

Generating Personal Travel Guides – and who wants them?

Cécile Paris, Stephen Wan, Ross Wilkinson, Mingfang Wu

CSIRO, Division of Mathematical and Information Sciences, Australia
{Cecile.Paris,Stephen.Wan,Ross.Wilkinson,Mingfang.Wu}@cmis.csiro.au

Abstract. In this paper we describe a system that generates synthesized web pages as a travel guide through integrating a discourse planner with a document retrieval system. We then present our investigation on whether the guide generated by such a system is actually preferred by users over a more general guide.

1 Introduction

Information retrieval is a proven technology for searching large document collections. However, the search result is usually a list of discrete units that are not easy for users to assimilate. In contrast, the discourse-oriented technologies aim at ensuring that the result is coherent and appropriate through explicit text planning. In our work, we have been designing a new approach to the tailored information delivery by combining both technologies. This approach assumes a user profile, a discourse model, and a set of web data sources, and delivers information from those sources in a coherent form to a variety of media including paper, hand-held devices, and the web. Embodying this new approach, we have developed a prototype system for generating tailored travel guides in the tourism domain. In this paper, we first describe our approach and then present one of our experiments to test the value of tailored travel information delivery.

2 Tailored Information Delivery Using the Tiddler System

The architecture of Tiddler is shown in Figure 1. The core elements include a discourse planner, a presentation planner, and a surface generator. A user dialogue system is also needed to capture the user model and to deliver the information.

The **discourse planner** takes the user model and discourse plans to create a discourse tree that represents the rough content of the guide and specifies which information is to be included in the text, and its organisation. The discourse planner, modeled on the one described by Moore and Paris [2], uses a library of discourse plans, which indicate how a discourse goal can be achieved. Mann and Thompson's Rhetorical Structure Theory (RST) [1] is used here to represent coherency.

The discourse plans were designed based on a corpus analysis and represent the prototypical structure of a travel guide. In our application, we studied and analyzed a

variety of travel guides. The resulting overall structure of the guide consists of a general introduction, and, depending on the user model, the information about accommodation, restaurants, special events when available, and activities. The user model contains the characteristics about the user that can affect the production of a personalized travel guide, namely: the type of accommodation required, the interests in terms of activities, the location and date of the trip, whether or not to show restaurant information, and the medium in which to present the guide.

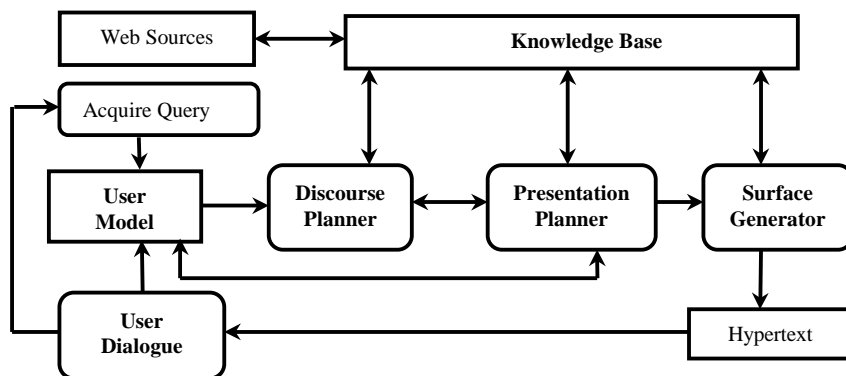


Figure 1. Tiddler System Architecture

The **presentation planner** decides how best to express the intermediate discourse tree resulting from the discourse planner, such as the amount and conciseness of the information to be included in each page and the navigation needed. Our presentation plans are based on both goal decomposition, which expands the tree, and pattern matching, which manipulates and transforms the discourse tree. The presentation planner populates the discourse tree with a series of queries against web sources. The Norfolk system [3] is used to create a set of virtual web pages by evaluating queries lying in the nodes of the discourse tree structure. Norfolk also synthesizes the results of these queries synthesized into a set of XML answer fragments. Exploiting this document and information retrieval technology allows Tiddler to query all available information (as if using information retrieval by itself), without having to first produce the knowledge sources appropriate for language generation (the main limitation of that technology), while still preserving coherence a discourse planning approach provides.

In addition to the constraints of the data itself, the presentation planner also considers the constraints of the delivery channel. For example, a hand-held device can present much less information on one screen than a standard web page. Its presentation should thus have a lot of navigation but very concise information on each particular page, whereas the Web form will use less navigation but more layouts.

Finally the **surface generator** traverses the generated discourse tree. At each node, the surface generator maps the RST relationship to its final presentation form as defined by a lexicon dependent on medium chosen by the user. Leaf nodes may also contain text retrieved by the search engine that is also appended to the document. Both the textual realization of the RST, which may be in the form of XML markup,

and the text retrieved by the search engine are appended to the final document during the traversal to produce a linearization of the discourse tree.

3 Evaluation

An experiment was conducted to examine whether a group of users preferred a tailored guide to a generic guide. We wished to explore how useful two guides were for people planning a short holiday on the coast in South-Eastern Australia. One guide was generic for the selected location(s), and the other guide was tailored to the user's particular information needs or preferences around the type of accommodation and activities. Both guides were delivered in the paper format. Our null hypothesis, therefore, was that there is no preference difference between two guides.

Of the 19 participants, 11 people preferred the tailored guide, 4 people preferred the generic guide, and the remaining 4 people had no preference. The preference difference between two guides is only significant at the level 0.059 (sign test). Although this difference is not significant enough to reject the null hypothesis, we see a promising tendency of preference to tailored delivery that could prove statistically significant with an experiment of greater power.

4 Discussion

We have described a system that delivers travel guides tailored to individual needs. By basing this system around a series of transformations of a basic discourse plan, different applications, different user models, and different delivery platforms can be supported by just changing the data and the transforms. There is no need to change the software. While engineering benefits accrue from this approach, we also believe that the key issue is the level of benefit of a discourse approach to tailored information delivery, which ensures a coherent presentation no matter what the circumstances are. Our key observation from an initial evaluation showed that users have a strong tendency to prefer tailored documents as a whole.

Reference

- [1] Mann, W.C., & Thompson, S.A. (1988). Rhetorical structure theory: Towards a functional theory of text organization. In *TEXT*, 8(3), pp 243- 281.
- [2] Moore, J.D. & Paris, C.L. (1993). Planning Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information. In *Computational Linguistics*, Cambridge, MA. Vol 19(4), pp.651- 694.
- [3] Paradis, F., Vercoustre, A.M., and Hills, B. (1998). A Virtual Document Interpreter for Reuse of Information, in *Lecture Notes in Computer Science 1375, Proc of Electronic Publishing*, Saint-Malo, France (pp. 487-498).