SUEDE: Iterative, Informal Prototyping for Speech Interfaces

Anoop K. Sinha, Scott R. Klemmer, Jack Chen, James A. Landay, Cindy Chen
Group for User Interface Research
CS Division, EECS Department
University of California at Berkeley
Berkeley, CA 94720-1776 USA
+1 510 643 3043
{aks, srk}@cs.berkeley.edu • http://guir.berkeley.edu

ABSTRACT

SUEDE is a speech interface prototyping tool that enables rapid, iterative creation of prompt-response speech interfaces. It explicitly supports iterative design, allowing a designer to quickly create an interface prototype, conduct user studies, and analyze the test data in a single tool. SUEDE offers an electronically supported Wizard of Oz (WOz) technique that captures user test data. Before building the tool that we present here, we built a paper and an interactive prototype and tested the prototypes with a total of fifteen users. We also interviewed six professional speech UI designers at their workplace about their current design practice. These designers encouraged us to build a tool that lets them start with examples when designing and simulate recognition error during testing. Our current tool is a culmination of this iterative design.

Keywords

Wizard of Oz, speech user interfaces, prototyping, design, low-fidelity, informal user interfaces, design tools

INTRODUCTION

Speech-based user interfaces are more appropriate than graphical user interfaces in many settings and will likely become a more common user interface paradigm in the near future. Speech UIs are hard to prototype due to the substantial time and technology expertise traditionally required to build them. We have developed an interactive prototyping tool called SUEDE to address this difficulty.

Most successful design and usability specialists iteratively create interfaces. SUEDE is a tool that explicitly supports each stage of this process. In addition to an interface for designing, our tool provides an interface for testing and collecting test data. We provide a visualization of the test data directly in the design as a mechanism for interface analysis. We believe that this explicit support for iterative Design, Test, and Analysis holds value as a general principle for user interface design tools.

An informal interface approach, using unrecognized, natural input (e.g., sketching or speech) can successfully support designers in the early stages of design [1]. With informal tools, designers are encouraged to explore the design space rather than just detail one single design idea. SUEDE embodies this approach.

HOW SUEDE WORKS

SUEDE's direct manipulation interface offers a mode for each of the three main phases of the iterative design cycle: Design, Test, and Analysis.

Design Mode

In the *Design* phase, a speech designer often begins by creating dialog examples. These examples evolve into, and provide the foundation for, the actual interface design. In SUEDE, they are created horizontally in the top area, called the *script area*, of Design mode (Figure 1).

After constructing several example scripts, the designer starts to construct a *design graph* representing a more general design solution. Designs are created by dragging cards from the script onto the canvas, creating new prompts on the canvas, and linking prompts with user responses. The designer records audio for each element in the graph.

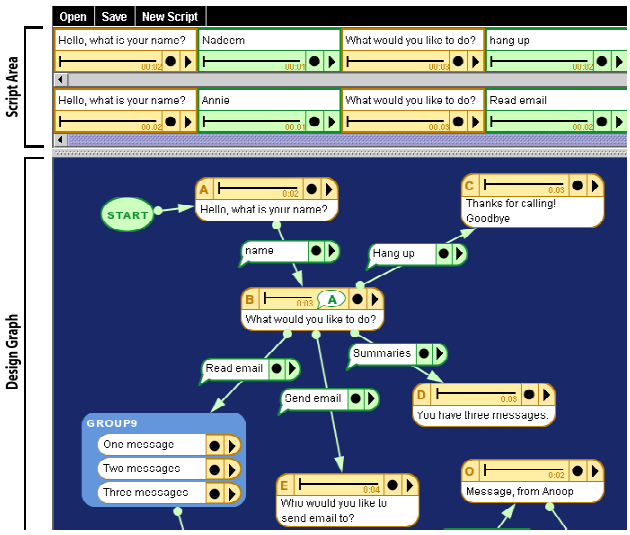

Figure 1. SUEDE's Design mode allows the easy creation of example scripts *(top)* and speech UI designs *(bottom)*.

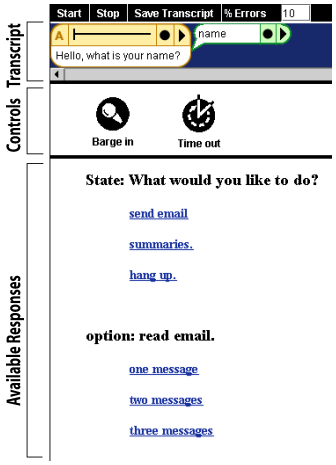

Figure 2. Test mode allows the wizard to focus on the available responses.

Test Mode

In the *Test* phase, a designer tries out a design with target users, following in the tradition of Wizard of Oz methodologies in speech systems [2]. Because the wizard "recognizes" participant responses, no speech recognition or speech synthesis is necessary to test SUEDE prototypes.

When the Wizard starts a test session, SUEDE automatically plays the pre-recorded audio from the current prompt and captures the response audio. The Wizard waits for the test participant to respond, and then based on the response, clicks on the appropriate hyperlink in the web-based test interface (Figure 2). Test mode displays and saves the actual transcript of the test session as it progresses. These user-generated transcripts containing system audio output and participant spoken audio input are the test-time analog of designer generated scripts.

SUEDE incorporates functionality to automatically insert simulated speech recognition errors. At the specified error rate, SUEDE overrides the wizard's choice, informs him of this fact, and randomly chooses one of the other links.

Analysis Mode

During *Analysis*, designers examine collected test data, deciding how it should influence the next design iteration. The Analysis interface is similar to the design interface, except the top of the screen contains user *transcripts* from the last test session and the design graph is annotated with test data (Figure 3). The annotated design includes information on the number of participants who took a particular path, the actual audio responses they made, and how long they took to respond to each prompt. The designer can review all test session responses to specific prompts directly in the Analysis interface, a substantial improvement over current practice, which often involves reviewing audio tapes.

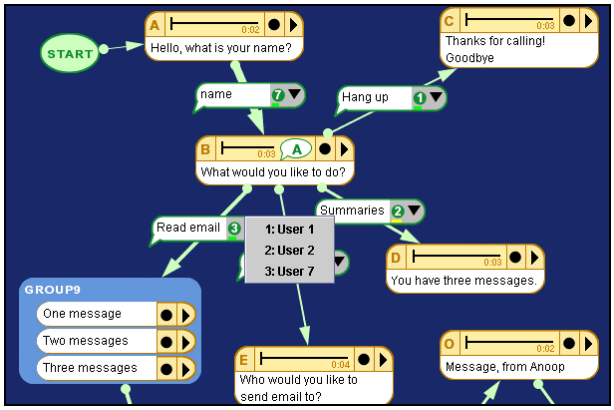

Figure 3. Analysis mode displays an annotated version of the design that summarizes the aggregate test results and provides the ability to hear the set of audio responses for a particular link.

RELATED WORK

SUEDE has been inspired by previous work in informal design tools [1] and Wizard of Oz studies [2]. We also took inspiration from two existing Speech UI construction tools which are geared towards rapid prototyping rather than informal prototyping: the CSLU Rapid Application Developer (RAD) [3] and the Unisys Natural Language Speech Assistant (NLSA) [4].

FUTURE WORK

We have released SUEDE to the speech design community to further evaluate its use (http://guir.berkeley.edu/suede/). Some speech UI designers who have seen SUEDE have asked for a way to bridge to their particular development environments. As a first step, we have created an add-on to SUEDE that creates a preliminary grammar for a standard speech recognition system.

CONCLUSIONS

The speech interface design problem is complicated, and SUEDE makes significant progress in supporting the early stages of this process. Based on our interviews, SUEDE's Design, Test, and Analysis paradigm maps quite well onto the speech designer's mental process.

The designers we interviewed told us that problems with speech interfaces often arise when users do not know what to say, when they say invalid words, or when the system makes recognition errors. In SUEDE, playback of the transcript allows the designer to hear what participants actually said, including places where the participant paused or had trouble. Error simulation allows a designer to see how participants cope with recognition error. Using SUEDE's Analysis mode, a designer can make decisions about appropriately redesigning the flow of the interface.

REFERENCES

1. Hearst, M.A., M.D. Gross, J.A. Landay, and T.E. Stahovich, Sketching Intelligent Systems. *IEEE Intelligent Systems*, 1998. 13(3): p. 10-19.
2. Kelley, J.F., An Iterative Design Methodology for User-Friendly Natural Language Office Information Applications. *ACM Transactions on Office Information Systems*, 1984. 2(1): p. 26-41.
3. Sutton, S. and R. Cole, The CSLU Toolkit: rapid prototyping of spoken language systems, in *Proceedings of UIST '97: the ACM Symposium on User Interface Software and Technology*. p. 85-6, 1997.
4. Unisys, *Natural Language Speech Assistant*, 1999. Unisys Corp. http://www.unisys.com/marketplace/nlu/nlaproductinfo.html