Sketch-interpreting software

February 19 2010, by Larry Hardesty

(PhysOrg.com) -- Science writers know as well as anyone how much information a diagram can contain. We often labor to express in words what a researcher was able to convey in a single image. But while a drawing can be rich in information, it's information that's usually inaccessible to computers. If you draw a diagram on the screen of a tablet computer, like the new Apple iPad, the computer can of course store the drawing as an image. But it can't tell what the image means.

MIT researchers intend to change that, with a new system that can interpret sketches. If a chemist, for example, uses a stylus — an inkless plastic pen — to draw a molecule on a tablet computer, the software can identify different types of chemical bonds and element symbols and determine the structure of the molecule. Similarly, if an electrical

engineer draws a circuit diagram, the software will identify the circuit's separate components — like resistors, capacitors, batteries, and simple wires — and display them in different colors.

Once a sketch has been interpreted by computer, it becomes much more useful. A chemical sketch, for instance, could be the basis for a literature search, to see whether there's any prior research on the same molecule; analysis software could determine whether the circuit depicted in a sketch will perform as intended. Or design software could simply clean up and standardize a sketch for display in a journal or PowerPoint presentation.

The writing's on the wall

The application of sketch recognition to chemistry grew out of a collaboration with a prominent drugmaker, says Tom Ouyang, a PhD student in MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), who developed the new system together with CSAIL professor Randall Davis. "We once visited their labs, and we noticed that on all their whiteboards and even on some of their windows they had all these chemical structures drawn using dry-erase markers, and when we talked to them they mentioned that they used these graphical diagrams all the time." Currently, Ouyang explains, the only way to translate such diagrams into a format that a computer can understand is to use software that requires the researcher to select an element — like a bond or a chemical symbol — from an on-screen palette, click it, drag it across the screen, drop it into place, and then repeat the process for each successive element. "That's not as intuitive or as fast as just being able to jot it down on paper," Ouyang says.

Most of today's tablet computers and even some smart phones come with software that can recognize handwriting. But interpreting a diagram is "completely different from handwriting recognition," says Tom

Stahovich, an associate professor of mechanical engineering at the University of California, Riverside, who researches sketch recognition. "When you do handwriting recognition, there's a natural temporal and spatial order to it. In English, you write left to right, top to bottom. And so figuring out what comes next is much easier." In a circuit diagram, on the other hand, a resistor might be oriented horizontally or vertically, and it might appear above, below, or next to the preceding circuit element. "With handwriting recognition," Stahovich says, "you keep looking to your right, and you see the next letter." Similarly, Stahovich explains, handwriting recognition systems exploit regularities that are unique to language. "They have a lexicon, just a giant word list, and they find the word most similar to what the recognizer produces," he says. "So if the recognizer recognizes the word as 'tbe,' that's not in the lexicon, but the most similar word from the lexicon is 'the,' so that will get replaced."

Anatomy of a sketch

To meet the particular demands of sketch recognition, the MIT researchers combine information about the physical appearance of the final sketch with information about how it was drawn: the system can recall the direction in which the stylus was moving when a particular stroke was made. That gives it a better sense of whether a stroke was intended to be horizontal, vertical, or diagonal. The system then decomposes a symbol into its constituent parts: its horizontal elements, its vertical elements, its diagonal elements from both upper left to lower right and upper right to lower left, and the endpoints of the strokes. Algorithms automatically refine the components to eliminate stray marks and enhance intentional ones. Finally, the system searches through a database of similarly decomposed sample symbols, looking for matches. Davis and Ouyang say that samples from only 10 or 12 subjects were enough to make both the molecular-sketch and circuit-diagram systems highly reliable, even for first-time users.

"Traditionally, there's been some distinct flavors of shape or symbol recognizers. Some looked at how the shape was drawn — how many pen strokes and in what direction — and some looked at the final image," Stahovich says. "What Tom has managed to do is come up with a technique that combines strengths of both approaches in a unique way."

The researchers have already developed an additional program that translates hand-drawn chemical sketches into a format recognizable by chemical-design software, but they haven't yet done the same for electrical-engineering sketches. And while their system recognizes standard symbols for chemical elements — H for hydrogen, C for carbon — it hasn't yet been trained on the large number of abbreviations that chemists use for more common molecular structures — "like AC for acetyl groups, or ME for methyl groups," Ouyang explains.

Ultimately, however, the researchers see the software as part of a larger project to make interactions with computers as natural as interactions with human beings. "We want to interconnect this with some of the other things we've done with speech and web-based lookup so that one could walk up to the whiteboard and sketch a molecule and say, 'Has anybody published anything like this?'" Davis says. "And then there's the multimodal aspect of that, which is, I draw it, ask if it's ever appeared, and the system says, 'I can't find anything like it.' And I point at the corner of the molecule and I say, 'What if I put a methyl group there?' Not draw it, but just gesture at it." Davis says that other members of his research group are working on the disparate technologies that would help enable such a flexible system.

In the short term, however, if the iPad helps bring tablet computers to a broader audience, sketch recognition could also come into its own. "Previously, the technology was looking for places to be used," says Stahovich. "Now, there's hardware everywhere in need of this technology."

Provided by Massachusetts Institute of Technology