Breakthroughs - The DSP
Making computers listen, speak, and sing

Imagine being able to talk to your computerÐand have it talk back. You could say commands, rather than typing them or selecting them from menus. Verbally dictate correspondence, database entries, and other information for the computer to type and file. Ask the computer to look up information and read it to you. And do any of these things remotely, via the telephone.

This kind of ªnatural interactionº between humans and computers isn't just the stuff of science fiction. The technologies that will make it a reality are being developed today. Exciting work is already being done in areas such as speech recognition and music synthesisÐthe beginnings of natural interactions with computers.

Much of this work is being done with NeXT Computers, thanks to the Motorola DSP56001 Digital Signal Processor (DSP) that's included in every one of the computers. The DSP chip, which is installed alongside the central processing unit and floating-point processor, offers extremely fast processing of digital signalsÐincluding sensory data such as sound, music, speech, and images. The DSP also handles data compression and expansionÐan especially valuable feature when dealing with the very large amounts of data typical of sound files. (For technical details on the DSP, see [sidebar].)

Because the DSP is included in every NeXT Computer, developers and researchers have a standard platform for exchanging music applications, sound files, and other softwareÐand the opportunity to routinely include advanced ªsensoryº interactions in their applications. To meet the requirements of sophisticated sensory interactions such as music and speech synthesis, the DSP is supported by hardware and software that enables it to generate sound of the same quality offered by compact disc players.

Some of the most promising work with DSP technology and the NeXT Computer is being done at universities. The following is a sampling of these academic projects.

The Sphinx Speech Recognition system developed at Carnegie Mellon University is widely regarded as one of the best of its kind. Utilizing the DSP chip, Sphinx runs in real time for 40-50 word vocabularies. It currently has a recognition rate of better than 96 percent for a 1,000-word task with moderately difficult grammar. Raj Reddy is the project leader, Eric Thayer is the DSP programmer, and Fil Alleva did the search programming. The Sphinx system running on the NeXT Computer is based on recent thesis research by Kai-Fu Lee at Carnegie Mellon. For more information, contact Alex Rudnicky at air+@cs.cmu.edu.

At the University of Michigan, Gregory Wakefield, and John Feng are developing NeXT courseware for a class in signal processing. The class will cover spectrum analysis and digital filter design. Wakefield will be offering a 2-day workshop in late July on the package in Ann Arbor, Michigan. For more informaton, contact Wakefield at ghw@caen.engin.umich.edu.

The Computer Audio Research Laboratory (CARL) at the University of California, San Diego's Center for Music Experiment has ported its cmusic program to the NeXT Computer. cmusic is a general-purpose acoustic program which has been modified to produce and process NeXT-style soundfiles. In addition, a new phase vocoder program (a tracking spectrum analysis system) has been implemented on the NeXT, both in software (pv) and DSP-supported (pvdsp) forms. The DSP-supported version can handle FFT sizes up to 1024 real points and speeds along about 3 to 5 times faster than its software counterpart. To obtain the CARL software package, which consists of about 115 programs for computer music, including complete sources and documentation, send e-mail to F. Richard Moore at frm%plexus@ucsd.edu or write to him at CARL, Center for Music Experiment, Q-037, UCSD, La Jolla, CA 92093.

At Stanford University's Center for Computer Research in Music and Acoustics (CCRMA), NeXT is the development environment for a variety of research projects:

- Perry Cook has developed an interactive vocal synthesis program, called SPASM, that allows the user to directly manipulate a physical model of the human vocal tract and immediately hear changes in the synthesis. (See [Stanford article].)
- Glen Diener has developed an easy-to-use and powerful music notation environment for the NeXT, which also makes use of the Music Kit's DSP sound synthesis.
- Mike Malcomb helped develop an application for measuring the perception of group-delay distortion

in digital filtersÐa good example of the value of CD-quality sound support in psychoacoustics experiments.

· Bill Schottstaedt is developing a high-level music compiler, written in Lisp, that drives the DSP directly in array-processing mode. The main purpose of this is to give music composers a workstation environment that creates music directly from high-level software.

As these examples show, the DSP and the NeXT Computer are already changing the way we interact with computers. Before long, the work being done with these technologies will lead to a world in which talking with and listening to computers is as routine as using a keyboard is today.

The DSP, technically speaking

To understand what the Digital Signal Processor (DSP) is, you first need to know what ªdigital signal processingº means. A ªsignalº is some physical variable used to convey information, measured over time. Often, it is necessary to ªclean upº a signal by smoothing it, removing noise, separating it from unwanted signals, or otherwise transforming it into a more useful form. The collection of mathematical techniques for processing signals is called, naturally enough, ªsignal processing.º For analog signals to be stored in a computer, they must be ªdigitized,º or converted into a set of numerical values. Processing signals in digital form is thus called ªdigital signal processing.º

The DSP56001, built into every NeXT Computer, is designed specifically for digital signal processing. It is a fixed-point chip, so it is not for floating-point number-crunching. (A floating point coprocessor, the Motorola 68882, is also built into every NeXT Computer.) What it excels at is processing music, speech, and other signals. Its combination of an elegant architecture and 24-bit data makes it the leading chip for digital audio processing.

For developers who will be writing DSP assembly code, a variety of development tools are provided with every machine. However, it is not necessary to know 56000 assembler to use the DSP in your programs. Each NeXT Computer also has intuitive tools including the Music Kit, Sound Kit, and an array processing library.