

Technical Report No. DCSE/TR-2005-01
April 2005

Text Summarization via Latent Semantic Analysis and Anaphora Resolution

Josef Steinberger

Abstract

The information-overloading problem can be reduced by automatic summarization. I present the state of the art in this area. Special attention is paid to evaluation methods. Then I describe principles of Latent Semantic Analysis (LSA) and basics of Anaphora Resolution (AR). Further, I discuss a new summarization approach that uses LSA and AR. A summary evaluation method based on LSA is proposed as well. Finally, I show preliminary results and discuss further research directions. The final goal is to integrate the summarizer into a natural language processing system capable of searching and presenting documents in a concise and coherent form.

This work was supported by grants MSM 235200005, ME494 and FRVŠ 1326/2005.

Copies of this report are available on <http://www.kiv.zcu.cz/publications/> or by surface mail on request sent to the following address:

University of West Bohemia in Pilsen
Department of Computer Science and Engineering
Univerzitni 8
30614 Pilsen
Czech Republic

Copyright © 2005 University of West Bohemia in Pilsen, Czech Republic