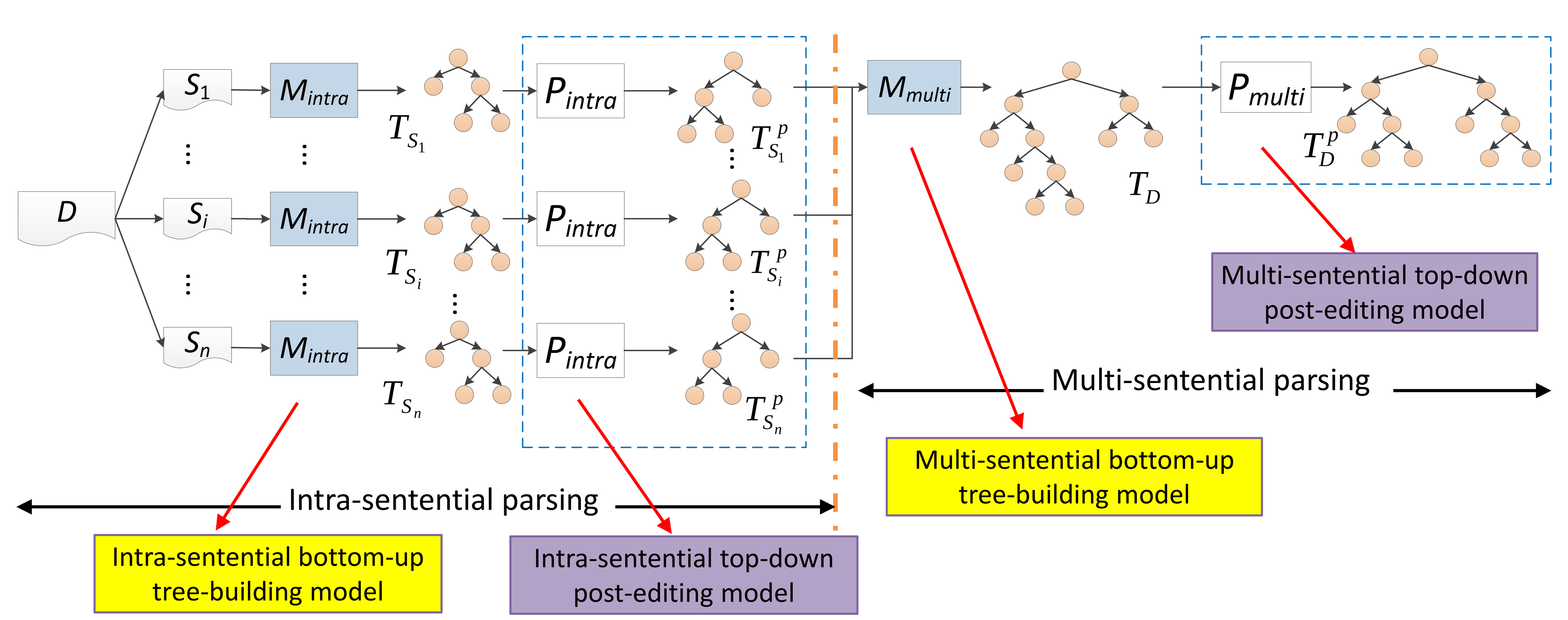
- Joty et al., 2013[2]. Decompose text-level discourse parsing into intra- and multi-sentential parsing.
- Joint modeling of the structure and the relation of adjacent text units.
- CKY-like parsing algorithm to build the discourse tree from bottom up.

Pros
- Takes into account the interaction between structures and relations.
- Globally optimal tree.

Cons
- Inefficient for large documents. CKY-parsing: $O(n^3)$ time complexity.

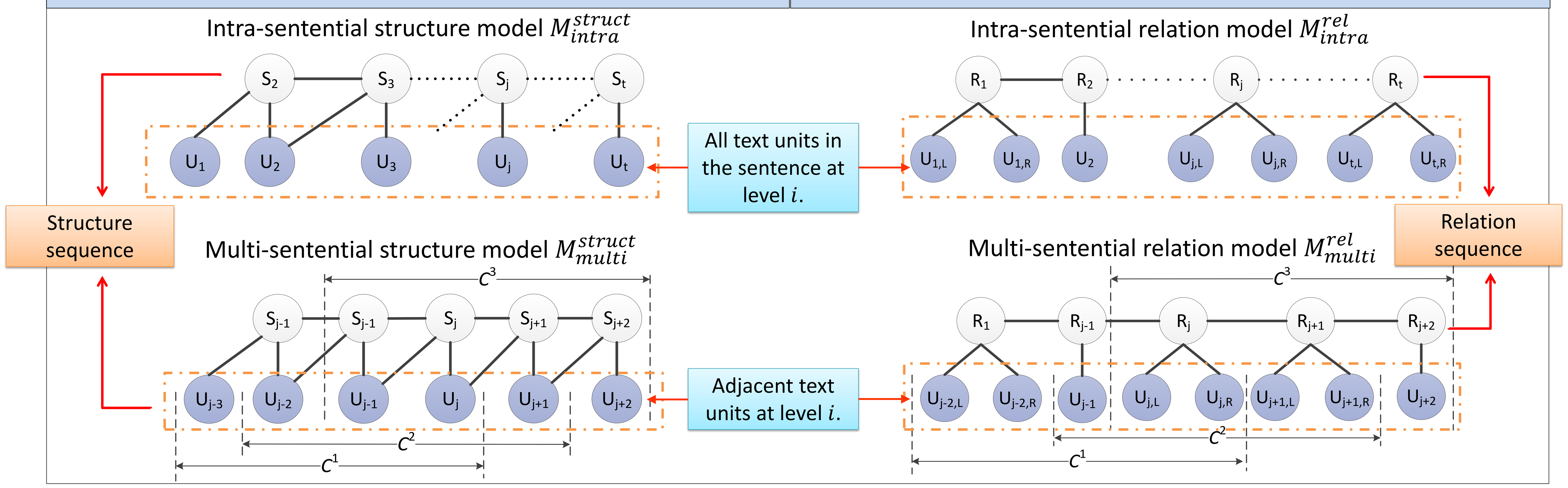
- **Our objective:** to develop a text-level discourse parser that is both accurate and efficient.
- **Our idea:** two-stage parsing with greedy tree-building and post-editing.
  - Post-editing: Incorporates additional features derived from the initial tree.

3. Overall Workflow



4. Linear-chain CRFs with Constraints as Local Models

- Two-stage pipeline in local models:
  - Decompose $P(structure, relation)$ into $P(structure) \times P(relation \mid structure)$.
  - First identify the pair of adjacent text units to be related, then assign the relation to the pair.
- Efficient in practice: $O(n)$ time complexity.



5. Experiments

- Features
  - Organization, textual, n-gram, dominance, contextual, substructure, entity transition, cue phrases, and post-editing features.
- Data
  - The RST Discourse Treebank: 347 for training and 38 for testing.
  - # of sentences: 2 to 187, average 26.
- Evaluation
  - Parsing accuracy using constituent precision and recall[3].
  - Parsing time.

Estimated time consumption of Joty et al. (2013) [2] on the largest document is over 16 hours.

Our previous work.

6. Results

- Parsing accuracy

Model	Span	Nuclearity	Relation
Joty et al., 2013[2]	82.5	68.4	55.7
Feng and Hirst, 2012[4]	82.8	67.1	52.0
Ours (no post-editing)	**84.9***	**69.9***	**57.2***
Ours	**85.7*†**	**71.0*†**	**58.2*†**
Human	*88.7*	*77.7*	*65.8*

* Significantly better than Joty et al., 2013 ($p < .01$).
† Significantly better than ours (no post-editing) ($p < .01$).

- Parsing time in seconds

Model	Min	Max	Average
Feng and Hirst, 2012[4]	0.42	124.86	11.19
Ours (no post-editing)	**0.05**	**40.57**	**5.52**
Ours	**0.12**	**84.72**	**10.71**

Parsing time excludes the time for necessary pre-processing.

References

[1] William Mann and Sandra Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.
[2] Shafiq Joty, Giuseppe Carenini, Raymond Ng, and Yashar Mehdad. 2013. Combining intra- and multi-sentential rhetorical parsing for document-level discourse analysis. *In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)*, pages 486–496, Sofia, Bulgaria.
[3] Daniel Marcu. 2000. *The Theory and Practice of Discourse Parsing and Summarization*. The MIT Press.
[4] Vanessa Wei Feng and Graeme Hirst. 2012. Text-level discourse parsing with rich linguistic features. *In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL 2012)*, pages 60–68, Jeju, Korea.

Acknowledgements: We thank Professor Gerald Penn and the reviewers for their valuable advice and comments. This work was financially supported by the Natural Sciences and Engineering Research Council of Canada and by the University of Toronto.