Indicative Abstractive Summaries of Meetings

Thomas Kleinbauer and Stephanie Becker and Tilman Becker

DFKI GmbH, Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany
<firstname.lastname>@dfki.de

Abstract. We present ongoing research on the generation of indicative meeting abstracts supporting quick relevance assessment of meetings.

1 Introduction

Applying extractive summarization approaches to meetings may lead to results vastly different from hand-written summaries: a concatenation of speaker contributions taken from the original dialog and made from each speaker's own perspective is likely to contain first-person wording, speech disfluencies and–in automated settings–speech recognition errors. Our approach uses NLG techniques to address these issues.

2 Representing Meeting Contents

In the AMI project[1], circa 100 hours of meetings have been recorded, annotated and stored in a freely available multimodal corpus [1]. In addition to multiple video and audio streams, a number of annotations are included in the corpus and one of the goals of the AMI project is to develop automatic recognition systems for all of these annotation layers. Additionally, we have annotated a small subset of the AMI corpus with categories from the AMIMATTER domain ontology [2] to represent the propositional content of speaker utterances.

We currently concentrate on three annotation layers: topic labels, dialog acts and propositional content. For topic annotation, the recordings were split into larger segments and labeled with predefined topics from the design scenario, e. g. "presentation of prototype(s)". These segments are used by our system as the basic structuring unit for the summaries yielding one sentence per topic.

On the basis of the segments induced by dialog acts, we perform a frequency analysis of all annotated ontology instances and select the three items that occur most often. These items are integrated into the summary sentence (s. fig. 1).

3 Text Generation

The generation of the text is done in a three-step pipeline: analysis of meeting annotation layers, sentence planning, and surface realization. In the first step,

[1] http://www.amiproject.org

"The meeting was opened and the meeting group talked about the user interface, the remote control and the design. They debated the costs, the company and the project while discussing the project budget. The signal, the remote control and the beep were mentioned afterwards. They talked about meeting before closing the meeting."

Fig. 1. Example of a generated meeting summary.

information drawn from the annotation layers is transformed into expressions in a propositional logic-like formalism. These assertions are used as a knowledge base by the sentence planner PrePlan[3] which additionally uses a library of plan operators, each of which encodes strategies how to reach a given goal. PrePlan successively finds matching plan-operators until all goals and subgoals are resolved. The outcome of this process is an XML-encoded description of instructions in a logical form which is passed to the surface realizer, NipsGen [4], a template-based generator. NipsGen converts the semantic input into typed feature structures which are then transformed into a natural language utterance. The actual syntax tree is constructed from a derivation tree for the XTAG-grammar [5] which in turn is created using transformation rules which are applied to the input structure. The generation of the correct morphological inflections is achieved by percolating the morphological features through the XTAG tree and a lexicon look-up of the correct inflections for all lexical leaves.

4 Current and Future Work

We are currently developing the summarization system further by adding more annotation layers to the processing pipeline. In parallel, we have also started work on multimedia summaries that will combine text with pictures from the video signals and links into the meetings. In addition, work has begun on an extrinsic evaluation. Here, we evaluate how summaries can improve access to recorded meetings when available in a meeting browser.

References

1. McCowan, I., Carletta, J., Kraaij, W., et. al.: The Ami meeting corpus. In: Proceedings of the Measuring Behavior 2005 symposium on Annotating and Measuring Meeting Behavior, Wageningen, NL (2005)
2. Lochert, H., Alexandersson, J., Becker, T., Kleinbauer, T.: Ontomatters–ontology-based representation of multi-modal multi-party interaction. In: 2nd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms, Royal College of Physicians, Edinburgh, UK (2005)
3. André, E.: Ein planbasierter Ansatz zur Generierung multimedialer Präsentationen. Infix, St. Augustin (1995)
4. Engel, R.: SPIN: A semantic parser for spoken dialog systems. In: Proceedings of Fifth Slovenian and First International Language Technology Conference. (2006)
5. XTAG Research Group: A lexicalized tree adjoining grammar for english. Technical Report IRCS-01-03, IRCS, University of Pennsylvania (2001)