Sebeok, T., ed. 1968. *Animal communication: Techniques of study and results of research*. Bloomington: Indiana University Press.
———. 1977. *How animals communicate*. Bloomington: Indiana University Press.
———, and R. Rosenthal, eds. 1981. *The clever Hans phenomenon: Communication with horses, whales, apes, and people*. New York: New York Academy of Sciences.
Seyfarth, R. M., D.L Cheney, and P. Marler. 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication, *Science* 210 (4471): 801–803.
Snowdon, C., and M. Hausberger, eds. 1997. *Social influences on vocal development*. New York: Cambridge University Press.

Keywords

animal communication
animals and language

7 Can computers learn language?

As I am revising this chapter in 2008, humans can interact verbally with computers in multiple ways. To take a simple example, when I dial the operator at my college and request the campus directory, a recorded voice asks me for the name of the person I'd like to speak to. When I say the name, the recording responds, "Did you say [the name]?" If I respond positively, the phone number of that person is automatically dialed for me. If I respond negatively, the recording apologizes, and a human operator comes on the line. Here a voice recognition program is at play. Such programs have been around for quite some time, and they are used in a variety of situations. One of my friends, who has cerebral palsy and is a quadriplegic, has a computer that has been programmed to recognize her voice and make written files of whatever she says. She's used this program for years to "type" with her voice. There's no doubt, then, that voice recognition by computers is possible.

Therefore, when we hear the common claim that computers can learn language, we might not be skeptical, although we should be. Learning a language is a complex process that involves more than voice recognition. When we ask whether or not computers can learn language, we're asking whether we will be able to have conversations with a computer that are indistinguishable from those with another person.

People have been asking that question and trying to devise computer programs that will allow a positive answer since as early as 1950, based on the work of Alan Turing, a British mathematician. Turing developed a kind of competition in which an interrogator (by keyboard) tried to figure out which of two respondents was a human and which was a computer. He predicted that a computer program would be judged "human" in such a competition within fifty years.

In the 1960s the program ELIZA was developed, which did little more than convert an input statement into a question. The result was that conversations with ELIZA were jokingly claimed to resemble conversations with a psychotherapist. Between then and now, other programs have been developed. Every time you use a search engine, you engage in a rudimentary kind of conversation with a computer.

If you'd like to try out some current attempts at computer-human conversations, there are at least two that might interest you. One is a.l.i.c.e. (http://www.alicebot.org). To any statement or question you input, a.l.i.c.e. responds with a preprogrammed statement or question from its database—and that database is impressively large. Nevertheless, one of my students in the spring of 2001 managed to stymie a.l.i.c.e. with the question "How much wood could a woodchuck chuck if a woodchuck could chuck wood?" Another possibility is Daisy (http://www.leedberg.com/glsoft), which is not preprogrammed at all. Rather, Daisy stores your input and manipulates it. So when you first talk with her, she seems to have no intelligence whatsoever (all she can do is repeat what you type in), but if you spend absurdly large amounts of time with her, she gradually comes to form relatively coherent responses.

Despite what seemed to be a reasonable prediction at the time, now more than fifty years after Turing made his claim we still have not come up with a program that allows a computer to participate like a human in a conversation. Will we ever be able to? At this point, the answer to that question is, of course, speculation. However, some guesses are more informed than others. To help you make an informed guess, let's analyze some sample conversations and ask whether computers could have produced them. Consider conversation 1:

A. Where are you going?
B. I am going to school.

In thinking about what went into this conversation, we can begin with the most obvious facts. The first is that someone produced utterance A. We know computers can produce language—whether preprogrammed (like a.l.i.c.e.) or not (like Daisy)—so utterance A could have been said by either a human or a computer.

Response B is based on interpreting utterance A. Again, we know computers can do this to a certain extent. That is, they can analyze sentences to some degree, recognizing verbs, *you* as the subject, and *where* as a location question word. They can then match that sentence with the same verb(s), an appropriate subject, and possible location responses—often those that begin with *to* (such as *to work, to the grocery*, and in this conversation *to school*). However, if you say the sentences aloud, you will find them stilted. Instead of response B, what would sound more natural? Probably this:

C. School.

That is, in casual conversation, we typically answer in fragments rather than whole sentences. And even if we were to answer in a whole sentence, we wouldn't say, "I am going" but rather the contracted form, "I'm going."

These kinds of discrepancies between ordinary and stilted language highlight the fact that spoken language and written language differ. Both forms change over time, but changes in written language usually lag behind changes in spoken language. You might object that response B would be acceptable in conversation if an authority figure had said A. That is, the person who produced response B might be using polite or formal language because of the situation. We exploit these discrepancies in language in our various types of conversation—using more formal styles of language with teachers, employers, and doctors, for example, and less formal ones with siblings, best buddies, and other people from our own age group and background. Computer programs that aim for naturalness must face these issues, which human speakers take for granted.

However, these matters do not interfere with comprehension, and if they were the only types of problems computer programs had to face, the outlook for natural computer-human conversations would be good. So let's look at other conversations that raise thornier problems. Consider conversation 2, which starts with this question:

D. What's up?

Now consider these possible responses (and don't worry about how the question in D would be written):

E. Not much.
F. Got an exam in the morning.
G. Party on Parrish Beach.
H. A preposition.
I. North, of course.
J. Google.

If you're a student, you might answer E, F, G, or a whole range of other possibilities. If you're taking a linguistics class, you could

well answer H. If you're standing in front of a wall map, you might answer I. If you're reading the stock page, you could answer J.

In other words, utterances take place within a context. Although sometimes the context imposes only minimally on the response (as in the situation in which a student is asking another student whether anything special is going on), at other times it allows for only a small range of appropriate answers.

The sensitivity to context covers not just particular situations but also information about cultural habits and facts about nature, mathematics, history, and so on. In short, just about anything can form the relevant context for a conversation. For example:

K. We're getting married, Dad.
L. Honey, come on in here, and bring four champagne glasses.

The father is communicating his approval and joy in sentence L by drawing on our knowledge that drinking champagne is a celebratory act.

M. I vomited again this morning.
N. Oh, my god. When's it due?

In interpreting sentence N as an appropriate response to sentence M, we are relying on the fact that morning sickness is often an early sign of pregnancy.

O. He's trying to draw a map of hypothetical countries that requires more than four colors in order to allow every contiguous country to be a different color.
P. The fool.

To see sentence P as a sensible conclusion given sentence O, we're relying on the fact that it is mathematically impossible to need more than four colors to draw such a map. A computer program that aims for natural conversation must somehow be sensitive to context in all of these different ways. These are gigantic problems, and they won't go away because people move into different places, using their various senses to understand their surroundings. They learn by experiencing the world in a number of ways. The reservoir of their accumulated knowledge feeds their language interactions (see also the discussion of pragmatics in chapter 2). Computers, on the other hand, do not gain experiential knowledge that allows them to understand contexts like those in the preceding examples because they don't have senses or cognitive abilities.

We could easily make mobile computers for home and the office; robots are already here, of course. In addition, we could equip them with cameras, so that they have something comparable to vision; tape recorders, so that they have something comparable to hearing; sensors for heat and weight and a variety of other physical characteristics, so that they have something comparable to touch; sensors for various scents, so they have something comparable to smell; and sensors for sweetness, saltiness, acidity, and so forth, so that they have something comparable to taste. If we program in enough information, computers might be able to detect the commonest things that human senses detect. But what about the less common ones? What about the nuances, many of which we rely on in understanding our world?

Computers may very well learn. One famous example is the chess-playing computer Deep Blue, which can adapt to new strategies as a chess game progresses. Nonetheless, the ways in which computers learn will limit what they can learn. Deep Blue, for example, can search through millions of possible move sequences, but it is by no means clear whether the heuristics and strategies Deep Blue uses amount to understanding a chess position. It is fast

but narrowly limited. Analogous limitations will affect any computer's ability to manipulate language.

You might object that problems involving context are not so much problems of language as problems of communication in general. Perhaps computers would fare better if we limited ourselves to purely linguistic matters. Consider the following examples:

Q. John ate the pie on the windowsill.

Certainly the pie was on the windowsill, but was John also? He might have been. Compare these sentences:

R. John ate his dinner in the restaurant.
S. John ate the candy in that cute little bag.

For sentence R, John was in the restaurant when he ate. For sentence S, John was not in the bag when he ate. A computer must be programmed to allow the prepositional phrase in all of these sentences to be analyzed as either modifying the noun that immediately precedes it (yielding the sense in which the noun was in that place but the eating did not occur there) or as modifying the verb (yielding the sense that the eating occurred in that place). How does the computer make such a choice when it doesn't know what a restaurant is and what a cute little bag is and when it has no knowledge of John's relative size to either of these places? Certainly we could feed the computer a database that notes which words co-occur with which structures since in the real world there are probabilities associated with situations (people are more likely to eat in restaurants than in bags). However, if our sentence were describing a situation with a low probability, the computer would assign the wrong structure even when a human might well understand the correct one from the context. Also, if the database does not happen to cover the relative probabilities of two structures for a given set of words (for example, if we were talking of eating candy with respect to closets), the computer would be at a loss. In that case, we could build in a default structure since it is likely that one structure occurs with more frequency than another in general. However, the chance of accuracy is only as good as the probability of that structure occurring in general.

It has been estimated that if we considered only sentences with twenty or fewer words in English, we'd have a database of approximately 10^{30} sentences, and many of them could be understood in multiple ways. Consider just this silly little sentence:

T. I saw her duck.

In example T, I might have seen an animal or an action, or I might be performing the gruesome action of sawing a duck. In some varieties of English, I could also have seen a bra strap (*duck* can have that meaning.) And sentence T has only four words. The task with longer sentences is daunting.

Moreover, think about these sentences:

U. Why's Virginia so mad?
V. My sister lost her book.

In response V the sister could have lost her own book or Virginia's book or even some third person's book, although to get this last sense we would probably need some preceding sentences or other context. Contrast sentence V to the possible answer W:

W. My sister lost her cool.

In contrast to sentence V, sentence W has only one interpretation for *her*—my sister. How does the computer recognize this fact?

Commands present additional types of problems. There's a considerable difference between asking a computer to follow a command such as:

X. Record "Law and Order" at 9 P.M. on Channel 10.

and the following command, which my students thought of:

Y. If there's a movie on tonight with Harrison Ford in it, then record it. But if it's *American Graffiti*, then don't bother because I already have a copy of that.

Whereas I'm optimistic about computers being able to follow command X someday, a command like Y is more difficult. Command X involves activating the "record" function on a DVR (digital video recorder) and selecting a particular channel at a particular time. It doesn't even ask the computer to scan a list of TV programs. That is, "Law and Order" will be the only program shown at 9 P.M. on Channel 10. However, command Y asks the computer to scan a list of TV programs, recognize which ones are movies, filter out the particular movie *American Graffiti*, determine whether Harrison Ford is an actor in the remaining movies, and then activate the "record" function on the DVR at all the appropriate times on all of the appropriate channels. These are Boolean tasks, and Web search engines perform them all the time. However, normally we feed search engines operators like *and* and *not* and key vocabulary items that the operators have scope over (that is, have in their domain). In an order like Y, we'd be asking the computer to work from ordinary sentences, extracting the operations and then properly associating them with the correct vocabulary items, a much harder task.

All of the issues discussed in this chapter will arise no matter what human language a computer is dealing with. The problems are compounded when we ask computers to deal with more than one language, as in computer translation programs (see chapter 3 for issues in translating). For these reasons, I seriously doubt that people will ever be able to have conversations with computers that are indistinguishable from those with humans. I want to leave you with one final conversation, in which I give you a question and a list of possible answers. I leave it to you to think about the difficulties these examples present for computers:

Question: Why won't you go into that room?
Answers: Spiders.
Superstition.
No shoes.
No windows.
No reason.
I won't tell you.
Guess.

Further Reading

Deep Blue. http://www.research.ibm.com/deepblue/meet/html/d.3.html (accessed May 3, 2009).

Harley, T. A. 2008. *The psychology of language: From data to theory*, 3rd ed. New York: Psychology Press.

IEEE International Workshop on Robot and Human Interactive Communication Proceedings. 2000. New York: Institute of Electrical and Electronics Engineers.

Jurafsky, D., and J. H. Martin. 2000. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, N.J.: Prentice Hall.

McCarthy, J. 2007. What is artificial intelligence? http://www-formal.stanford.edu/jmc/whatisai/whatisai.html (accessed May 3, 2009).

Meeden, L. Developmental robotics. http://www.cs.swarthmore.edu/program/research/meeden.html (accessed May 3, 2009).

Neural information processing systems. http://nips.djvuzone.org/index.html (accessed May 3, 2009).

Sixth International Conference on Neural Information Processing. 1999. *Minds and machines: Journal for artificial intelligence, philosophy, and cognitive science*. Norwell, Mass.: Kluwer Academic.

Smith, G. 1991. *Computers and human language*. New York: Oxford University Press.

Keywords

computational linguistics
computers and language