

**Document Summarization for Answering Non-Factoid Queries**

**Abstract:**

We formulate a document summarization method to extract passage-level answers for non-factoid queries, referred to as answer-biased summaries. We propose to use external information from related Community Question Answering (CQA) content to better identify answer bearing sentences. Three optimization-based methods are proposed: (i) query-biased, (ii) CQA-answer-biased, and (iii) expanded-query-biased, where expansion terms were derived from related CQA content. A learning-to-rank-based method is also proposed that incorporates a feature extracted from related CQA content. Our results show that even if a CQA answer does not contain a perfect answer to a query, their content can be exploited to improve the extraction of answer-biased summaries from other corpora. The quality of CQA content is found to impact on the accuracy of optimization-based summaries, though medium quality answers enable the system to achieve a comparable (and in some cases superior) accuracy to state-of-the-art techniques. The learning-to-rank-based summaries, on the other hand, are not significantly influenced by CQA quality. We provide a recommendation of the best use of our proposed approaches in regard to the availability of different quality levels of related CQA content. As a further investigation, the reliability of our approaches was tested on another publicly available dataset.