From Graphical to Voice User Interface: The Next Revolution

Chin-Hui LEE
Dialogue Systems Research Department, Bell Labs, Lucent Technologies
Murray Hill, New Jersey, USA
chl@research.bell-labs.com

The end of the 20th century witnesses the explosive growth of internet usage. Human beings are no longer satisfied with simple connectivity among people which is the main thrust for technology advancement in communications in a large part of the last one hundred years after the invention of the telephone. In addition to talking among people, they now have an increasing appetite for more information and content. The most common vehicle for accessing information residing on many websites across the globe is still the dominating interface of point and click with a mouse using the graphical user interface (GUI). However due to the emerging need for wireless and mobile internet appliances, such as cellphones and personal data assistants, the traditional GUI is having difficulties in delivering content effectively because of the miniaturization of keyboards and displays. On the other hand, voice input and output devices are usually built-in and heavily used in such mobile communicators. Since speech is the most natural means of interface and communication especially in eyes-busy and hands-busy situations, voice will be a dominating mode in newly designed multi-modal user interfaces for future devices. This calls for a revolutionary design of a voice user interface (VUI) to supplement the conventional GUIs.

In this talk we review the core voice technology components, including automatic speech recognition (ASR), text-to-speech (TTS), speaker verification (SV) and utterance verification (UV). We discuss technology capabilities and limitations of each technology. In order to build effective VUIs, dialogue management and natural language understanding are two critical components that are not yet mature and many research issues are still open and need to be addressed. Even with all the constraints, a number of large scale applications and services have been implemented and deployed in recent years. Voice portal companies have also been established to take advantage of the increasing application needs. We will present a few application examples and discuss why they are successful.

Finally recent approaches to web and phone service integration and personalized voice portals are illustrated. They point to new technology challenges in order to realize VUIs. They also offer tremendous business opportunities and benefits to the society in the global village of the coming information age in the new millennium.