Turning reviews into ratings

February 3 2011, By Larry Hardesty

Graphic: Christine Daniloff

The proliferation of websites such as Yelp and CitySearch has made it easy to find local businesses that meet common search criteria -- moderately priced seafood restaurants, for example, within a quarter-mile of a particular subway stop. But what about the not-so-common criteria? How big are the portions? Are diners packed too closely together? Does the bartender make a good martini?

That kind of information often turns up in reviews posted by site users, but finding it can mean skimming through pages of largely irrelevant text. A new system from the Computer Science and Artificial Intelligence Laboratory's Spoken Language Systems Group, however,

automatically combs through users' reviews, extracting useful information and organizing it to make it searchable.

The first thing the system does is determine the grammatical structure of the sentences that compose the reviews and sort the words used into adjective-noun pairs. If, for instance, someone has written, "I found the martinis to be excellent," the algorithm extracts the phrase "excellent martinis."

As the group's name might imply, its principal area of research is computer systems that respond to spoken language, and indeed, the interface for the new system is speech-based: A user looking for seafood restaurants, for instance, simply says "Show me seafood restaurants" into the microphone of either a computer or a cell phone. Likewise, the algorithm that does the grammatical analysis is one that Stephanie Seneff, a senior research scientist with the group, began developing 20 years ago as a component of speech-recognition systems. Seneff and her grad student Jingjing Liu applied the algorithm to the substantially different problem of parsing written text with very little modification and even less certainty about how it would fare. "We ran it, and we were absolutely delighted with how well it worked," Seneff says.

Seeing sense

The algorithm produces its adjective-noun pairs — like "excellent martinis" or "friendly vibes" — based purely on the words' positions in sentences; it has no idea what the words mean. Fortunately, many review sites allow users to provide numerical scores for some aspects of their customer experience. In work presented at several different conferences sponsored by Association for Computational Linguistics, Liu and Seneff developed a second set of algorithms that use numerical ratings to infer adjectives' meanings. If people who describe food as "excellent" consistently give it five out of five stars, and people who describe food

as "horrible" consistently give it one out of five stars, then the system deduces that "excellent" probably indicates greater customer satisfaction than "horrible."

Once the system has calibrated a set of adjectives against numerical scores, it uses them to infer the meanings of still other words. For instance, if the service at enough restaurants is consistently described as both "horrible" and "rude," the system concludes that "rude," like "horrible," is a term of opprobrium. Similarly, if the adjective "rude" is frequently paired with nouns like "service," "waiters" and "staff" — but not with nouns like "view" or "parking" — then the system deduces that "service," "waiters" and "staff" are thematically related terms.

As a consequence, if a user asks the system to identify restaurants with nice ambiance, its list of search results will include restaurants described as having, say, a "friendly vibe." The system can also use information gleaned from the sites of the businesses under review to expand its semantic repertory. If, for instance, the foie gras and bisque at some restaurant are consistently praised, and they both turn up, on the restaurant's website, under the menu heading "appetizers," then the system will include the restaurant among those with good appetizers, even if the word "appetizer" never appears in any of its reviews.

Xiao Li of Microsoft's Speech Research group says that extracting quantitative ratings from unstructured reviews is a hot research topic both in the academy and in industry and that several commercial products already offer some version of the same functionality. "But you can always do it better," she says. The MIT researchers' work is distinct, she says, in that "they do a lot of linguistic analysis." Other systems, for instance, might try to infer relationships between words without first determining their parts of speech. Which approach will prevail remains to be seen, she says, but she adds that the abundance of research in the area demonstrates that the work has obvious practical import.

Two prototypes of the MIT system, both with speech interfaces, can currently be found online. One takes commands in Chinese and contains information on businesses in Taipei, Taiwan, and the other takes commands in English and includes information on businesses in Boston.

Another grad student in the group, Alice Li, has used similar techniques to extract information from online discussions of patients' experiences with pharmaceuticals. In a yet-unpublished paper, Li, Seneff and Liu present evidence that certain types of cholesterol-lowering drugs may pose a significantly higher risk of some neurological side effects than their alternatives.

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology