Chapter 8:

SUMMARY AND FUTURE RESEARCH

Introduction

Sounds have been and will continue to be a part of our computing world just as they are a part of our broader everyday world. From the reassuring click of the keys we hit on keyboards to the steady hum of a healthy processor, sounds provide a multitude of clues about our environment. The topics and work presented here represent a start at designing the use of non-speech sounds in the computer-human interface. In this way, sounds can play an expanded role in presenting information and we designers, developers, and users of interfaces can take a more active role in accessing and using sounds.

The work presented in this tutorial has shown that sounds already contribute to our ability to gather information about the world around us. Sounds can be used effectively to present information in computing environments and applications. Sounds have potential for being structured to represent complex information. Sounds and visuals may play complementary roles in providing information. Technologies for producing sounds are both capable and available, and much is already known about the nature and perception of sounds. Our challenge is to integrate that knowledge and technology into useable and useful computer-human interfaces.

Consider six areas of future research: the use of non-speech sounds, the mapping of information to sounds, sounds in relationship to visuals, user manipulation of sounds, the structure of sounds, and system support for sounds.

The first area of research that we suggest for the future is to develop a better understanding of the *use* of non-speech audio. Although Gaver (1989, also section 3 of these notes) considers how sounds are used in our everyday lives, much more of this sort of work needs to be done. We need studies of everyday listening to provide an understanding of the way in which sounds are currently used and to inform decisions about the ways in which sounds might be used. We have tried, in section 2, to provide a brief summary of issues in acoustics and psychoacoustics that are likely to be relevant to auditory interfaces. But we also need to explore the perception of higher-level musical structures in order to assess their potential to encode information for users. For instance, Blattner et al. (1989) suggest that variations on motives can be used to signal complex information. What sorts of variation will prove most useful, easiest to relate to the original, best associated with a particular meaning? Patterson (1982) has studied the perceived urgency of sounds, but what about their annoyance? Understanding how people listen to and react to music seems likely to aid us in designing effective and aesthetic auditory cues. Finally, as prototypes of various applications of sounds appear, we need to observe the resulting work practice and the use of sounds in that practice.

A second area of research is to explore more completely the mapping of information to sounds. Each of the examples in section 5 takes a different means of translating the events or data into sounds. In order to present information effectively in sounds, we need a more general understanding of the relationship between kinds of sounds and the kinds of information presented. Gaver suggests that mapping events in the computer to sound-producing events is intuitive and natural. But there are events in the computer that will not map readily to everyday events. When might it be useful to hear the sound itself? Clearly some musical properties map naturally to properties in the world -- for instance, a high pitch means up, a low pitch means down. Mezrich et al. (1984) implicitly acknowledge this principle in translating graphic representations of data to sound. Does a high pitch also stand for a large data value? What are other examples of such mappings that people spontaneously generate and understand? How difficult is it for people to learn and use wholly new mappings?

The third area of future research we offer is more study of the relationship between sound and graphics in interfaces. More generally, we might ask how sounds work in relationship with any other feedback in computer-human interfaces. Sounds may complement or work with other interface information; they may replace other interface feedback; they may exist simultaneously but independently of other interface outputs. Bly (1982) and Mezrich et al. (1984) have noted that sounds may be particularly well suited to displaying multidimensional data, and may make some aspects of data more apparent than graphic displays do. Mansur (in Bly, 1985) and Gaver (1989) noted that sounds are useful in that they do not rely on the directed attention of the user; this is of course why so many alerts are auditory (Patterson, 1982). Gaver further points out that sounds are useful for displaying information about changing events, and about occluded events or internal mechanisms of objects. Future work will address the question of whether auditory and graphic components can be designed to create one coherent informational system.

Certainly auditory interfaces offer potential aid to visually-disabled users in gaining access to computers. Morrison & Lunney (in Bly, 1985) have shown that visually-disabled students are able to learn to use auditory representations of infrared spectrograms, and Edwards (1989) has created an auditory wordprocessor that visually-disabled people could use. Will we be able to take this research further and convey spatial information using sounds -- can we create add-on auditory interfaces that allow the visually-disabled to use existing graphic interfaces, as Edwards (1989) suggests?

As a fourth area of future reseach, consider what principles might govern the user manipulation of sounds in interfaces. Surely users must be able to control at least the volume of auditory interfaces -- even emergency alerts in aircraft, nuclear power plants and hospitals have on/off switches. How much and what other kinds of overall control will be needed? What interactive techniques give users access to sounds and control over sounds? How do users refer to sounds, store sounds, retrieve sounds, designate sounds attached to information? Given that both musical and everyday sounds are difficult to design, how will new "sound palettes" be created and distributed? The answers to many of these questions will rely on the user model of the sounds which in turn depends on many of the other research areas suggesteed here: the kinds of sounds (those that are manipulated and perceived in terms of their acoustic properties or their sources), the applications of sounds, and the relationship between sounds and graphics.

A fifth area of future research should consider the structure of sounds. Can useful sounds be built out of other sounds? In Section 3, we suggest that many complex auditory events, such as slamming doors or footstep sounds, are the result of a constrained series of more primitive sound-producing events. How can these complex events be understood, and can they be mapped to information such that both the simple and the complex events are meaningful? How is complex structure to be mapped to sound? Blattner et al. (1989) suggest that variations on motives can be used to create structured audio messages, while Gaver (1989) claims that everyday sounds are structured because sound-producing events are themselves structured, and that the structure inherent in everyday sounds can be mapped readily to complex information. Will both (or either) of these claims prove true, and how easy will it be for listeners to perceive and learn the structures to be conveyed?

Finally, our sixth area of research is concerned with the needed computing systems to support sounds. What are the appropriate software architectures for user interfaces including sounds? What does a tool kit for sounds include? What capabilities of sounds are needed? Will users of audio interfaces need to purchase MIDI controllers, synthesizers, and other peripheral hardware, or will the increased interest in providing good audio from personal workstations allow suitably complex manipulations of sound in the near future? The more we understand about the system needs for non-speech sounds, the better positioned we will be to influence the integration of sounds with future hardware and software.

We can imagine the sounds of tomorrow: a multi-processing environment in which background tasks indicate their status by audio. Or a sound cue that signals a remote colleague is accessing the same text file you're editing. Consider an expansion of scientific visualization using graphics to include sounds, so that mapping data dimensions to sounds will become a standard option in many graphing programs. We hope that, after this tutorial, such applications will not only seem sensible, but obvious.