Theories of discourse and dialogue and their applications

Ani Nenkova
CS224N/Ling 237 2006

Slides borrowed from Advaith Siddharthan and Julia Hirschberg

Lecture overview

- Coherence and cohesion
- Centering theory
- Rhetorical structure theory and applications
- Dialogue

Terminology

- Coherence
  - Is text interpretable by the reader?
  - Mental/semantic phenomenon (reader-centric)
- Cohesion
  - Textual phenomenon
  - Use of linguistic devices to link discourse elements
  - Contributes to coherence

Cohesion

- Conjunctive (Intentional) Approaches
  - Model: Rhetorical relations between text spans
  - Linguistic devices: coordination/subordination/cue phrases
  - Dominant theory: RST (Mann and Thompson, 1988)
- Referential (Attentional) Approaches
  - Model: discourse focus and changes in focus
  - Linguistic devices: pronouns, referring expressions, voice
  - Dominant theory: Centering (Grosz et al, 1995)

Centering theory

- Consecutive utterances are linked
  - by the entities mentioned in then
  - the forms of the entity mention (pronoun, NP)
- Forward looking centers (Un)
  - A list ranked by salience
  - SUBJ>IND OBJ>OBJ>OTHER
- Backward looking centers (Un+1)
  - The highest ranked entity from Un realized in Un+1

Example (Hudson-D'Zmura 1988)

- John went to his favorite music store to buy a piano. He had frequented the store for many years. He was excited that he could finally buy a piano. He arrived just as the store was closing.
- John went to his favorite music store to buy a piano. It was a store John had frequented for many years. He was excited he could finally buy a piano. It was closing just as John arrived.

Transitions

	Cb(Ui)=Cb(Ui-1)	**Cb(Ui) ≠ Cb(Ui-1)**
Cb(Ui) = Cp	Continue	Smooth shift
Cb(Ui) ≠ Cp	Retain	Rough shift

Continue > Retain > Smooth Shift > Rough Shift

Applications of centering

- Uses for generation by computers
  - Choice of anaphora/referring expressions
  - Sentence ordering
- Uses for analyses by computers
  - Anaphora resolution
  - Assessment of text cohesion

Evaluation of writing skills (Miltskaki and Kukich 2000)

- 100 GMAT essays
- Marked for centering transitions
- High percentage of ROUGH shifts correlates with low essay score

Automatic models of local coherence (Barzilay&Lapata 2005)

- Grid representation of text
  - Rows correspond to sentences
  - Columns indicate the type of occurrence of a given entity
    - S: subject
    - O: object
    - X: neither subject nor object
    - -: does not occur in the sentence
- Use automatic co-reference resolution to equate for example *Microsoft Corp*, *Microsoft*, *the company*, *it*.

- Tasks
  - Text ordering
  - evaluating summary coherence
- Features
  - The percentage of each transition type
    - SS, SO, S-, XX etc
  - Salience

Results: pairwise ranking accuracy

Model	**Ordering**	**Summarization**
Coreference+Syntax+ Salience	**87.3**	68.8
Coreference+Salience	86.9	62.5
Syntax+Salience	83.4	**81.3**
Coreference+Syntax	76.5	75.0

Rhetorical structure theory

- Online intro: http://www.sfu.ca/rst
- Mann and Thompson (1988)

… for every part of a coherent text, there is some function, some plausible reason for its presence, evident to readers….

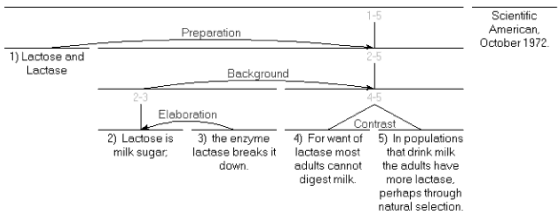
  - RST models intentional structure
  - Every span of text is linked to the rest of the discourse by an RST relation
  - Originally 24 relations: motivation, antithesis, background, elaboration, circumstance, contrast…
  - RST does not consider referential relations

Main ideas: text structure

- Text spans are linked by RST relations
- Text spans are of two types
  - Nucleus: more essential to the author's purpose
  - Satellite: often incomprehensible without the nucleus
- A text is a hierarchical structure of text spans connected by relations

Analysis of the title and abstract

from the beginning of a Scientific American article



Uses of RST

- Text generation
  - Description of tourists sites/museum artifacts
  - Tutoring systems
- Automatic summarization
  - Deleting satellites for shorter summaries
  - Identifying important claims in legal summaries, patent applications
- Evaluation of student essays
- Analysis of scientific articles

Automatic grading of student essays (Bernstein et al 2003)

- Part of standardized tests
- Expensive to grade---can the task be automated
- Yes!
  - A system assigned score coincides with that of a human as often as two different humans assign the same score
  - One of the graders is a machine

Essay structure

- Introduction
- Thesis statement
- Main points (supporting the statement)
  - Elaboration of each point
- Conclusion

➔ Detecting organizational problems can be used to give user feedback

Analysis of scientific articles

- Overall predictable global structure
  - Introduction
  - Method
  - Results
  - Discussion
  - Conclusion
- Makes searching more efficient
  - In a Psycholinguistics paper, where do you quickly find the number of experimental subjects?
  - In a Chemistry paper, where do you look for spectroscopy tables?

Argumentative Zoning Teufel and Moens 1999

- BACKGROUND Generally accapted knowledge
- OTHER Specific other work
- OWN Methods, results, future work
- AIM Specific research goal
- CONTRAST Comparison, weakness of other solution
- BASIS work that has been improved

- Annotators trained to tag these classes
- Automatic classifiers trained
  - Features
    - Cue phrases
      - propose, present, suggest…
      - adopt, agree, originate…
    - Tense, voice, modality
    - Location
    - Section heading
    - Context

Uses for summary lists

- Aim
  - In this work we propose a method for establishing the probabilty of such previously unseen word combinations using available information on “most similar” words.
  - …
- Basis
  - We present a different method that takes as a starting point the back-off scheme of Katz (1987).
  - …
- Contrast
  - Finally, while we have used our similarity model only for missing bigrams in a back-off scheme, Essen and Steinbiss (a992) used linear interpolation for alll bigrams.

Uses for information retrieval

“Representing Txt Chunks”. Erik. F. Tjong, Kim Sang and Jorn Veenstra

- Aim
  - In this paper we will examine seven different data representations for the problem of recognizing noun phrases
- Based on
  - Bosch 1998, Argamon et al. 1998
- Contrasts with
  - Ramshaw and Marcus 1995, Daelemans et al. 1999, Cardie and Pierce 1998
- Also cites
  - Abney 1991, Veenstra 1998

Dialogue

- Both dialogue and monologue need to be coherent/cohesive
  - Why is information conveyed
  - How is information structured
- More issues in dialogue
  - Turn taking
  - Grounding and repairing misunderstanding
  - Initiative and confirmation strategies

Segmenting Speech into Utterances

- Why is EOU detection harder than EOS?
- How does speech differ from text
- Single syntactic sentence may span several turns
  - A: We've got you on USAir flight 99
  - B: Yep
  - A: leaving on December 1
- Multiple syntactic sentences may occur in a single turn
  - A: We've got you on USAir flight 99 leaving on december 1. Do you need a rental car?
- Intonational definitions: intonational phrase, breath group, intonation unit

Turns and utterances

- Dialogue is characterized by turn taking
  - Who should speak next
  - When they should speak
- How do we know when a speaker is giving up or taking a turn? Holding the floor? Can we interrupt?

Dialogue acts and adjacency pairs

- Dialogue act: greeting, question, answer
- Adjacency pairs set up expectations
  - GREETING/GREETING
  - QUESTION/ANSWER
  - COMPLIMENT/DOWNPLAYER
  - REQUEST/GRANT
- Long silences are dispreferred
  - A: Is there something bothering you (1.0)
  - A: Yes or no? (1.5s)
  - A: EH?
  - B: No.

Intonational cues to turn taking

- Continuation rise (L-H%) holds the floor
- H-H% requests a response
  - L*H-H% yes/no question
  - H*H-H% highrise question contour
- Intonational contours signal dialogues acts in adjacency pairs

Initiative strategies

- System initiative
  S: Please give me your arrival city name
  U: Baltimore
  S: Please give me your departure city name…
- User initiative
  S: How may I help you?
  U: I want to go from Boston to Baltimore on November 8.
- Mixed initiative
  S: How may I help you?
  U: I want to go to Boston.
  S: What day do you want to go to Boston?

Grounding

- The HEARER must making it clear to the speaker if understanding has occurred
- How is this achieved?
  S: I can upgrade you to an SUV at that rate.
  - Continued attention
    (U gazes appreciatively at S)
  - Relevant next contribution
    U: Do you have RAV4 available
  - Acknowledgement/backchannel
    U: Ok/Great!
  - Request for repair
    U: I beg your pardon?

System misconceptions reflected in user response (Krahmer et al 99)

- Responses to incorrect verifications
  - Contain more words (or are empty)
  - Contain more repetition
- NO after incorrect verification vs. other yes/no questions
  - Has higher boundary tone
  - Longer duration
  - Longer pauses before and after
  - More additional words after it

Grounding and confirmation strategies

U: I want to go to Baltimore

- Explicit
  S: Did you say you want to go to Baltimore?
- Implicit
  S: Baltimore (H*L-L%)
  S: Baltimore? (L*H-H%)
  S: What time do you want to leave for Baltimore?
- No confirmation

Non-Understanding Error Recovery (Skantze '05)

- Collected human/human interactions
- Humans tend not to signal non-understanding:
  - O: Do you see a wooden house in front of you?
  - U: ASR: YES CROSSING ADDRESS NOW
    (I pass the wooden house now)
  - O: Can you see a restaurant sign?
- This leads to
  - Increased experience of task success
  - Faster recovery from non-understanding

Dialogue system challenges

- More complexities
  - Requires speech recognition
  - Turn taking, grounding, error-recovery
- But imagine the possible benefits
  - Talking robots?
  - Games
- Check out some talking heads demos
  - http://www.speech.kth.se/multimodal/

More NLP applications

- Speech to speech translation systems
  - Verbmobil (http://verbmobil.dfki.de/overview-us.html)
- Analyzing meetings
  - http://labrosa.ee.columbia.edu/mapmeet/
  - Identifying agreement and disagreement
  - Meeting segmentation
  - Browsing

References

- Mann, W.C., & Thompson, S.A. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. Text, **8** (3). 243-281.
- Jill Burstein, Daniel Marcu, and Kevin Knight (2003). Finding the WRITE Stuff: Automatic Identification of Discourse Structure in Student Essays. IEEE Intelligent Systems, pp. 32-39, Jan/Feb, 2003.
- Eleni Miltsakaki and Karen Kukich. The role of centering theory's rough-shift in the teaching and evaluation of writing skills. ACL 2000
- Regina Barzilay and Mirella Lapata. Modeling Local Coherence: An Entity-Based Approach. ACL 2005.

- S. Teufel, M. Moens, Argumentative classification of extracted sentences as a first step towards flexible abstracting. I. Mani, M. Maybury (eds.), *Advances in automatic text summarization*, 1999
- S. Teufel, A. Siddharthan, D. Tidhar. 2006a. Automatic classification of citation function. In proceedings of EMNLP-06
- E. Krahmer, M. Swerts, M. Theune, M. Weegels, Error Spotting in Human- Machine Interaction. In: *Proceedings Eurospeech'99*, September 1999, Budapest
- Exploring Human Error Recovery Strategies: Implications for Spoken Dialogue Systems. Skantze, G. (2005). *Speech Communication*, 45(3) (pp. 325-341).