Content Realization

Ling573
Systems and Applications
May 11, 2017

Roadmap

- Content realization

  - Linguistic quality
    - Improving referring expressions

  - Compression approaches
    - Heuristic techniques
    - Linguistically motivated methods
    - Learning compression

Content Realization: Referring Expressions

Referring to People in News Summaries

- Intuition:
  - Referring expressions common source of errors
  - References to people prevalent in news data, summaries
  - Information status constrains realization
  - Targeted rewriting can improve readability

Referring to People in News Summaries

- Intuition: (Nenkova, '08; Siddarthan et al, 2011)
  - Referring expressions common source of errors
  - References to people prevalent in news data, summaries
  - Information status constrains realization
  - Targeted rewriting can improve readability

- Approach:
  - Exploit information status distinctions
    - Automatically identified
  - Use to guide rule-based generation of referring expressions

Challenges

- Lack of training data:
  - No summary data labeled for information status

- Readers sensitive to referring expressions
  - Prior work on NP rewriting has shown mixed results
    - Some improvement, some failures

- Relies on potentially errorful coref, other processing

NP Rewrite: very good example

- While the British government defended the arrest, it took no stand on extradition of Pinochet to Spain, leaving it to the courts.

- While the British government defended the arrest in London of former Chilean dictator Augusto Pinochet, it took no stand on extradition of Pinochet to Spain, leaving it to British courts.

NP Rewrite: mixed example

- *Duisenberg* has said growth in the euro area countries next year will be about 2.5 percent, lower than the 3 percent predicted earlier.

- *Wim Duisenberg, the head of the new European Central Bank*, has said growth in the euro area countries next year will be about 2.5 percent, lower than just 1 percent in the euro-zone unemployment predicted earlier.

Information Status

- Build on three key distinctions:
  - Discourse-new vs discourse-old:
    - First mention handling vs others

  - Hearer-new vs hearer-old:
    - Distinguish well-known individuals from others
      - Don't waste space describing well-known individuals

  - Major vs minor character:
    - Salience of the person in the event

Corpus Analysis

- Assess relation between:
  - information status and referring expressions

		Discourse-new	Discourse-old
Name Form	Full name	0.97	0.08
	Surname only	0.02	0.87
	Other (e.g., Britney, JLo)	0.01	0.05
Pre-Modification	Any	0.51	0.21
	None	0.49	0.79
Post-Modification	None	0.60	0.89
	Apposition	0.25	0.04
	Relative clause	0.07	0.03
	Other	0.08	0.04
Any Modification (Either Pre- or Post-)	Some Modification	0.76	0.30
	No Modification	0.24	0.70

Siddharthan et al, 2011, p. 818, Table 1

Generating Discourse-New/Old

- If discourse-new,
  - If the NP head is a person name,
    - If appears with pre-modifier in text, write as:
      - Longest pre-modifier + full name

    - Else if it appears with an apposition modifier
      - Add that to the reference
  - Else don't rewrite

- Else use surname only

- Significantly preferred over original forms

Example Rewrite

Honecker has come under investigation for charges of corruption and living in luxury at the cost of the state. **Former East German leader Erich Honecker** may be moved to a monastery to protect him from a possible lynching by enraged citizens. As protests gathered strength last fall, **Erich Honecker, East Germany's longtime orthodox leader** "lost touch with reality," according to the man who succeeded him as Communist leader only to be ousted later. **Ousted East German leader Erich Honecker**, who is expected to be indicted for high treason, was arrested Monday morning upon release from a hospital and taken to prison.

Former East German leader Erich Honecker has come under investigation for charges of corruption and living in luxury at the cost of the state. **Honecker** may be moved to a monastery to protect him from a possible lynching by enraged citizens. As protests gathered strength last fall, **Honecker** "lost touch with reality," according to the man who succeeded him as Communist leader only to be ousted later. **Honecker**, who is expected to be indicted for high treason, was arrested Monday morning upon release from a hospital and taken to prison.

Siddharthan et al, 2011, p. 818, Table 1

Hearer & Salience

- Discourse-new status:
  - Obvious from summary

- How do we establish hearer or major/minor status?

- Categorize based on human summaries (gold)
  - Specifically by their referring expressions:
    - Hearer-old (i.e. familiar)
      - Title/role+surname or unmodified fullname
    - Major:
      - Referred to by name in some human summary of topic
        - 258 major/3926 minor by data

Training

- Trained classifiers to recognize
  - Using features in document set
    - Frequency, lexical, syntactic
  - Classifiers:
    - SVM, Decision trees

  - Hearer-New/Old: F-measure: 0.75 on both classes
  - Major/Minor: F: Major: 0.6; Minor: 0.98

  - All significantly better than baseline

Application

- If discourse-new and NP head is person name:
  - If MINOR:
    - Exclude name, use only role, modifiers, etc
  - If MAJOR and Hearer-Old:
    - Include name and role/temporal (only)
  - If MAJOR and Hearer-New:
    - Include name and role/temporal
    - Also include affiliation, post-mod (classifier)

- If discourse-old:
  - Surname ONLY

Evaluation

- Created (nearly) deterministic rule set
  - Based on information status classification
  - To rewrite referring expressions in extractive summaries

- Evaluated in paired preference tests over:
  - Original Extractive and Rewritten Summaries

- Where a preference was expressed,
  - Rewritten summaries rated as more coherent
  - Extractive rated as more informative
    - Why? Rewrite rules generally shrink rather than add content

Discussion

- Pros:
  - Intuitive, interpretable model
  - Solid results: ~0.75 accuracy, higher if humans agree
  - Often preferred to extract

- Cons:
  - Limited: only applies to person names
  - Error propagation: coreference, NP extraction
  - Ignores other aspects of realization, i.e. length

Summary

- Can identify particular correlates of readability scores

- Can automatically predict linguistic quality scores

- Build systems that focus on frequent violations
  - Yield systematic improvements in linguistic quality

Sentence Compression

Sentence Compression

- Main strategies:
  - Heuristic approaches
    - Deep vs Shallow processing
    - Information- vs readability- oriented

  - Machine-learning approaches
    - Sequence models
      - HMM, CRF
    - Deep vs Shallow information

  - Integration with selection
    - Pre/post-processing; Candidate selection: heuristic/learned

Shallow, Heuristic

- CLASSY 2006
  - Pre-processing! Improved ROUGE
    - Previously used automatic POS tag patterns: error-prone
- Lexical & punctuation surface-form patterns
  - "function" word lists: Prep, conj, det; adv, gerund; punct
- Removes:
  - Junk: bylines, editorial
  - Sentence-initial adverbials, conj phrase (up to comma)
  - Sentence medial adverbials ("also"), ages
  - Gerund (-ing) phrases
  - Rel. clause attributives, attributions w/o quotes
- Conservative: < 3% error (vs 25% w/POS)

Deep, Minimal, Heuristic

- ICSI/UTD:
  - Use an Integer Linear Programming approach to solve

- Trimming:
  - Goal: Readability (not info squeezing)
  - Removes temporal expressions, manner modifiers, "said"
    - Why?: "next Thursday"
  - Methodology: Automatic SRL labeling over dependencies
    - SRL not perfect: How can we handle?

D3

- Code/results due Sunday
  - Tag as D3

- Report due Tuesday morning
  - Tag as D3.1

- Presentations next week
  - Please respond to scheduling Doodle

Deliverable #4

- Final system
  - Continue system improvement
  - Add content realization

- Evaluation:
  - Devtest (2010)
  - Evaltest (2011)
    - New blind test
    - New document set: Documents from Gigaword
    - New evaluation models, ROUGE config file