CLMA
Thesis-in-Progress

Robust Minimal Recursion Semantics and Aspect-Oriented Multi-Document Update Text Summarization

Prescott Klassen
University of Washington
DELPH-IN Summit
June 25, 2011

Hypotheses

Transforming query and document text into Robust Minimal Recursion Semantics (RMRS) for the generation of an automated aspect-oriented multi-document update text summary will improve its responsiveness and readability.

Text Analysis Conference (TAC)

- Annual Evaluation Workshops
  - Conducted by *National Institute of Standards and Technology* (NIST)
  - Sponsored by *US Department of Defense*
- Previously know as *Document Understanding Conference* (DUC)
- TAC 2011 has 3 Tasks:
  - Knowledge Base Population
  - Recognizing Textual Entailment
  - Summarization
- TAC 2011 Guided Summarization Task had 23 participants

TAC 2011 Guided Summarization Schedule

June 3	Deadline for TAC 2011 track registration
July 1	Release of test data
July 17	Deadline for submission
Sept 7	Release of individual evaluated results
Sept 25	Deadline for TAC 2011 workshop proposals
Oct 25	Deadline for system reports
Nov 14-15	TAC 2011 Workshop

TAC 2011 Guided Summarization

- 2 100 word summaries (truncated for evaluation)
  - Initial summary of 10 document corpus
  - Update summary of 10 document corpus
- 5 pre-determined topic categories
- Category specific aspects (WHAT, WHEN, WHERE, etc)
- Doc set is newswire portion of TAC Knowledge Base Population 2010 Source Data (Distributed by LDC)
- Human evaluation (Pyramid, Responsiveness)
- Automated evaluation (AESOP for Readability)
- 2 runs per team

Example Category & Aspects

1. Accidents and Natural Disasters:

1.1 WHAT: what happened
1.2 WHEN: date, time, other temporal placement markers
1.3 WHERE: physical location
1.4 WHY: reasons for accident/disaster
1.5 WHO_AFFECTED: casualties (death, injury), or individuals otherwise negatively affected by the accident/disaster
1.6 DAMAGES: damages caused by the accident/disaster
1.7 COUNTERMEASURES: countermeasures, rescue efforts, prevention efforts, other reactions to the accident/disaster

Example Topic Statement

```
<topic id = "D0906B" category = "1">

        <title> Rains and mudslides in Southern California </title>

        <docsetA id = "D0906B-A">
                <doc id = "AFP_ENG_20050110.0079" />
                …
        </docset>

        <docsetB id = "D0906B-B">
                <doc id = "AFP_ENG_20050221.0700" />
                …
        </docsetB>

</topic>
```

Example Document

```
<DOC>
<DOCNO> APW19990331.0161 </DOCNO>
<DATE_TIME> 1999-03-31 16:21:02 </DATE_TIME>
<BODY>
<CATEGORY> usa </CATEGORY>
<HEADLINE>  NYC Cops Charged in Shooting Death </HEADLINE>
<TEXT>
<P>
        NEW YORK (AP) -- Four white police officers were charged with
murder Wednesday for killing an unarmed African immigrant in a hail
of 41 bullets -- a shooting that has led to months of protests and a
painful examination of police tactics and race relations.
</P>
<P>
        Officers Kenneth Boss, Sean Carroll, Edward McMellon and Richard
Murphy pleaded innocent in a Bronx courtroom to second-degree
murder. They could get 25 years to life in prison on the murder
charges.
</P>
<P>
        Amadou Diallo, a 22-year-old street vendor from Guinea with no
criminal record, was shot 19 times Feb. 4 in his vestibule by
members of an elite street-crime unit looking for a rape suspect.
Through their lawyers, the officers have said they thought Diallo
had a gun.
</P>
<P>
        The slaying has frayed Mayor Rudolph Giuliani's already strained
relationship with the black community, and many have accused him of
failing to understand the depth of New Yorkers' anger.
</P>
<P>
        Giuliani said Wednesday: ``We should allow the criminal justice
system to now operate.''
</P>
```

TAC 2010 Evaluation

- All runs evaluated manually
  - Responsiveness (1 -5)
  - Readability (1 – 5)
  - Pyramid (Passonneau et al, 2005)
- 23 Participants, 41 runs
- Baseline_1: first 100 words from most recent document
- Baseline_2: summary generated by MEAD
- System score = mean of all summaries
- System ranking (ANOVA, multiple comparison)

Results

	Model	Automated
Pyramid Initial	0.785	0.302
Pyramid Update	0.673	0.199
Responsiveness Initial	4.761	2.565
Responsiveness Update	4.712	2.102
Readability Initial	4.908	2.837
Readability Update	4.821	2.765

- Delta between human and automated
  - All model summaries out-performed best automated
  - TAC 2009 paper: HEXTAC the Creation of a Manual Run (Genest et al, 2009)

TAC 2010 Themes

- Almost all participants employing extractive approaches
  - Exception**(Genest & Lapalme, University of Montreal) *Text Generation for Abstractive Summarization*
- Singular Value Decomposition, Topic Guided Manifold Ranking, Graph-based measurements of topic/term overlap and density, "Wikification", etc
- Best system is still CLASSY (Clustering, Linguistics, and Statistics for Summarization Yield)

Genest & Lapalme

- Information Items based on verbs (triples plus date and location are extracted using dependency parser)
  - Use aspects for InIt selection
  - Rank InIt by salience
  - Order InIts in logical way
  - Generate sentences using extractive fragments and SimpleNLG
- Evaluated middle-of-the-pack
- For TAC 2011 will be investigating a more formal semantic representation

Proposed Approach for Master's Thesis

- RMRS representation (semantics over surface)
- Event detection, relation extraction via RMRS structures
- RMRS fragments from category/aspect TAC templates (Q/A strategy)
- Sentence compression/fusion via RMRS comparison and frequency-based fragment/ relation selection
- Generation from RMRS