

Student Research Symposium 2004

Nikesh Garera

Towards a Personal Briefing Assistant

The preparation of summary reports from raw information is a common task in research projects. A tool that highlights useful items for a summary would allow report writers to be more productive, by reducing the time needed to assess individual items. It has further potential benefit in that it can be used to create user-specific or audience-specific digests. In the latter case, multiple tailored reports could in principle be generated from the same input information. With this motivation, we present a design of an adaptive system that learns to extract important items from weekly interviews by observing the behavior of human summary authors.

Our application scenario involves a report writer producing digests on a week-to-week basis and our goal is to make this person more efficient over time. We propose to do this by presenting the writer with successively better ordered lists of items (such that digest-worthy items appear at the top of the ordered list).

We identified salient features used for learning in this new domain by studying the corpus of project interviews. This corpus consisted of weekly progress interviews of project members collected over a period of 4 months. The features were then annotated in the corpus and were used as parameters in a regression model. This model is incrementally trained from user input and is used to reorder items in successive weeks. We measure the user effort in terms of how far down the user has to go in the list in order to select all important items in a weekly set.

In our evaluation study, 7 expert subjects (project members, managers) were asked to create 5-item summaries for 12 successive weeks, using a selection interface. The results with the assistance of our system show an improvement in average precision by a factor of more than 2.21 by the end of the learning period as compared to the baseline of no learning. Other evaluation metrics also show significant improvement. A low inter-rater agreement ($Kappa=0.26$) indicates that the subjects are selecting different items and the learned models are individual. Moreover, the different feature weights in the regression models for each subject identify their summarization differences. We also report our ongoing work of automatic feature extraction to make this approach domain independent.

The talk will include a short demonstration of our system showing how the learned models can be used to populate a template for a standard quarterly report.