# 1  Introduction

Automatic coherence assessment has typically examined the effect of local discourse features on coherence. These features are typically sentence-to-sentence transitions of entities, discourse relations, or other sentential information. Additionally, this task is often framed as selecting the most coherent ordering of a set of sentences in a document, where the original ordering as written by a human is assumed to be the most coherent ordering.

# 2  Related Work

Local coherence has been strongly influenced by Centering Theory [5], which places constraints on discourse entities and their transitions between text units. Moreover, it establishes the importance of entity salience, and syntactic role for local coherence. While not directly implementing these constraints, [2] uses discourse entities in formulating an *entity grid* model for learning sentence-to-sentence transition features that contribute to document coherence. Additionally, [2] establish the sentence ordering task such that training data can be generated in an unsupervised fashion. The entity grid model, and its variants, have remained popular in the discourse coherence literature, as various authors have tried to augment it with global features [3], more direct coreference features [4], discourse connectives [6], and syntax patterns [7].

Few authors have dealt directly with the issues of intractability. [1] prove the sentence ordering task is an NP complete problem by showing its reducibility to a traveling salesperson problem. [8] uses an IDL (Interleave-Disjunction-Lock) graph formalism to compactly represent the search space of sentence-to-sentence transitions in the ordering problem, and experiment with A* and beam search algorithms for learning/decoding. The IDL formalism allows for a complex scoring for sentence to sentence transitions, while using an admissible heuristic to prevent *unfolding* unnecessary search states. Unfortunately, their implementation exhibits run times of over a minute for problems of only eleven sentences, indicating this approach may not scale for some applications.

Put sauper and barzilay 2009 in here for topic level discourse ordering and ILP solver implementation.

# 3  Entity Grid Model

The *entity grid* model is a popular local coherence model where a document rendering is represented by a matrix or grid of discourse entities and their syntactic role, such as the *subject*, *object*, etc. Each entity in the document has a column and each sentence has a row in the grid. A role transition is simply a column subsequence of any length. For example, in Table 1 the entity *winds* has a length 2 transition {-,o} in sentences 2–3.

| | Karen | winds | Friday | Wash., D.C. | Monday | Gulf of Mexico | coast | weekend | authorities | orders | New Orleans |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | s | - | - | - | - | x | x | x | - | - | - |
| **2** | s | - | x | - | - | - | x | - | s | o | x |
| **3** | s | o | x | - | - | - | - | - | - | - | - |
| **4** | s | - | - | x | x | - | - | - | - | - | - |

Table 1: An example entity grid representation.

Let a document $d$ be an unordered set of $n$ sentences. We define the set $\mathcal{E} = \{e_1, e_2, \ldots, e_{|\mathcal{E}|}\}$ to be the set of $|\mathcal{E}|$ entities that occur in $d$ and the set $\mathcal{R}$ to be a finite set of syntactic roles. We also define the set of document renderings $\mathcal{P} = \{\rho_1, \rho_2, \ldots, \rho_{n!}\}$ to be the set of ordered permutations of these sentences. A document rendering $\rho_i = \{s_{i,j} : 1 \leq j \leq n\}$ is an ordered sequence of sentences where a sentence is a vector $s_{i,j} = \{r_{i,j,e_k} : 1 \leq k \leq |\mathcal{E}|\}$ of syntactic roles.

We formally define the entity grid model of a document rendering $\rho_i$ to be the sets $\mathcal{E}$, $\mathcal{R}$, and the set of transition vectors $t_{i,j}$ for all $j$, $2 \leq j \leq n$. A transition vector $t_{i,j} \in \{\mathcal{R} \times \mathcal{R}\}^{|\mathcal{E}|}$ is a vector of entity role transitions from $s_{i,j-1}$ to $s_{i,j}$ where each element corresponds to a transition $q \in \{\mathcal{R} \times \mathcal{R}\}$. Define the function $\boldsymbol{\phi} : \mathcal{P} \to \mathbb{R}^{|\mathcal{R} \times \mathcal{R}|}$ with $\boldsymbol{\phi}(\rho_i) = \{\phi_q(\rho_i) : \forall q \in \mathcal{R} \times \mathcal{R}\}$ and

$$\phi_q(\rho_i) = \frac{\sum_{j=2}^{n} \sum_{k=1}^{|\mathcal{E}|} I(t_{i,j,e_k} = q)}{|\mathcal{E}|(n-1)}. \tag{1}$$

Given a vector of weights $\mathbf{w} \in \mathbb{R}^{|\mathcal{R} \times \mathcal{R}|}$, we define our objective function $f : \mathbb{R}^{|\mathcal{R} \times \mathcal{R}|} \to \mathbb{R}$ where $f(\boldsymbol{\phi}(\rho_i); \mathbf{w}) = \mathbf{w} \cdot \boldsymbol{\phi}(\rho_i)$. The sentence ordering task can further be defined as finding the rendering that maximizes this function, or

$$\underset{i \in \{1 \ldots n!\}}{\arg\max} f(\boldsymbol{\phi}(\rho_i); \mathbf{w}). \tag{2}$$

Solving this problem for larger documents is intractable as the time complexity of computing the $\arg\max$ is $O(|\mathcal{P}||\mathcal{R}|^2) = O(n!|\mathcal{R}|^2)$; to compute each ranking, we must perform $|\mathcal{R}|^2$ multiplications, and to find the maximum we must check $n!$ possible renderings.

## 4  TSP Formulation

We formulate this problem as a Traveling Salesperson Problem where each sentence $s$ is a vertex in a fully connected graph. Adding an edge $(s_i, s_j)$ to

the solution path represents selecting $s_j$ as the next sentence in the document rendering. In order to calculate edge weights, we define a function $\boldsymbol{\psi} : E \to \mathbb{Z}^{|\mathcal{R} \times \mathcal{R}|}$ where $\boldsymbol{\psi}(s_i, s_j) = \{\psi_q(s_i, s_j) : q \in \mathcal{R} \times \mathcal{R}\}$ and

$$\psi_q(s_i, s_j) = \sum_{k=1}^{|\mathcal{E}|} I\left(\langle s_{i,e_k}, s_{j,e_k} \rangle = q\right). \tag{3}$$

Edge weights are determined by the function

$$g(\boldsymbol{\psi}(s_i, s_j); \mathbf{w}) = \mathbf{w} \cdot \boldsymbol{\psi}(s_i, s_j) = \sum_{q=1}^{|\mathcal{R}|^2} w_q \left(\sum_{k=1}^{|\mathcal{E}|} I(\langle s_{i,e_k}, s_{j,e_k} \rangle = q)\right). \tag{4}$$

where $\langle s_{i,e_k}, s_{j,e_k} \rangle$ is a transition $q \in \{\mathcal{R} \times \mathcal{R}\}$ for entity $k$ obtained from ordered sequence of $s_i, s_j$.

Under this formulation, the optimal path is equivalent to the solution to the optimal rendering in the entity grid formulation. The total edge weights for the optimal path are the sum of $g(\boldsymbol{\psi}(s_i, s_j); \mathbf{w})$ for all $(s_i, s_j)$ in the path. Without loss of generality, let the optimal path $\rho = \{s_1, s_2, \ldots, s_n\}$.

$$\text{Optimal Path Edge Weight} = \sum_{j=2}^{n} g(\boldsymbol{\psi}(s_{j-1}, s_j); \mathbf{w})$$

$$= \sum_{j=2}^{n} \sum_{q=1}^{|\mathcal{R}|^2} w_q \left(\sum_{k=1}^{|\mathcal{E}|} I(\langle s_{j-1,e_k}, s_{j,e_k} \rangle = q)\right)$$

$$= \sum_{q=1}^{|\mathcal{R}|^2} w_q \left(\sum_{j=2}^{n} \sum_{k=1}^{|\mathcal{E}|} I(\langle s_{j-1,e_k}, s_{j,e_k} \rangle = q)\right)$$

Equivalently the arg max of the Entity Model objective function is

$$\operatorname*{arg\,max}_{i \in \{1...n!\}} f(\boldsymbol{\phi}(\rho_i); \mathbf{w}) = \operatorname*{arg\,max}_{i \in \{1...n!\}} \sum_{q=1}^{|\mathcal{R}|^2} w_q \left(\frac{\sum_{j=2}^{n} \sum_{k=1}^{|\mathcal{E}|} I(t_{i,j,e_k} = q)}{|\mathcal{E}|(n-1)}\right)$$

$$= \operatorname*{arg\,max}_{i \in \{1...n!\}} \sum_{q=1}^{|\mathcal{R}|^2} w_q \left(\sum_{j=2}^{n} \sum_{k=1}^{|\mathcal{E}|} I(t_{i,j,e_k} = q)\right)$$

$$= \operatorname*{arg\,max}_{i \in \{1...n!\}} \sum_{q=1}^{|\mathcal{R}|^2} w_q \left(\sum_{j=2}^{n} \sum_{k=1}^{|\mathcal{E}|} I(\langle s_{i,j-1,e_k}, s_{i,j,e_k} \rangle = q)\right)$$

give an example. NetworkX

3

# References

[1] Ernst Althaus, Nikiforos Karamanis, and Alexander Koller. Computing locally coherent discourses. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, page 399. Association for Computational Linguistics, 2004.

[2] Regina Barzilay and Mirella Lapata. Modeling local coherence: an entity-based approach. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ACL '05, pages 141–148, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.

[3] Micha Elsner, Joseph L Austerweil, and Eugene Charniak. A unified local and global model for discourse coherence. In *HLT-NAACL*, pages 436–443, 2007.

[4] Micha Elsner and Eugene Charniak. Coreference-inspired coherence modeling. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers*, pages 41–44. Association for Computational Linguistics, 2008.

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

[6] Ziheng Lin, Hwee Tou Ng, and Min-Yen Kan. Automatically evaluating text coherence using discourse relations. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 997–1006. Association for Computational Linguistics, 2011.

[7] Annie Louis and Ani Nenkova. A coherence model based on syntactic patterns. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 1157–1168. Association for Computational Linguistics, 2012.

[8] Radu Soricut and Daniel Marcu. Discourse generation using utility-trained coherence models. In *Proceedings of the COLING/ACL on Main conference poster sessions*, pages 803–810. Association for Computational Linguistics, 2006.