Abstract

Text summarization is the task of generating a shortened version of the original text where core ideas of the original text are retained. In this work, we focus on query focused summarization. The task is to generate the summary from a set of documents which answers the query. Query focused summarization is a hard task because it expects the summary to be biased towards the query and at the same time important concepts in the original documents must be preserved with high degree of novelty.

Graph based ranking algorithms which use *biased random surfer model* like Topic-sensitive LexRank have been applied to query focused summarization. In our work, we propose look-ahead version of Topic-sensitive LexRank. We incorporate the option of look-ahead in the random walk model and we show that it helps in generating better quality summaries.

Next, we consider assessment of machine translation. Assessment of a machine translation output is important for establishing benchmarks for translation quality. An obvious way to assess the quality of machine translation is through the perception of human subjects. Though highly reliable, this approach is not scalable and is time consuming. Hence mechanisms have been devised to automate the assessment process. All such assessment methods are essentially a study of correlations between human translation and the machine translation.

In this work, we present a scalable approach to assess the quality of machine translation that borrows features from the study of writing styles, popularly known as Stylometry. Towards this, we quantify the characteristic styles of individual machine translators and compare them with that of human generated text. The translator whose style is closest

to human style is deemed to generate a higher quality translation. We show that our approach is scalable and does not require actual source text translations for evaluation.