Modelling knowledge summarization by evolving fuzzy grammar

ABSTRACT

Summarized text is a simplified and condensed version of the original text containing highlighted information to help the audience get the gist in a short period of time. Typically, text summarization produces abstract or a paragraph-like outputs by omitting details and irrelevant information. However, the text summary can also be produced in a visualized form, such as a chart, graph or table representing a collection of similar cases. The visualized version generates a statistical-like presentation, which often involves numerical and ordinal observation of the gathered knowledge from the text. This requires lexical syntactic understanding of the text. Essential to achieve this goal is topic identification, message analysis/interpretation and knowledge summarization generation. The objective of this study is to model knowledge summarization problem using the evolving fuzzy grammar technique and we focus on metadata generation for producing visualized knowledge summarization. The process comprises of: (i) identifying the underlying structure of the texts for knowledge summarization, (ii) represent the identified knowledge for summarization manipulation and (iii) presentation of the summarized knowledge. A prototype called FTCat© is developed as a proof of concept and we demonstrate its practicality in summarizing news reports.

Keyword: Text summarization; Evolving fuzzy grammar; Text mining.