Description

A text is more than a sequence of sentences. To understand a text, one needs to recognize how sentences are connected with each other and why they appear in a particular order. To capture the particular characteristics of text, we developed a range of methods in computational linguistics: local and global coherence models, anaphora and coreference resolution algorithms, methods for recognizing the rhetorical, the temporal, the causal and the argumentative structure of texts. In the seminar we first talk about linguistically well-founded classical models. Then we continue with more recent machine learning and graph based models before we turn to current neural models for text processing and understanding. The usefulness of such models can only be determined when they are integrated into applications. Hence the module will also deal with evaluating text processing algorithms within applications such as information extraction, machine translation, question answering, readability prediction, essay scoring, etc.

References

• Webber, Bonnie et al. (2012). Discourse structure and language technology, Natural Language Engineering, 18, pp 437-490.

selected chapters from:


Journal Papers

**Dates, Topics**

**27.10.2021**

**Introduction, historical overview . . .**
(Sede, 2012; Webber et al., 2012; Webber & Joshi, 2012)

**Introduction into discourse structure**
(Grosz & Sidner, 1986; Morris & Hirst, 1991; Passonneau & Litman, 1997)

**Introduction into local coherence analysis: Coreference**

**Introduction into centering**
(Brennan et al., 1987; Grosz et al., 1995; Tetreault, 2001; Poesio et al., 2004)

**03.11.2021**

**Local coherence: Information status**

**Linguistic background:** (Daneš, 1974; Chafe, 1976; Prince, 1981; 1992; Arnold et al., 2000)

**Determining and using information status:** (Nissim et al., 2004; Nissim, 2006; Cahill & Riester, 2009; Riester et al., 2010; Riester & Baumann, 2011; Rahman & Ng, 2011; Cahill & Riester, 2012; Markert et al., 2012; Eckart et al., 2012; Baumann & Riester, 2013; Roesiger, 2016; Hou, 2016; 2020)

to be prepared: ((Markert et al., 2012) or (Baumann & Riester, 2013) or (Hou, 2021) – https://arxiv.org/pdf/2109.02753)

**10.11.2021**

Cancelled, MS @ EMNLP/CODI

**17.11.2021**

Cancelled, MS @ HITS Group Leaders’ Retreat

**24.11.2021**

**The structure of documents in different genres**

**Linguistic background:** (Swales, 1990; Kessler et al., 1997; Webber, 2009)

**Determining and using document structure:** (Teufel & Moens, 2002; Power et al., 2003; Sporleder & Lapata, 2004; Graham et al., 2005; Filippova & Strube, 2006; Chung, 2009; Teufel et al., 2009; Liakata et al., 2010; 2012)

**Document structure: Automatic summarizaton**
(Liakata et al., 2012; Contractor et al., 2012; Liakata et al., 2013; Xiao & Carenini, 2019; Xing et al., 2021)

to be prepared: (Teufel & Moens, 2002, pp.409-424) or (Teufel et al., 2009)
01.12.2021
Topic segmentation: Lexical chains
(Morris & Hirst, 1991; Barzilay & Elhadad, 1997; Hirst & St-Onge, 1998; Barzilay & Elhadad, 1999; Silber & McCoy, 2002; Galley et al., 2003; Stokes et al., 2004; Ye & Chua, 2006; Ye et al., 2007; Medelyan, 2007)

Hierarchical topic segmentation
(Grosz & Sidner, 1986; Hsueh et al., 2006; Eisenstein, 2009; Carroll, 2010)

Topic-Segmentation: Distributional approaches
(Hearst, 1997; Beeferman et al., 1999; Reynar, 1999; Choi, 2000; Pevzner & Hearst, 2002)

Probabilistic models for topic segmentation
(Utiyama & Isahara, 2001; Eisenstein & Barzilay, 2008; Shafiei & Milios, 2008; Chen et al., 2009; Purver, 2011; Du et al., 2013; Simon et al., 2013)

Applications of topic segmentation: Automatic summarization
(Goldstein et al., 2000; Teufel & Moens, 2002; Narayanan & Harabagiu, 2004; Stokes et al., 2004)

to be prepared: (Barzilay & Elhadad, 1997) and (Galley et al., 2003) or (Du et al., 2013)

08.12.2021
Cancelled, MS @ NeurIPS

15.12.2021
Modeling local coherence: Centering, entity Grid with applications (information ordering, evaluating the quality of summaries)
(Poesio et al., 2004; Karamanis et al., 2004; Barzilay & Lapata, 2008; Karamanis et al., 2009)

Modeling local coherence: Entity graph with applications (information ordering, evaluating the quality of summaries)
(Guinaudeau & Strube, 2013; Petersen et al., 2015)

Applications of local coherence models: Information ordering, evaluating the quality of summaries
(Elsner et al., 2007; Filippova & Strube, 2007; Pitler et al., 2010; Elsner & Charniak, 2011; Lin et al., 2012; Li & Hovy, 2014)

Applications of local coherence models: Information retrieval
(Petersen et al., 2015)

Applications of local coherence models: Readability – (Higgins et al., 2004; Pitler & Nenkova, 2008; Feng et al., 2009; Wang et al., 2013; Mesgar & Strube, 2015; 2016;
Applications of local coherence models: Essay scoring – (Miltsakaki & Kukich, 2004; Burstein et al., 2010; Beigman Klebanov & Flor, 2013; Beigman Klebanov et al., 2014; Somasundaran et al., 2014; Mesgar & Strube, 2018)

Applications of local coherence models: Coreference and coherence for automatic translation
(Hardmeier et al., 2013; 2015; Born et al., 2017; Jean et al., 2017; Lapshinova-Koltunski et al., 2020)

to be prepared: (Barzilay & Lapata, 2008)

12.01.2022
Local coherence models: Neural approaches (Li & Hovy, 2014; Tien Nguyen & Joty, 2017; Li & Jurafsky, 2017; Joty et al., 2018; Mesgar & Strube, 2018; Moon et al., 2019; Jeon & Strube, 2020a; 2020b)
Applications of neural coherence models (information ordering, readability, essay scoring): (Joty et al., 2018; Mesgar & Strube, 2018; Pishdad et al., 2020; Jeon & Strube, 2020a; Laban et al., 2021; Mohiuddin et al., 2021)
to be prepared: ((Moon et al., 2019) or (Jeon & Strube, 2020a)) and ((Pishdad et al., 2020) or (Mohiuddin et al., 2021))

19.01.2022
Introduction into global coherence analysis
(Hobbs, 1979; 1985; Mana & Thompson, 1988; Knott & Dale, 1994; Webber & Joshi, 1998; Kehler et al., 2008; Webber et al., 2012)

Global coherence: Rhetorical Structure Theory (RST)
(Marciu & Echihabi, 2002; Carlson et al., 2003; Soricut & Marcu, 2003; duVerle & Prenderg, 2009; Subba & Di Eugenio, 2009; Hernault et al., 2012; Feng & Hirst, 2012; Ji & Eisenstein, 2014; Huber & Carennini, 2020; Nguyen et al., 2021; Huber et al., 2021)

Global coherence: Penn Discourse Treebank
(Prasad et al., 2008; Pitler & Nenkova, 2009a; 2009b; Lin et al., 2009; 2010; 2011; Ghosh et al., 2012; Lin et al., 2014; Prasad et al., 2014; Ji et al., 2015; Xue et al., 2016; Mihailov & Frank, 2016; Prasad et al., 2017)

Global coherence: Applications (automatic summarization, sentiment analysis, coreference resolution)
(Marciu, 1997; Maslenikov & Chua, 2007; Clarke & Lapata, 2010; Zirn et al., 2011; Bhatia et al., 2015; Khosla et al., 2021a)

26.01.2022
Local coherence: Rule-based system for coreference resolution: (Raghunathan
Local coherence: Machine learning for coreference resolution: (Soon et al., 2001; Ng & Cardie, 2002; Yang et al., 2008; Ng, 2008; 2010; Fernandes et al., 2012; Durrett & Klein, 2013; Fernandes et al., 2014; Martschat & Strube, 2015)

to be prepared: (Lee et al., 2013) and (Durrett & Klein, 2013) – compare the two approaches; which one might work better in another genre, in another domain?

02.02.2022
Local coherence: Neural coreference resolution: (Clark & Manning, 2015; Wise-man et al., 2015; 2016; Clark & Manning, 2016b; 2016a; Lee et al., 2017; Kundu et al., 2018; Zhang et al., 2018; Sahu et al., 2019; Yu et al., 2020; Jiang & Cohn, 2021; Khosla et al., 2021a; Kirstain et al., 2021)

Local coherence: Critique on coreference resolution systems – (Moosavi & Strube, 2017; 2018)

Local coherence: Coreference resolution in other languages, multilingual coreference resolution – (Luo & Zitouni, 2005; Recasens & Martí, 2009; Chen & Ng, 2013; Kong & Ng, 2013; Martins, 2015; Bernard & Han, 2020; Wilkens et al., 2020; Aloraini & Poesio, 2020; Jwalapuram et al., 2020)

Local coherence: Coreference resolution in other genres and domains – (Sahu et al., 2019; Bamman et al., 2020; Pagel & Reiter, 2020; Martin et al., 2020; Khosla et al., 2021b)

Local coherence: Error analysis for coreference resolution – (Uryupina, 2008; Kummerfeld & Klein, 2013; Martschat & Strube, 2014)

Local coherence: Evaluating coreference resolution algorithms
(Vilain et al., 1995; Bagga & Baldwin, 1998; Popescu-Belis, 2003; Luo & Zitouni, 2005; Recasens & Hovy, 2011; Pradhan et al., 2014; Tuggener, 2014; Moosavi & Strube, 2016)

to be prepared: ((Lee et al., 2017) or (Kirstain et al., 2021)) and (Moosavi & Strube, 2017)

09.02.2022
Bias in coreference resolution – (Rudinger et al., 2018; Webster et al., 2018; Zhao et al., 2018)

Local coherence: Bridging
(Clark, 1975; Hahn et al., 1996; Poesio et al., 1997; Vieira & Teufel, 1997; Vieira & Poesio, 2000; Bunescu, 2003; Fan et al., 2005; Lassalle & Denis, 2011; Hou et al., 2013b; 2013a; Rösiger & Teufel, 2014; Hou et al., 2014; Rösiger, 2018; Hou et al., 2018)
Local coherence: Event coreference resolution
(Bejan & Harabagiu, 2010; Chen & Ji, 2009; Chen et al., 2010b; 2010a; Goyal et al., 2013)

16.02.2022
Discussion about the whole term on different aspects of discourse

Optional:

Global coherence: Temporal structure
(Lapata & Lascarides, 2004; 2006; Ng et al., 2013)

Global coherence: Argumentation structure: – (Reed & Long, 1998; Katzav & Reed, 2008; Stab & Gurevych, 2014; Peldszus & Stede, 2015a; 2015b; Habernal & Gurevych, 2017; Stab & Gurevych, 2017; Wachsmuth et al., 2017)

The debater:

Further Remarks:

Assessment:

1. For each class read the material marked in the schedule as to be prepared. Formulate two questions about the material and send them to me via email until the Tuesday before the class, 8am at the latest. Participate actively in the class (important!).

2. Choose a topic in the schedule you want to present in class. Select one or more papers from the reading list. Present this work in the class (30 minutes presentation, 15 minutes discussion).

3. Choose a second topic and present it in class (as in (3.)). Or: Write a report/an essay towards the end of the term either about the topic you presented in class or about a new topic (6LP: 8-10 pages; 8LP (if possible): 12-15 pages)). Or: Implement a discourse processing algorithm, evaluate it in comparison to a baseline, and write a very short report (3-4 pages) about it.

Literature: Most papers can be downloaded from the ACL Anthology (http://acl.ldc.upenn.edu/), in particular all papers presented at (E/NA)ACL, Coling and EMNLP conferences, all workshops organized during these conferences and the journals TACL and Computational Linguistics. Papers published through AAAI (AAAI conference, AAAI workshops, AAAI symposia, etc.) are available through the AAAI Digital Library verfügbar (http://www.aaai.org/Library). – Other journals are available electronically at the university library (https://www.ub.uni-heidelberg.de/, http://rzblx1.uni-regensburg.de/ezeit/search.phtml?bibid=UBHE&colors=3&lang=de).

Office hours: Right after class, or in my office at HITS (https://www.h-its.org/en/).
References


Hungary, 14 April, 2003, pp. 47–52.


Filippova, Katja & Michael Strube (2007). Extending the entity-grid coherence model to seman-


Hou, Yufang (2016). Incremental fine-grained information status classification using attention-based LSTMs. In *Proceedings of the 26th International Conference on Computational Lin-


Mesgar, Mohsen & Michael Strube (2018). A neural local coherence model for text quality as-


