Algorithm

Minh Lê

February 8, 2017

foreach *sentence* $s \in$ *document* **do**
  compute coreference score s(mention, candidate) for all entities in the sentence;
  sort mentions in descending order of score (easy-first);
  foreach *mentions* $m \in$ *sorted mentions* **do**
    link m to the cluster with highest score;
    foreach *event* v_m *that takes* m *as participant/argument* **do**
      compute relation scores between v_m and events related to the cluster;
      create relations that are greater than a threshold;
    end
  end
end

Algorithm 1: Joint entity coreference resolution and event relation identification.

Scores are computed based on the neural network model of [Clark and Manning, 2016].

Consider this excerpt from CoNLL-2012 for example:

"The film is a copy of the film which was taken after [[his]$_1$ execution]$_2$ on **the morning of the blessed Eid** . [...] A new video recording depicting the corpse of the late Iraqi president [Saddam Hussein]$_3$ after [[his]$_4$ execution]$_5$ on **the morning of the first day of the Eidul Adha** was broadcasted " (a2e_0010_part_000)

By the time the system starts processing the last sentence ("A new video..."), it already stores a cluster for Saddam Hussein and an event of his execution. The system proceeds on by:

1. Link [Saddam Hussein] to the Saddam Hussein cluster
2. Link [his] to Saddam Hussein
3. Since [execution] is an event related to [his], it is taken into consideration
4. Connect [execution] to Saddam Hussein's execution with relation type Identity
5. Consider [the morning of the first day...], since it is the same morning of the same execution, it corefers to [the morning of the blessed Eid]

In this example, we see how entity coreference resolution and event relation identification work in tandem to uncover the semantic structure of a story. Without the help of event relation, it's very hard to work out the coreferring relation between the two [morning]s – the overlapping content was small and broad while the two names appear different. On the other hand, considering pairs of events that have coreferring participants/arguments is a good way to narrow down the search space while maintaining good coverage.

Consider not only Identity event relation but also weaker types: causation, subevent, before/after

References

[Clark and Manning, 2016] Clark, K. and Manning, C. D. (2016). Improving Coreference Resolution by Learning Entity-Level Distributed Representations. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 643–653.