Sheffield-Trento System for Sentiment and Argument Structure Enhanced Comment-to-Article Linking in the Online News Domain

Ahmet Aker, Fabio Celli*, Adam Funk, Emina Kurtic, Mark Hepple, Rob Gaizauskas
University of Sheffield
University of Trento*

Motivation

- Commenting to online news articles has become a much used way of communication between online media outlets and their readers
- At present it is not easy to determine which parts of the news article a comment relates to
- However, knowing that relationship is a crucial step in higher level comment processing tasks, like automatic comment summarization
  - That relationship can be used to group topically related contributions to conversations and representative comments from the groups can be used to build summaries
- Each related article (a segment of the article) and comment can be enriched with “argument” and “sentiment” information
  - Help to know whether a particular comment agrees or disagrees with the article or if it is in favour of the opinions voiced in the article or not

Outline

- Task
- Methods
- Evaluation & Results
- Conclusion

Task

- Article *A* is divided into n segments S(A) = s1 , ..., sn, we treat article sentences as segments
- *A* is also associated with a set of comments C(A) = c1 , ..., cl
- The task is to link comments $c \in C(A)$ with article segments $s \in S(A)$
- We express the strength of link between a comment *c* and an article segment *s* as their linking score (*Score*)
- A comment *c* and an article segment *s* are linked if and only if their *Score* exceeds a threshold, which we experimentally optimized
- *Score* has the range [0, 1], 0 indicating no linking and 1 defining a strong link
- For the argument structure detection, we assign each segment-comment pair (s, c) to agree, disagree or neutral categories.
- Likewise, the sentiment assignment classifies the segment-comment pairs as in favour, against and indifferent.

Method – linking

- Pairing every c ∈ C(A) with every s ∈ S(A)
- Extracting features
- Quote

  quoteScore = len(quote)/len(S)

  with len returning the length of the argument
- *c ∈ C(A)* and *s ∈ S(A)* are considerred as pairs if *s ∈ S(A)* contains at least 10 words and *quoteScore >= 0.5*. In this case nothing else is computed.
- Otherwise computing futher features:

Method – linking (cont.)

- Cosine (f1): $V(s) * V(c) / |V(s)| * |V(c)|$
- Dice (f2): $2 * len(I(s, c)) / len(s) + len(c)$, $I(.,.)$ is the intersection of words/terms
- Jaccard (f3): $len(I(s, c)) / len(U(s, c))$, $U(.,.)$ is the union of words/terms
- NE-overlap (f4): $len(I(s, c)) / len(U(s, c))$, instead of words/terms NEs are used

Method – linking (cont.)

- DISCO semantic (f5): DIStributionally similar words using CO-occurrences) assumes words with similar meaning occur in similar context. Context is derived from large corpora such as Wikipedia and represented as vector.
- f1 to f5 are computed only when *quoteScore < 0.5.* We combine them using:

Score = w1 * f1 + w2*f2 + w3*f3 + w4*f4* w5*f5

- w1 to w5 are trained using linear regression + training data

Method – linking (cont.)

- Training data:
- Total 3362 news articles collected automatically from The Guardian using an in-house tool
- For each article *A* paired each $s \in S(A)$ with $len(s) >= 10$ every $c \in C(A)$ and computed only the *quoteScore*
- If *quoteScore* >= *0.5* the pair <*s,c*> was taken as positive pair
- Total positive pairs: 43300
- Also collected negative pairs by pairing an *s* from A with an *c* from A'
- Total negative pairs: 43300
- quoteScore is the outcome for each <*s,c*>

Method – argument & sentiment extraction

- Trained regression models for argument extraction on 2260 comments extracted from CorEA (Celli et al., 2014), an Italian news blog corpus manually annotated with arguments (1000 disagreement, 783 agreement and 215 neutral) labels
  - A feature vector with 84 shallow statistical dimensions about text encoding, characters, ngrams, punctuation, numbers, parentheses, uppercases, lowercases, word freq, word length, string similarity, emoticons, parentheses, tf*idf, similarity of uppercase words and sine of the frequency of word pairs
- For sentiment extraction we used an existing GATE pipeline that combines named entity recognition, event detection, and sentiment detection Maynard and Funk, 2012; Maynard et al., 2014

Evaluation & Results

- Performance of our system (USFD UNITN) was evaluated within the MultiLing 2015 Online Forum Summarization (OnForumS) task and reported relative to a baseline system and 3 further competing systems
- The evaluation was performed with English and Italian data and results are reported in precision
- Each participant was allowed to submit two runs
- Our runs differed in how we set a threshold for linking similarity: first run was set to 0.3, second run to 0.5 – anything above the threshold was regarded as linked
- For Italian our second run with the threshold 0.5 was not considered
- For argument and sentiment extraction we only participated on the English data

Evaluation & Results (cont.)

- Performance for linking (*left is for English and right for Italian*)

Participant and run	Precision score
BASE-overlap	0.928
USFD_UNITN-run2	0.892
JRC-run1	0.857
UWB-run1	0.851
JRC-run2	0.8291
USFD_UNITN-run1	0.818
BASE-first	0.738
CIST-run2	0.709
CIST-run1	0.702

Participant and run	Precision score
BASE-overlap	0.590
UWB-run1	0.25
USFD_UNITN-run1	0.2
JRC-run1	0.152
CIST-run1	0.084
CIST-run2	0.33
BASE-first	0.010

Evaluation & Results (cont.)

- Performance for argument (left) and sentiment extraction (right) – both results are for English only

Participant and run	Precision score
CIST-run2	0.990
CIST-run1	0.988
UWB-run1	0.974
BASE-first	0.915
JRC-run2	0.896
USFD_UNITN-run1	0.891
JRC-run1	0.884
BASE-overlap	0.881
USFD_UNITN-run2	0.859

Participant and run	Precision score
CIST-run1	0.946
CIST-run2	0.933
BASE-first	0.927
BASE-overlap	0.922
UWB-run1	0.897
JRC-run2	0.895
USFD_UNITN-run2	0.885
USFD_UNITN-run1	0.880
JRC-run1	0.874

Conclusions

- We report the details of the Sheffield-Trento system for argument structure and sentiment enhanced comment-to-article linking in the online news domain for English and Italian
- The system links readers' comments to news article sentences that triggered them and is based on a combination of quotation detection and a combined similarity computation between comment and article sentence.
- In addition argument structure (agreement, disagreement, neutral) and sentiment (in favour, agianst, indifferent) are assigned to comment-article sentence pairs
- For the linking task in English our system outperforms all other competing systems
- For Italian linking as well as for argument structure and sentiment assignment in both languages, there is a substantial scope for improvement compared to other competing systems