Referring Expressions & Alternate Views of Summarization

Ling 573
Systems and Applications
May 24, 2016

Roadmap

- Content realization:
  - Referring expressions

- Alternate views of summarization:
  - Dimensions of the TAC model

  - Other methods, goals, data
    - Abstractive summarization

    - Summarizing reviews

    - Summarizing speech

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

- 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
      - E.g. President Obama, Kim Kardashian

  - Major vs minor character:
    - Salience of the person in the event
    - E.g., Former East German leader Erich Honecker vs
    - "the man who succeeded him as Communist leader only to be ousted later"

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

Summary Example

- 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.....

Summary Example

- **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 Honecke**r, who is expected to be indicted for high treason, was arrested Monday morning.....

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

Summary Example

- **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. **Honecke**r, who is expected to be indicted for high treason, was arrested Monday morning.....

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
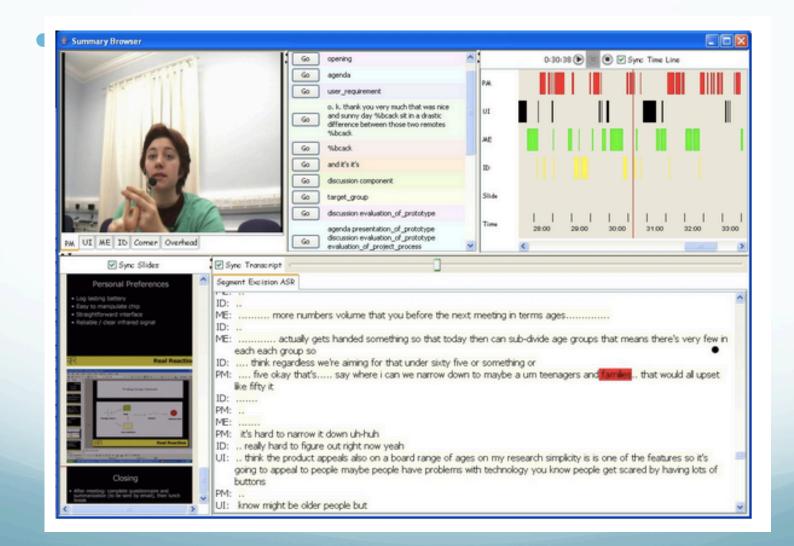
Alternate Views of Summarization

Dimensions of
TAC Summarization

- Use purpose: Reflective summaries

- Audience: Analysts

- Derivation (extactive vs abstractive): Largely extractive

- Coverage (generic vs focused): "Guided"

- Units (single vs multi): Multi-document

- Reduction: 100 words

- Input/Output form factors (language, genre, register, form)
  - English, newswire, paragraph text

Meeting Summaries

- What do you want out of a summary?

Example

Meeting Summaries

- What do you want out of a summary?

- Minutes?

- Agenda-based?

- To-do list

- Points of (Dis)agreement

Dimensions of Meeting Summaries

- Use purpose: Catch up on missed meetings

- Audience: Ordinary attendees

- Derivation (extactive vs abstractive): Extractive or Abstr.

- Coverage (generic vs focused): User-based?

- Units (single vs multi): Single event

- Reduction: ?

- Input/Output form factors (language, genre, register, form)
  - English, speech+, lists/bullets/todos

Examples

- Decision summary:
  - 1. The remote will resemble the potato prototype
  - 2. There will be no feature to help find the remote when it is misplaced;
  - instead the remote will be in a bright colour to address this issue.
  - 3. The corporate logo will be on the remote.
  - 4. One of the colours for the remote will contain the corporate colours.
  - 5. The remote will have six buttons.
  - 6. The buttons will all be one colour.
  - 7. The case will be single curve.
  - 8. The case will be made of rubber.
  - 9. The case will have a special colour.

Examples

- Action items:
  - They will receive specific instructions for the next meeting by email.
  - They will fill out the questionnaire.

Examples

- Abstractive summary:
  - When this functional design meeting opens the project manager tells the group about the project restrictions he received from management by email. The marketing expert is first to present, summarizing user requirements data from a questionnaire given to 100 respondents. The marketing expert explains various user preferences and complaints about remotes as well as different interests among age groups. He prefers that they aim users from ages 16-45, improve the most-used functions, and make a placeholder for the remote...

Abstractive Summarization

- Basic components:
  - Content selection
  - Information ordering
  - Content realization
    - Comparable to extractive summarization

- Fundamental differences:
  - What do the processes operate on?
    - Extractive? Sentences (or subspans)
    - Abstractive? Major question
      - Need some notion of concepts, relations in text

Levels of Representation

- How can we represent concepts, relations from text?
  - Ideally, abstract away from surface sentences

- Build on some deep NLP representation:

  - Dependency trees: (Cheung & Penn, 2014)

  - Discourse parse trees: (Gerani et al, 2014)

  - Logical Forms

  - Abstract Meaning Representation (AMR): (Liu et al, 2015)

Representations

- Different levels of representation:
  - Syntax, Semantics, Discourse

- All embed:
  - Some nodes/substructure capturing concepts
  - Some arcs, etc capturing relations
  - In some sort of graph representation (maybe a tree)

- What's the right level of representation??

Typical Approach

- Parse original documents to deep representation

- Manipulate resulting graph for content selection
  - Splice dependency trees, remove satellite nodes, etc

- Generate based on resulting revised graph

- All rely on parsing/generation to/from representation

AMR 2

- AMR Bank: (now) ~40K annotated sentences

- JAMR parser: 63% F-measure (2015)
  - Alignments b/t word spans & graph fragments

- Example: "I saw Joe's dog, which was running in the garden."

Liu et al, 2015.