ALSO FEATURED THIS ISSUE

FINE-TUNING "SMART" RADIOS

Machine Translation Inching toward Human Quality

Jan Krikke

After 50 years of research and tinkering, machine translation might be ready to compete with human translators. Several companies have announced breakthroughs or substantial progress in MT research in recent months. In January, for example, Steven Klein, CEO of New York-based Meaningful Machines, announced that his company successfully tested new translation algorithms that he says could lead to translation engines replacing human translators.

"While our current prototype is already outperforming other systems on limited resources," says Klein, "we expect to see significant improvement to our quality as both the target language corpus and the dictionary continue to increase in size, with a realistic goal of reaching human quality."

"Although the prototype is only partially complete," says Klein, "we recently began blind testing from Spanish to English, and our system is already performing at higher quality levels on the BLEU (Bilingual Evaluation Understudy) scale than any system we are aware of—0.6092. Systran, whose Spanish-to-English system is one of the best, scored a 0.5494 when we ran it through the same test, and the Systran system has been through many decades of development and incremental improvements."

Meaningful Machines' test has not been independently verified, and the goal of reaching near-human quality translation will probably depend on some degree of pre- and post-editing for years to come. But, the growing number of global corporations (such as Philips, Samsung, and HP) and international agencies and institutions (such as the UN and the European Commission) using the technology illustrates that machine translation—the first nonnumerical application of AI—is finally delivering practical solutions. Popular perception of MT has suffered from low-quality "gisting" translation that Web-based translation engines, such as Babelfish and other online services, generate. But MT engines designed for limited domains, and tailor-made systems that use controlled language, are already delivering services.

Rules and statistics

The Japanese Patent Office's Web-based MT engine instantly translates Japanese patents into readable English. The site makes available a wealth of information previously inaccessible to non-Japanese speakers.

MT has also made it to the desktop. Germany-based linguatec language technologies' MT system translates corporate email and other business documents between several European languages. The system is self-learning—it improves over time as its associative memory grows.

MT requires complex cognitive operations to perform a seemingly mundane task: decoding a source text and recoding into the target language. The three common methods are *rule-based MT* (RBMT), *statistical MT* (SMT), and *example-based MT* (EBMT). RBMT parses text and typically creates an intermediate symbolic representation to generate a translation in the target language. The method relies on large sets of rules and on syntactic, semantic, and morphological information. RBMT poses enormous challenges because it must deal with infinite exceptions to the rules.

SMT and EBMT rely on large collections of parallel (human) translated documents, or *bilingual corpora*. The translation engine looks for parallel phrases and ranks them probabilistically. In theory, the larger the corpus, the better the results.

SMT involves decomposition, matching, and extraction based on individual source language words, whereas EBMT involves decomposition, matching, and extraction based on word sequences and fragments.

Hybrid systems

Several companies are working on RBMT-SMT hybrid systems with the goal of creating systems that can develop an understanding of languages by analyzing human-translated documents. Language Weaver, a company founded by University of Southern California researchers, develops algorithms that handle both learning and translation.

"The learning aspect looks at words and phrases to build a database of phrases for possible substitution," says Language Weaver chief scientist Kevin Knight. "Then the translation algorithms work somewhat like computer chess does, looking at millions of options for word sequence and assigning probability scores to the most likely translation."

Scanning huge corpora for identical or similar phrases must be done in a limited time—the reason Language Weaver is refining algorithms to include syntactic data. "Bringing in some more grammar into the database will increase the likelihood of the accurate response," says Knight. "However, it is still different from the traditional rule-based methods because we continue to look at it computationally rather than strictly linguistically."

Language Weaver has released translation software for European and Asian languages, including Chinese. Prices for European language versions range from US$5,000 for a standalone system to $25,000 for a bidirectional server license. Asian languages range from $15,000 for a standalone system to $125,000 for a bidirectional server license.

Cutting costs with MT

The already huge global translation industry is growing rapidly. According to the American Localization Industry Standards Association, the translation industry's annual production value exceeds $13 billion.

MT can be an enormous time and money saver. Human translators average 3,000 to 5,000 words per day at a cost of $0.05 to $0.20 per word. The UN, the EC, and multinationals such as HP, Philips, and Samsung require translation of many millions of words a year. Microsoft's localization of its new Visual Studio 2005 for eight markets required translation of 120 million words.

With the Internet's continued expansion in the non-English-speaking world, MT's demand can only grow, especially in China. English is the native tongue for eight percent of the world population, but Chinese is the native tongue of 18 percent. Moreover, as Meaningful Machines' Klein notes, "There's a very large dormant demand for translation in areas that no one thinks about now because it would be time- and cost-prohibitive. With next-generation MT, market demand for these applications will suddenly materialize because fast, high-quality translation at very low cost will be available."

MT on the desktop

In Europe, hundreds of companies have integrated translation engines into desktop software. linguatec recently launched the latest versions of its Personal Translator, a customizable solution that translates business email, corporate documents, and Web pages into various European languages. Multinationals such as Siemens and Lease Plan use the software, which integrates into Microsoft Office applications. Customers can add customized dictionaries to tailor the software to specific domains.

linguatec's technology, a three-time winner of the European Information Technology Prize, has a rule-based backbone architecture. The architecture is extended by probabilistic components for automatic subject-area recognition, and syntactic analysis, where the evaluation of competing analysis results requires probabilistic information. "We call it neural transfer," says Gregor Thurmair, linguatec chief developer. "A statistical component is integrated into the rule-based possibilities and uses automatic learning methods to identify conceptual contexts and trigger the right translation."

The neural transfer technology mimics the associative powers of the human brain. Using linguistic and neuroinformatics methods, the program analyzes a corpus that now exceeds 1.9 billion words to identify which concepts commonly occur in context with each other. The program collects contexts from the corpus for all terms that undergo neural transfer. English words with multiple meanings, such as *plant*, trigger a different German translation depending on whether the context is industrial or botanical. The system improves as the corpus and the associative memory expand.

No single metric for quality

A pioneer of hybrid (RBMT-SMT) systems, linguatec was involved in building the corpus for Google's translation technology. Google now relies on Systran technology, but it's expected to launch its own translation technology in the coming months.

The search engine giant came out on top in last year's National Institute of Standards and Technology machine-translation test, achieving the highest rating in both Arabic-to-English and Chinese-to-English. The annual NIST test involves the translation of 100 newswire articles published by the Agence France Presse and China's Xinhua News Agency. Each source set had four sets of independently generated human translations. NIST used IBM's BLEU metric to measure quality.

"BLEU measures translation accuracy according to the N-grams, or sequence of N-words that it shares with one or more high-quality reference translations," NIST wrote in announcing the results. "The more co-occurrences the better the score. BLEU is an accuracy metric, ranging from 0 to 1 with 1 being the best possible score."

However, NIST adds that BLEU can't distinguish subtle differences in high-quality translations. "At the present time, there is no single metric that has been deemed to be completely indicative of all aspects of system performance," the organization explains. NIST also notes that the participants were required to submit only their translation system output, and that the systems themselves weren't evaluated.

In the latest test, Google outperformed established players such as Systran by a considerable margin. The company scored 0.5131 for Arabic-to-English and 0.3531 for Chinese-to-English. Systran's score was 0.1079 and 0.1471, respectively.

"Google works on a fully statistical translation system," Thurmair says, giving the scores some context. "They ran the competition on about 3,000 parallel machines. Even with this power, it is not clear the translation quality is superior to a well-designed rule-based or hybrid system. The measure used in the evaluation does not reflect a fundamental requirement of translation, namely that it should produce grammatically correct sentences."

Machines have no real understanding?

Some human translators are skeptical about MT. Steve Vitek, a professional patent translator of Japanese-to-English and German-to-English, wrote a widely circulated essay in 2003 entitled "Reflections of a Human Translator on Machine Translation." In it, he argued that human translators will always be needed. Many documents require an understanding of the source texts to produce an accurate translation, he wrote, citing as an example manuals with illustrations showing how to assemble equipment.

Vitek still remains unconvinced. "A machine by definition will never 'understand' anything," he says. "It can only simulate understanding based on rules that

must be constantly input and modified by a human being who is in fact capable of real understanding. People who think that MT will one day replace human translators are misled by companies who are trying to sell them their products."

He does admit, however, that progress is being made. He notes that the service of the Japanese Patent Office Web site, which offers free online translations of recent Japanese patents, comes "pretty close" to translating the real meaning of the original Japanese sentences.

Beyond MT

Domain-restricted translation has improved in large part because humans producing the source text understand that they must meet MT systems halfway. They avoid usage that's likely to confuse the system, such as figurative speech and abbreviations. Even then, MT might not generate perfect translations, but it will produce the next best thing: usable results.

Language Weaver's Knight says MT might yet surprise us by going beyond translations. "Computers have famously classified stellar data, proposing new classes of stars for astronomers," he says. "Likewise, many things could happen in linguistics. For example, we know that a parallel United Nations document contains the same basic ideas in five different languages. With enough text, the computer may be able to come up with a proposed representation of the underlying ideas themselves, not just the words and syntax that carry those ideas."

Fine-tuning "Smart" Radios

Benjamin Alfonsi

Outside the tech community, the big news in radio is, of course, satellite radio. Within the community and AI circles, cognitive or "smart" radios are making a splash.

Paving the way is Virginia Tech's Center for Wireless Telecommunications Cognitive Wireless Technology (CWT[2]) group. The team recently won a three-year US National Science Foundation grant to continue its research into developing and deploying *cognitive-radio transceivers*, or CRs.

A CR combines AI with *software-defined radio* (SDR) technology to create a transceiver that's simultaneously aware of the radio frequency environment, legal operation policies, its own capabilities, and user needs.

The AI approach

Smart radios learn from experience, says CWT[2] head Charles Bostian, a Virginia Tech professor of electrical and computer engineering. A CR consists of a *cognitive engine*—a hardware-independent software package that can fine-tune itself. The cognitive engine sets the SDR's operating parameters (turns the knobs), observes the results (reads the meters), and optimizes its operation under the governing rules.

The CWT[2] team, which includes Bin Le, David Maldonado, and Thomas Rondeau, stumbled on its discovery while working on another project. "In 2002, our group discovered that we needed CR capabilities for a rapidly deployable, high-data-rate communications system that we were developing for disaster response," recalls Bostian. "The concept of a cognitive engine followed."

The team's first breakthrough was developing a computationally efficient way of implementing rapid machine learning in a trial-and-error process. The process used genetic algorithms to set the knobs and hidden Markov models to represent the radio channels.

"We used a proof-of-concept prototype of our cognitive engine to control a 'dumb' legacy 'hardware' radio," Bostian explains. "The resulting CR could identify the presence of a jammer and change its modula-

tion index, transmitter power, and FEC [forward error correction] coding in a way that minimized the effects of the jammer. This prototype demonstrated learning, and we were off."

Bostian describes the cognitive-engine solution as a tiered system of highly integrated AI techniques. "We are trying to push AI systems into a realtime reconfiguration system that learns and optimizes itself based on the user's needs," he says.

The team's approach to realizing the machine intelligence the CR requires for proper operation is strongly rooted in standard AI methods, including genetic algorithms, case-based reasoning, and neural networks. In each of these categories, however, Bostian says his team is pushing beyond standard implementations.

"We have been criticized in the past for using genetic algorithms because of the length of time required to converge on the optimal solution," he says. "And while the cognitive radio is expected to optimize itself in real time, the CR does not need to find the optimal solution, just a better solution."

When the CR observes a new situation that requires action, the case-based decision maker compares the new observation with previous observations and actions in its case base. The case-based decision maker calculates the case-base item's similarity and utility to the incoming observation. It selects the item that maximizes both utility and similarity and sends its information about the actions to take to the genetic algorithm optimization process. The case-base item's action information isn't the exact action that the radio will take. Instead, the information gives the genetic algorithm a good starting point in its search for the optimal solution and directions for finding the optimal solution.

"How the case base calculates the similarity and utility functions is a large part of our research, as is the interaction with the genetic algorithms," Bostian says. "Our goal is to tie these two AI systems together to create a genetic algorithm that optimizes a radio in real time."

According to Bruce Fette, a chief scientist at General Dynamics Corp. and a founding member of the Software Defined Radio Forum, "The extension of the genetic algorithm, 'survival of the fittest,' working at the physical layer of the radio could be a very important development in AI and machine learning applications."

Possible applications

Although Bostian estimates that a large-scale, full-function version of the CR could be more than two years away, his team has already defined two likely applications.

The first is in the realm of public safety communications. "We are developing a prototype CR that will recognize networks using any of four common public safety waveforms and configure itself to communicate with them," says Bostian.

Fette says this type of application would fill a crucial role. "Being able to provide a common communications bridge between public safety organizations and various emergency responders addresses a critical concern of the Departments of Defense and Homeland Security," he says.

But technology analyst Monica Paolini has reservations. "In public safety, you want the most reliable and resilient hardware you can find, and cognitive radios are unlikely to meet this criterion in the foreseeable future," she says. "The risk here is that you have a cognitive radio that creates interference to other users, and that is really the last thing you want to happen during an emergency."

The second possible application is in dynamic spectrum allocation, using vacant TV channels as a test case—although some people disagree about whether such channels actually represent unused spectrum. This industry segment, of which CWT[2] is clearly a part, wants to deploy IEEE 802.11-like access points to operate in vacant TV channels.

Bostian says a CR would know its location and from that would be able to determine which channels were potentially available. It would then configure itself to avoid causing any interference to the licensed users and negotiate with its unlicensed peers to find a way to share the channel with minimal interference.

Paolini sees the benefits of this application since, at any given time, large swaths of spectrum sit unused. Her only caveat is the CR's ability to navigate the regulatory framework.

"From a regulatory point of view there is a need to limit the power of cognitive radios; otherwise, they will never be used," she says. "A CR that can use any type of spectrum available is unlikely to get regulatory approval because it would have to prove that it cannot interfere with licensed spectrum holders in any band." ■