

Ordering Circumstantials for Multi-Document Summarization

Michael Elhadad*, Yael Netzer*
Regina Barzilay**, Kathleen McKeown**

(*) Dept of Computer Science
Ben Gurion University of the Negev
84105 Beer Sheva, Israel
(elhadad,yaeln)@cs.bgu.ac.il

(**) Dept of Computer Science
Columbia University
New York, NY 10027, USA
(regina,kathy)@cs.columbia.edu

Abstract

We address the problem of ordering several circumstantials when generating or revising a clause. This problem occurs in the context of a multi-document summarization system that relies on language generation to incrementally reformulate the wording of fragments of sentences extracted from the documents. We present the results of an extensive corpus analysis of the relative position of different types of circumstantials. Our approach learns a set of rules using parameters that can be effectively used by our system. Evaluation indicates that these rules, which we have implemented in our text generator, attain a high level of precision (95.4% over a baseline of 78.6%).