

Generating Stories from Different Event Orders: A Statistical Approach

M.Sc. Thesis of Anastasia Shimorina

30 March 2016

Abstract

This research presents the strategy how to find statistically significant language patterns and make use of them in generating new texts. Namely, the temporal relations in narrative are explored. To investigate the narrative temporal structure, a specially designed corpus of stories is used. For each story, main events and their chronological and discourse-level orders are known. This corpus allows us to identify common temporal models for specific orders of events at the discourse level. The Conditional Random Fields method is applied to predict the best temporal model for each event order. The acquired temporal models are used in a template-based natural language generation system which outputs stories. The stories generated by the system are evaluated by human subjects. We demonstrate that stories generated according to the acquired temporal models are adequately interpreted by humans. Our results show that it is possible to apply machine learning methods to acquire useful heuristics for narrative generation.