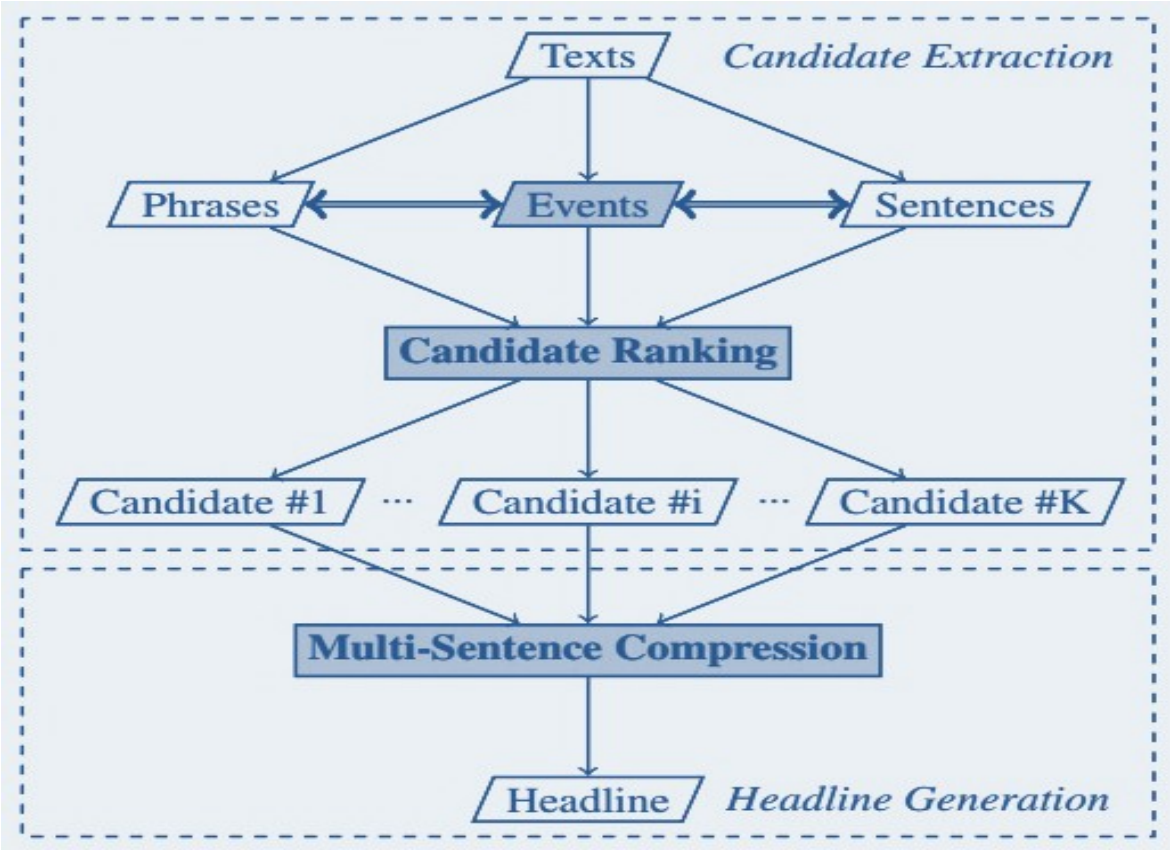
Event Driven Headline Generation

Rui Sun† , Yue Zhang‡ , Meishan Zhang‡ and Donghong Ji† † Computer School, Wuhan University, China ‡ Singapore University of Technology and Design



M.Arvind
12385
CS671A

Introduction

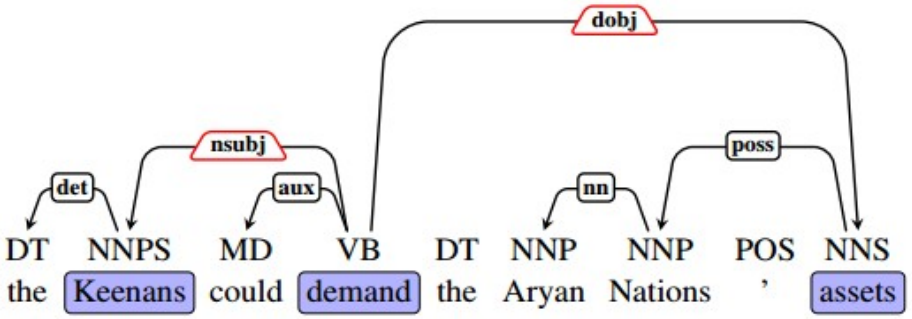
- Given an input document, the system identifies a key event chain by extracting a set of structural events that describe them.
- Then a novel multi-sentence compression algorithm is used to fuse the extracted events, generating a headline for the document.
- This model can be viewed as a novel combination of extractive and abstractive headline generation, combining the advantages of both methods using event structures.
- Headline generation models consist of two steps: **Candidate extraction** and **Headline generation**

Candidate extraction in the proposed model

- Candidate event extraction is performed on a bipartite graph, where the two types of nodes are lexical chains and events.

- An event is a tuple (S, P, O), where S is subject, P is predicate and O is object.

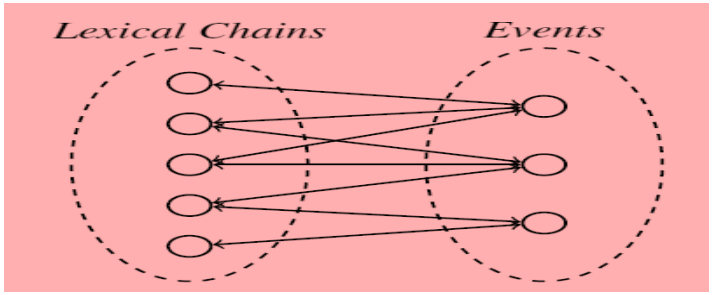
- **Extracting events:** Stanford dependency parser was used to obtain the Stanford typed dependency structures of the sentence.



- **Extracting lexical chains :** Lexical chains are used to link semantically related words and phrases.

Mutual Reinforcement Principle

- **Mutual Reinforcement Principle** is applied to jointly learn chain and event salience on the bipartite graph for a given input.

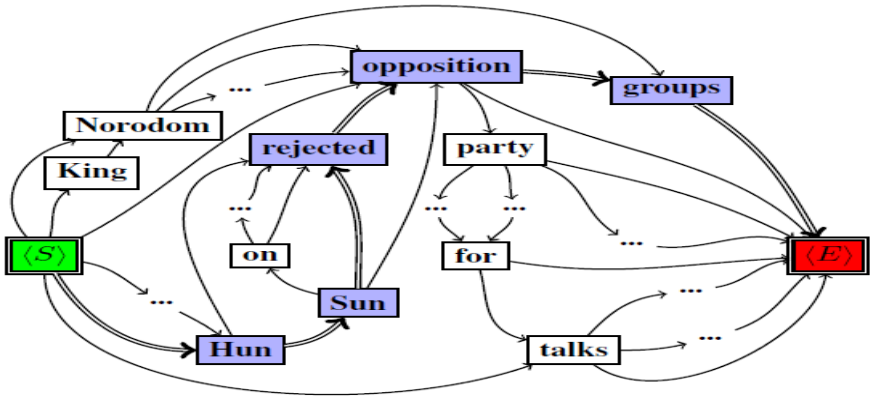


- Suppose that there are n events: {e1, · · · , en} and m lexical chains: {l1, · · , lm} :

$$sal(e_i) \propto \sum_{j=1}^{m} r_{ij} * sal(l_j) \qquad sal(l_j) \propto \sum_{i=1}^{n} r_{ij} * sal(e_i)$$

$$r_{ij} = \frac{w(l_j) * w(e_i)}{A} \qquad w(l_j) = \sum_{w \in l_j} sal_{abs}(w) \qquad w(e_i) = \sum_{s \in Sen(e_i)} sal_{ext}(s)$$

Word-Graph Construction



- Salience information is introduced into the calculation of the weights of vertices.

$$w(V_i) = \sum_{e \in CE} sal(e) * \exp\{-\text{dist}(V_i.w, e)\}$$

- Edge weights are computed as follows:

$$w'(E_{ij}) = \sum_s \text{rdist}(V_i.w, V_j.w)$$

$$w(E_{ij}) = \frac{w(V_i)w(V_j).w'(E_{ij})}{w(V_i)+w(V_j)}$$

Scoring Method

- The key to our MSC model is the path scoring function.

- The overall score of a path is compute by:

$$\text{score}(p) = edge(p) + \lambda * flu(p)$$

$$\text{edge}(p) = \frac{\sum_{E_{ij} \in p} \ln\{w(E_{ij})\}}{n}$$

$$\text{flu}(p) = \frac{\sum_{i} \ln\{p(w_i / w_{i-2} w_{i-1})\}}{n}$$

Beam Search Algorithm for Headline generation

Input: $G \leftarrow (\mathcal{V}, \mathcal{E})$, LM, B
Output: *best*
candidates $\leftarrow$ { {$\langle S \rangle$} }
loop do
  beam $\leftarrow$ { }
  for each *candidate* **in** *candidates*
    if *candidate* **endwith** $\langle E \rangle$
      ADDTOBEAM(*beam*, *candidate*)
      continue
    for each V_i **in** $\mathcal{V}$
      candidate $\leftarrow$ ADDVERTEX(*candidate*, V_i)
      COMPUTESCORE(*candidate*, *LM*)
      ADDTOBEAM(*beam*, *candidate*)
    end for
  end for
  candidates $\leftarrow$ TOP-K(*beam*, *B*)
  if *candidates* **all endwith** $\langle E \rangle$: **break**
end loop
best $\leftarrow$ BEST(*candidates*)

Parameter value optimization

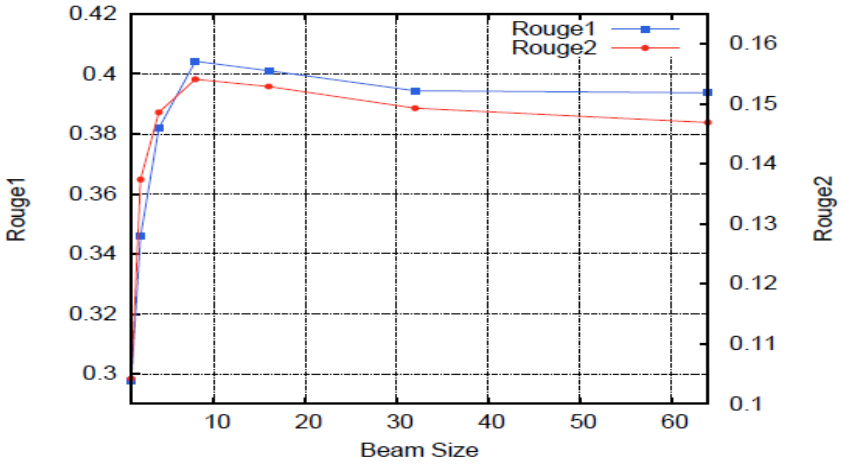


Figure 6: Results with different beam sizes.

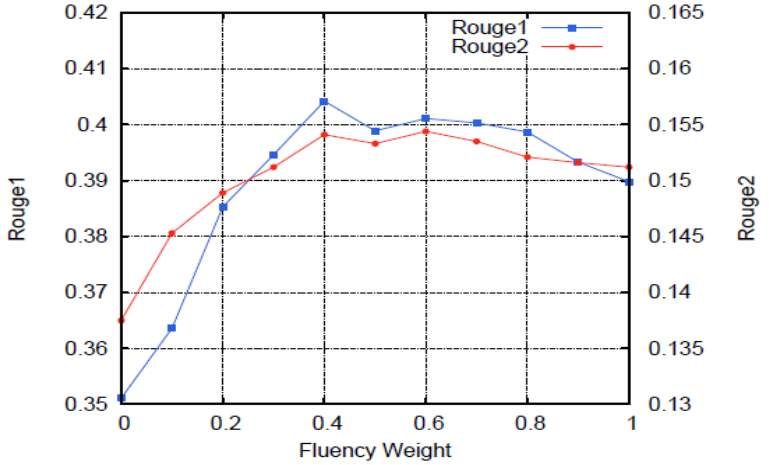


Figure 7: Results using different fluency weights.

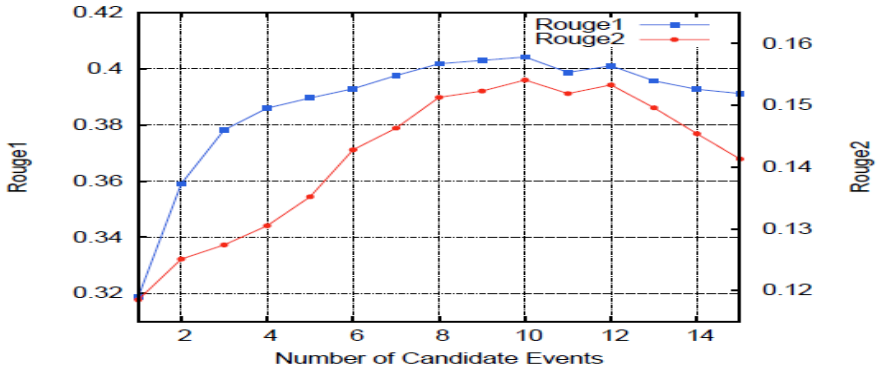


Figure 8: Results using different numbers of candidate events.

Comparison of headlines generated by the different methods

Method	Generated Headlines
Reference	Honduras, other Caribbean countries brace for the wrath of Hurricane Mitch
SentRank	Honduras braced for potential catastrophe Tuesday as Hurricane Mitch roared through northwest Caribbean
PhraseRank	Honduras braced catastrophe Tuesday Hurricane Mitch roared northwest Caribbean
EventRank	Honduras braced for Hurricane Mitch roared through northwest Caribbean
Reference	At Ibero-American summit Castro protests arrest of Pinochet in London
SentRank	Castro disagreed with the arrest Augusto Pinochet calling international meddling
PhraseRank	Cuban President Fidel Castro disagreed arrest London Chilean dictator Augusto Pinochet
EventRank	Fidel Castro disagreed with arrest in London of Chilean dictator Augusto Pinochet

Conclusion

- Experimental results demonstrate that event-driven model can achieve better results than extractive and abstractive models, and the proposed graph-based MSC model can bring improved performances compared with previous MSC techniques. Our final event-driven model obtains the best result on this dataset.

- The proposed graph-based MSC model is not limited to our event-driven model. It can be applied on extractive and abstractive models as well