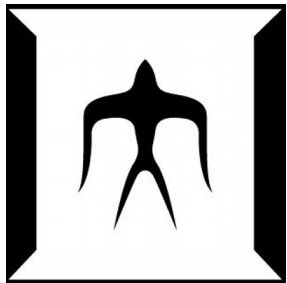# Evaluating Text Coherence Based on Semantic Similarity Graph

Jan Wira Gotama Putra and **Takenobu Tokunaga**

Tokyo Institute of Technology, Japan

`wiragotama.github.io | gotama.w.aa@m.titech.ac.jp`

Motivation

- Modeling coherence in linguistics theory into computational task (Barzilay & Lapata, 2008; Guinaudeau & Strube, 2013; Feng et al., 2014; Li and Hovy, 2014; Petersen et al., 2015, Nguyen and Joty, 2017)

- Approaches

  - Supervised – mostly

  - Unsupervised – infrequent

Coherence

- Coherent text is integrated as a whole, rather than a series of independent sentences (Bamberg, 1983; Garing, 2014)

  Graph structure

- Every sentence in a coherent text has relation(s) to each other (Halliday and Hasan, 1976; Mann and Thompson, 1988; Grosz et al., 1995; Wolf and Gibson, 2005)

  Evaluate coherence through cohesion

- Lexical and semantic (meaning) continuity are indispensable for coherent text (Feng et al., 2014)

  Semantic similarity

Related Work: Entity Graph (1)
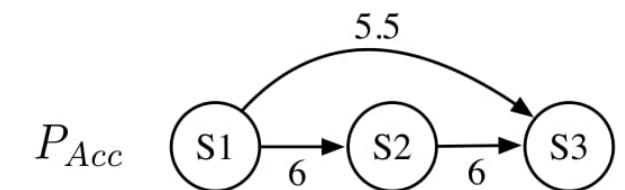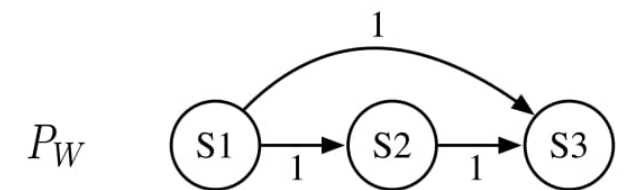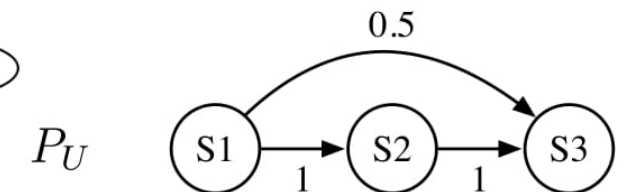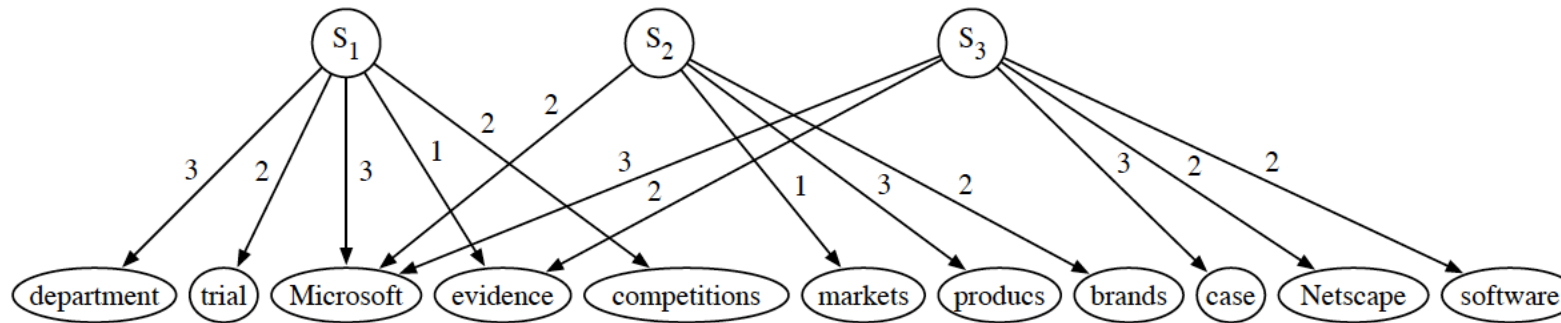
- Entity graph was introduced by Guinaudeau & Strube (2013)

- Text -> Bipartite Graph -> Projection Graphs

- Coherence is achieved by cohesion: considers repeated mention of entities and their syntactical role (weight)

Related Work: Entity Graph (2)

- Graph data structure can represent the structure of text and relations among sentences

- Coherence is achieved through lexical cohesion: repeated mention of entities.
  - Disadvantage: cannot capture the relation between related-yet-not identical entities (Li and Hovy, 2014; Petersen et al., 2015)
  - Solution: use distributed representation of words/sentences

- Relation between vertices in projection graph has to satisfy surface sequential ordering
  - Proposal: allows two directions (omit the constraint)

Proposed Method (1)

- Formally, text is a graph $G(V, E)$, where
  - V is a set of vertices, v_i represents i-th sentence.
  - E is a set of edges, e_{ij} represents relation (cohesion) from i-th to j-th sentence (weighted & directed).
  - Evaluate the coherence through cohesion

- Sentences are encoded into their meaning form
  Average of summation of word vectors (distributed representation of words)
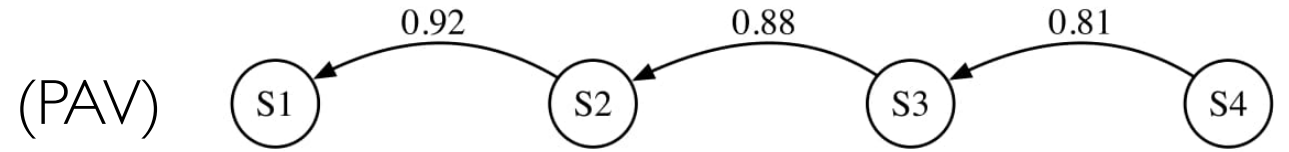
  $$\vec{s_i} = \frac{1}{M} \sum_{k=1}^{M} \vec{w_k}$$

- An edge represents cohesion among sentences
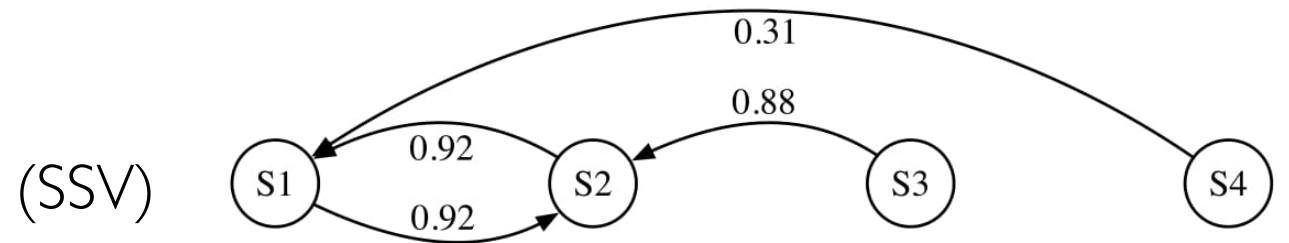 Establishment of edge is decided as the operation of vectors representation of sentences

Proposed Method (2)

- Preceding Adjacent Vertex   (PAV)

- Single Similar Vertex   (SSV)

- Multiple Similar Vertex   (MSV)
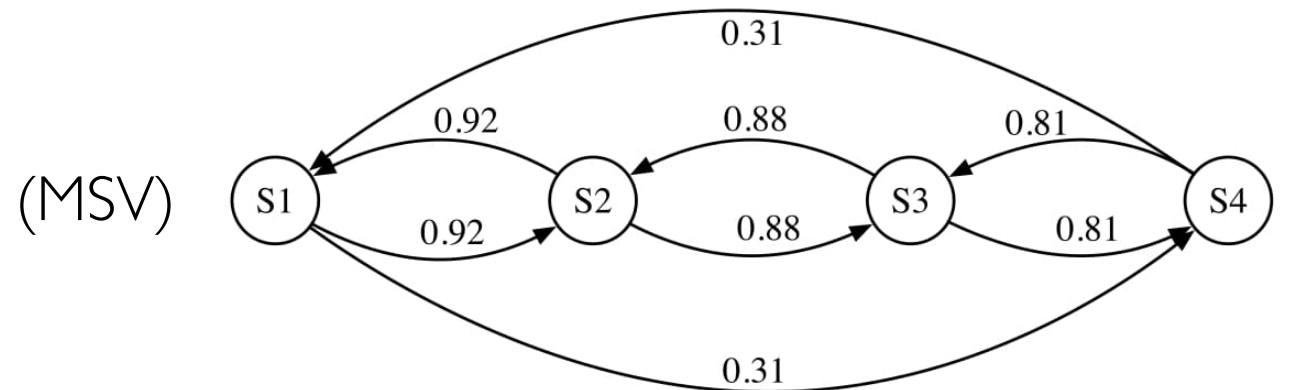
Proposed Method (3)

- An edge is established from the sentence vertex in question to the other vertex with the weight calculated by

$$weight(e_{ij}) = \frac{\cos\_\text{sim}(\vec{s_i}, \vec{s_j})}{|i - j|}$$

normalization

- Text coherence measure (higher is better) is calculated by averaging the averaged weight of outgoing edges from every vertex in the graph as

$$t_c = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{k=1}^{L_i} weight(e_{ij})$$

\# vertices

\# outgoing edges of vertex v_i

Evaluation

- Task 1:   Discrimination (Barzilay and Lapata, 2008)
- Task 2:   Insertion (Eisner and Charniak, 2011)

- Both tasks evaluate how well the methods in comparing coherence between texts

Evaluation: Discrimination Task

- The goal is to compare original vs. permutated text

- Program is considered successful when giving greater score to the more coherent (original) text

- Dataset: 683 WSJ (LDC) texts, 13586 permutations (avg. 24 sentences, 521 tokens)

S1	S4
S2	S3
S3	S1
S4	S2
original	permutated
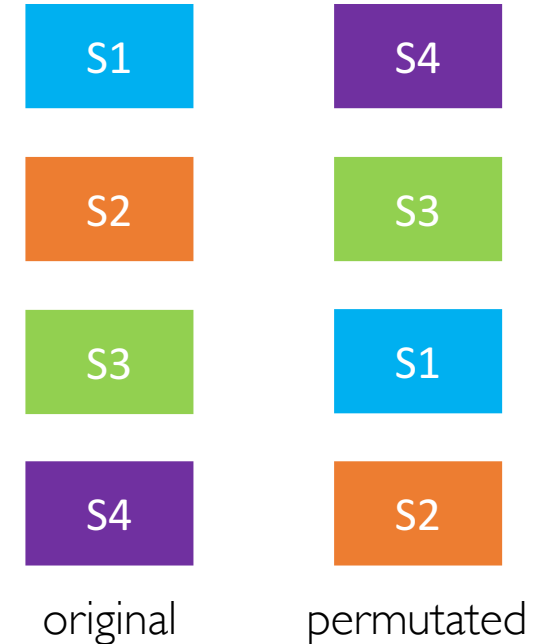
Result: Discrimination Task

- Difference of performance is statistically significant at $p < 0.05$

- PAV > MSV > Entity Graph
  Cohesion is not only about repeating mention of entities

- PAV – MSV pair shares 88.3% same judgement (largest).
  Local (adjacent) cohesion is possibly more important than long-distance cohesion

Method	Accuracy
PAV	**0.774**
SSV	0.676
MSV	0.741
Entity Graph	0.725

Evaluation: Insertion Task

- Insertion task is more important than discrimination task

- It was proposed by Eisner and Charniak (2011):
  - Given a text, take out a sentence (randomly), then place it into other positions
  - Program is considered successful if it prefers to insert take-out-sentence at its original position rather than arbitrary (distorted) positions
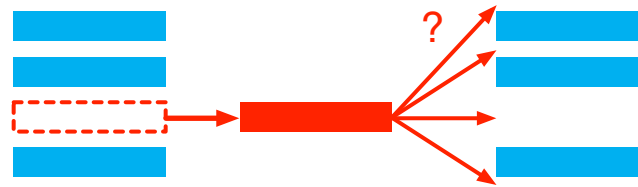
- Our Proposal: use TOEFL iBT insertion-type questions

TOEFL iBT Insertion‒type Question

- A text is coherent even without the insertion sentence

- Preservation of coherence is achieved when the question-sentence is inserted in the correct place but disrupt coherence otherwise

- 104 questions
  (avg. 7 sentences, 139 tokens)

(A) The raising of livestock is a major economic activity in semiarid lands, where grasses are generally the dominant type of natural vegetation. (B) The consequences of an excessive number of livestock grazing in an area are the reduction of the vegetation cover and trampling and pulverization of the soil. (C) This is usually followed by the drying of the soil and accelerated erosion. (D)

Question: Insert the following sentence into one of (A)-(D)
question sentence = "This economic reliance on livestock in certain regions makes large tracts of land susceptible to overgrazing."

correct answer = B

Result: Insertion Task

- Difference in every pair of methods is not statistically significant at $p < 0.05$

Method	Accuracy
PAV	**0.356**
SSV	0.346
MSV	0.327
Entity Graph	0.260

- 14 questions are answered incorrectly by PAV, but correctly by SSV.

- In these questions, SSV tends to establish the relationship between distance sentences (dist = 2.8). For example, exemplification text

Conclusion and Future Work

- Coherence can be achieved through cohesion (lexical and semantic continuity)

- Local cohesion is more important than long-distance cohesion in evaluating coherence, but long-distance cohesion can also contribute as well
  - (PAV > {SSV, MSV})
  - We need to introduce a more refined mechanism for incorporating distant sentence relations.

- The representation of sentences and method to establish edges would be direct targets of the refinement

Appendix

Discrimination Task	
Method	Accuracy
PAV	0.774
SSV	0.676
MSV	0.741
Entity Grid	**0.845**
Entity Graph	0.725

Insertion Task	
Method	Accuracy
PAV	**0.356**
SSV	0.346
MSV	0.327
Entity Grid	0.346
Entity Graph	0.260

References (1)

1. Alphie G. Garing. 2014. Coherence in argumentative essays of first year college of liberal arts students at de la salle university. *DLSU Research Congress.*

2. Barbara J. Grosz, Scott Weinstein, and Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2):203-225.

3. Betty Bamberg. 1983. What makes a text coherent. *College Composition and Communication* 34(4):417-429.

4. Camille Guinaudeau and Michael Strube. 2013. Graph-based local coherence modeling . In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).* Association for Computational Linguistics, Sofia, Bulgaria, pages 93–103.

5. Casper Petersen, Christina Lioma, Jakob Grue Simonsen, and Birger Larsen. 2015. Entropy and graph based modelling of document coherence using discourse entities: An application to IR. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval,* pages 191-200.

6. Dat Tien Nguyen and Shafiq Joty. 2017. A neural local coherence model. In *Proceedings of Annual meeting for association for computational linguistics.*, pages 1320-1330.

7. Florian Wolf and Edward Gibson. 2005. Representing discourse coherence: A corpus-based study. *Computational Linguistics* 31(2):249–288.

References (2)

8. Jiwei Li and Eduard Hovy. 2014. A model of coherence based on distributed sentence representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Doha, Qatar, pages 2039–2048.

9. Micha Eisner and Eugene Charniak. 2011. Extending the entity grid with entity-specific features. In *Proceedings of the 49th Annual Meeting of Association for Computational Linguistics: HLT short papers.*, pages 125-129.

10. M.A.K Halliday and Ruqaiya Hasan. 1976. Cohesion in English. Longman, Singapore.

11. Regina Barzilay and Mirela Lapata. 2008. Modeling local coherence: Entity based approach. *Computational Linguistics* 34(1):1–34.

12. Vanessa Wei Feng, Ziheng Lin, and Graeme Hirst. 2014. The impact of deep hierarchical discourse structures in the evaluation of text coherence . In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin City University and Association for Computational Linguistics, Dublin, Ireland, pages 940–949.

13. William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8(3):243–281.