the seed, or the atmosphere in the growth chamber—the group was able to control the composition of the resulting fiber. One of the impressive things about the process, Feigelson said, is that the fiber is superconducting as it crystallizes, and needs no further processing. Other techniques for making superconducting wires require an extra step to change the material into a superconducting form.

The Stanford tea[illegible] laser-heated pedestal growth process on more than 50 different types of materials, Feigelson said, and this process produces closer-to-perfect crystals than any other. He said he expects therefore that the superconducting bismuth fibers will prove to be very high quality crystals, which will make them quite valuable to researchers investigating the properties of the bismuth superconductors. So far these superconductors have proved quite difficult to grow in pure form.

The Stanford group has also grown fibers of the 1-2-3 superconductors, Feigelson said, but they required an extra annealing step (heating in the presence of oxygen, then cooling) after the wires were formed. Still to come are attempts to grow useful fibers from the 2223 phase of the bismuth superconductor as well as from the thallium compounds.

Whether the growth process will ever be commercially feasible for making superconducting wires depends partly on how much it can be speeded up, since it is very slow now. "In principle, there is no limit to the length of the wire," Feigelson said, but "right now it's more of an experimental tool for studying high-quality fibers."

■ ROBERT POOL

A Landmark in Speech Recognition

A computer science graduate student at Carnegie Mellon University, Kai-Fu Lee, appears to have toppled some of the largest barriers to building a genuinely useful computer system for speech recognition—a goal that has tantalized and frustrated the likes of IBM and the Defense Department for decades.

SPHINX, as Lee's experimental program is known, has demonstrated up to 96% accuracy in identifying spoken words, while simultaneously meeting some of the field's most stringent performance criteria: continuous speech (no pauses between words); speaker independence (no restrictions on who is talking, and no requirement that each user spend time training the system); a natural task (questions about naval ship movements); and a large vocabulary (997 words.)

While running on special-purpose hardware, moreover, SPHINX can produce its output at roughly one-third the speaker's rate of input, with every indication that further improvements will enable the system to keep up with the speaker word for word.

Individually, any one of these achievements would put SPHINX close to the state of the art in speech recognition; taken together they are unprecedented. "I think we're entering a phase when [speech recognition research] will show some real payoffs," says Lee's adviser Raj Reddy, a former president of the American Association for Artificial Intelligence and an active contributor to this field for more than 15 years. SPHINX is hardly perfect, he says—"The Holy Grail of totally spontaneous, unrestricted speech is still 10, 20, 30 years away." Yet it does point the way toward some genuinely useful applications in the near term, with the possibilities ranging from office dictation systems to the voice command of robots on the factory floor. In fact, says Reddy, "It begins to be possible to imagine speech hardware and software as standard on every personal computer."

From a theoretical point of view, moreover, it is significant that SPHINX owes its power not to any one conceptual breakthrough, but to a series of incremental improvements within a thoughtfully planned framework. "Kai-Fu's thesis shows that there is no substitute for having a careful understanding of the phenomenon—speech—and then representing it in the very foundations of your program," says Reddy.

As Lee himself explains it, those foundations consist of four principles:

■ **A good model of speech.** One of the many things that makes speech recognition so hard is that the same word can vary wildly in sound depending on where it is in a sentence, what words come immediately before and after, how much stress it receives, how rapidly the speaker is talking, *who* is talking, and a host of other variables. At least 50 distinct pronunciations have been recorded for the word "the"—when it is not smothered or swallowed entirely—and at least half a dozen are typically found even for such formal words as "Atlantic."

Thus, says Lee, the first priority in a computer system is to encode all these possible variations in a form that is simple enough and compact enough for the computer to handle. In SPHINX this is done by a technique known as Hidden Markov modeling, which has become very popular in the field since it was introduced about 15 years ago. In effect, this method treats every distinct sound in English as a little computer that can generate a whole array of pronunciations; thus it can be viewed as a speech generator. Given a vocal input, conversely, the Hidden Markov method picks out the sound most likely to have produced it; thus it can function as a speech recognizer.

By adding in some restrictions derived from the "grammar" of the 997-word vocabulary—that is, a statistical measure of which sounds are likely to follow after one another—Lee was able to use this method to achieve accuracies of up to 75%. Adding more words to the vocabulary does tend to slow the system down. However, one of his top priorities for future research is to improve the modeling algorithms so that they can handle many thousands of words.

■ **Use of human knowledge about speech.** When we humans listen to an utterance we decipher it not just by sound alone, as the standard Hidden Markov method does, but by factors such as stress, duration, and the rise and fall of pitch, all of which have to do with the way the sounds are changing in a given context. Lee accordingly integrated such factors into SPHINX, along with some simple phonological rules, and found that its accuracy improved to some 90%. Another item on the agenda for future research is to add higher level knowledge about grammar and the meaning of words, using a variety of artificial intelligence techniques.

■ **Choosing a good unit of speech.** The way we pronounce a given sound is strongly influenced by what comes before and after it: listen carefully to the initial sounds in *tea* and *tree*. Lee therefore used his Hidden Markov models to encode not single sounds, but sequences of three sounds—"triphones." He also used separate models for function words such as *the* and *by*. This boosted the accuracy of SPHINX to almost 96%.

■ **Adaptation to individual speakers.** In much the same way that a human listener can adapt his or her ear to a strange accent, a speech recognition system ought to be able to improve its performance by adapting itself to the vagaries of the individual speaker—although it has to do so automatically and unobtrusively, says Lee, lest it degenerate into a speaker-dependent system. For SPHINX he investigated several algorithms for doing this, and found that they made only a marginal difference. However, he believes there is still a good deal of progress to be made in this area.

■ M. MITCHELL WALDROP