Chapter 5

Visualization Tools for Designing Spoken Dialogs

Deborah A. Dahl

Conversational Technologies

Key words: speech application development tools, spoken dialog design, visualization

Abstract: Designing sophisticated spoken human-computer dialogs is difficult both because dialogs can become extremely complex and also because spoken dialogs by definition are auditory, not visual. Therefore there is no obvious mapping of a spoken dialog design to a visual representation that can be used in a design tool. Nevertheless, tools for dialog design are very important, and there have been many attempts to develop useful tools to aid developers in designing spoken dialogs. This paper discusses a number of approaches to dialog visualization and points out considerations that need to be taken into account when selecting a tool for a particular project.

1. PROBLEM

Spoken human-computer dialogs are difficult to design for many reasons. For example, there are many gaps in our understanding of the human factors of how people interact with speech interfaces. This problem has been the focus of considerable research. In addition, there is a great deal of practical knowledge in this area gained from the analysis of deployed applications (see, for example, the papers in [1]). This paper focuses on a second problem that has received much less attention. That is, how can a developer who is focused on the details of scripting a complex dialog, develop an understanding of the entire dialog, making sure that all paths through the dialog are reasonable and consistent, and that the dialog as a whole doesn’t contain dead ends or loops? Mixed initiative and user initiative dialogs in particular have many complexities introduced by the freedom of the user to
change the direction of the dialog at any time. When all of these situations have to be individually anticipated by the designer, the complexities can become overwhelming. It’s easy to assume that graphical dialog design tools will automatically provide a solution to this problem, but tools that aren’t a good fit to the design task can do more harm than good.

There are many design tools available in the speech industry now. These tools adopt different approaches to assisting designers in their tasks. How can we know what’s the best dialog design tool for a specific project? This paper looks at a number of approaches to dialog design tools, points out advantages and disadvantages of each approach and provides some dimensions for classifying tools.

Although there are many types of spoken human-computer dialogs, the focus in this paper is on designs for the form-filling tasks and navigational dialogs that are by far the most commonly encountered types of dialogs in commercial systems. By form-filling dialogs, we mean dialogs where the task centers on obtaining several pieces of information from a user and then acting on that information. Travel reservations, finding a certain type of restaurant, and placing orders are good examples of form-filling dialogs. An example of a navigational dialog is the kind of dialog that might be encountered in a voice portal where the user is asked questions that allow the system to navigate to a particular application. Other types of dialogs potentially will require quite different dialog design tools. In the future these other types of dialogs may be more frequently deployed, and may require different thinking about tools for dialog design.

Another type of dialog which this paper does not address is multimodal dialogs, which contain visual components as well as spoken ones. Due to the presence of the graphical interface, graphical tools for designing multimodal dialogs will be very different from tools for designing voice-only dialogs.

2. VISUALIZING DIALOGS

One important component of most design is visualization, or how a non-visual concept or data can be displayed visually in a meaningful and conceptually transparent way. The main point of this paper is to compare different approaches to visualization in designing dialogs, but it will also discuss other general considerations in tool selection.

The appropriate type of visualization for a particular task depends on how the underlying task is conceptually structured. There is considerable research literature on visualization for various kinds of abstract data, for example, the papers in [2]. However, very little work has directly focused on effective visualizations for spoken human-computer dialogs. Visualization is