Let's Go Lab: a platform for evaluation of spoken dialog systems with real world users

Maxine Eskenazi, Alan W Black, Antoine Raux, and Brian Langner

Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA. USA

`{max,awb,antoine,blangner}@cs.cmu.edu`

Abstract

This short paper is intended to advertise Let's Go Lab, a platform for the evaluation of spoken dialog research. Unlike other dialog platforms, in addition to example dialog data and a portable software system, Let's Go Lab affords evaluation with real users. Let's Go has served the Pittsburgh public with bus schedule information since 2005, answering more than 52,000 calls to date.

Index Terms: spoken dialog systems, evaluation

1. Overview

Let's Go Lab is a research lab at the service of the spoken dialog community that provides:

1. Collected data for offline study (52,000 dialogs).
2. The dialog system software itself.
3. The platform, which allows online dialog studies on a large volume of calls from real users.

The Let's Go Lab dialog system has been answering the phone for the Port Authority of Allegheny County nightly since early March 2005 and has collected over 52,000 calls. The callers are Pittsburghers from every walk of life who need bus scheduling information. The project website can be accessed at http://www.cmuletsgo.org/.

2. The Let's Go System

Let's Go Lab is built on top of Let's Go[1], a spoken dialog system for bus information. This in turn is based on Olympus, a dialog system architecture born out of Carnegie Mellon's existing expertise in speech recognition, speech synthesis and spoken dialog systems[2]. With around 60 calls a day LGL offers a unique experimental environment for evaluation of advances in spoken dialog systems.

Initially, Let's Go was a typical research mixed-initiative system used for in-lab experiments[1]. However, in order to serve the diverse user population of the Port Authority, the interaction was constrained to make it easier for novice users. Much effort was spent on the efficiency and robustness that is required for a continuous public service. Let's Go Public went "live" in March 2005[3], providing bus schedule information daily from early evening through early morning, when the Port Authority offices are closed. It is directly accessed through the Port Authority number (+1-412-442-2000). Although the system has not been heavily advertised, it receives around 60 calls per night on average.

3. A Platform for User Studies

In our experience in building and deploying spoken dialog systems, one of the hardest issues in dialog research is getting real users. DARPA Communicator evaluations consisted of paying callers to dial each system to plan a scripted flight schedule. At startup, Let's Go also had volunteers call the system with pre-specified tasks. Although in some cases this is the only way to carry out an evaluation, we found that dialogs with users who have no real need for the information differs dialogs with users who need the information in various ways[4]. While commercial systems do get calls from a lot of real users, their data is proprietary, and thus unavailable to the research community. Let's Go Lab offers a unique opportunity in spoken dialog system research in that its user base is large and consists of people who call to fulfill a genuine information need rather than to complete a predefined scenario.

The Lab enables grounded research in dialog systems at a time when the volume of research in the domain is at a high, productive level. It is a stable, deployed system with real users that researchers from the spoken dialogue community can use to experiment with new ideas, and to properly evaluate how they are advancing the state of the art. This type of community service for active experimentation is analogous to the Hubble telescope and the LearnLabs in the NSF Pittsburgh Science of Learning Center, where researchers apply for "live" time to run their own experiments.

The Let's Go Lab platform is unique. While other system may only offer software or recorded dialogs, Let's Go Lab has an existing large user base, thus allowing for more reliable studies than may be possible in a lab setting.

We see many projects that can be run with the Lab, beyond dialog research. For example:

- **TTS**: prompt variations varying level of information presentation
- **ASR**: investigation of frustrated speech
- **Dialog**: error recovery, implicit confirmation, etc

Let's Go Lab is open to dialog researchers throughout the world as a resource for their research. All researchers are encouraged to contact the authors to be a part of this exciting new approach to dialog research.

4. Acknowledgements

5. References

[1] A. Raux, B. Langner, A. Black and M. Eskenazi, "LET'S GO: Improving Spoken Dialog Systems for the Elderly and Non-natives", Eurospeech 2003, Geneva, Switzerland.

[2] A. Rudnicky, C. Bennett, A. Black, A. Chorimongkol, K. Lenzo, A. Oh and R. Singh, "Task and domain specific modeling in the Carnegie Mellon Communicator", ICSLP2000, Beijing, China, vil II, pp 130-133..

[3] A. Raux, D. Bohus, B. Langner, A. Black, and M. Eskenazi "Doing Research on a Deployed Spoken Dialogue System: One Year of Let's Go! Experience" Interspeech 2006, Pittsburgh PA.

[4] H. Ai, A. Raux, D. Bohus, M. Eskenazi and D. Litman, "Comparing Spoken Dialog Corpora Collected with Recruited Subjects versus Real Users", SIGDial, Antwerp, Belgium 2008

September 22 – 26, Brisbane Australia