Translators &
Computers

Steve Vlasta Vitek was born in Czechoslovakia, where he majored in Japanese studies and received his M.A. (Magister of Arts) degree from Charles University in Prague in 1980. He emigrated through West Germany, where he lived and worked in 1981-82, to the United States in 1982. He later moved from San Francisco to Tokyo, Japan, where he worked for a Japanese company as a translator from Japanese and German into English between 1985 and 1986, when he returned to US to start his freelance translation business in San Francisco in 1987. Since then, he has been earning a living as a freelance technical translator, mostly from Japanese, but also from German, Czech and Slovak, and occasionally also from Russian, Polish and French. He now lives in the city of Santa Rosa in the heart of the Wine Country in Northern California, 1 hour north of San Francisco. After almost 20 years of living in different countries with no one to speak to in his native Czech, he does not have a native language any more. His sister, who is Czech, told him that he speaks funny Czech, his wife, who is Japanese, told him that he speaks funny Japanese, and although his monolingual children who speak only English don't dare to tell him that his English is funny too, this is probably because they want to stay on his good side as they find his ability to translate the text on their Japanese Pokemon cards into English very handy. He has three dogs who are trilingual—they obey commands in English, Japanese, and Czech, although they seem to prefer Japanese, which is a much more expressive language than English or Czech, and in spite of its shortcomings, clearly a superior language for the purposes of communication with intelligent canines. Steve Vlasta Vitek can be reached at stevevitek@pattran.com; his website is at http://www.japanesetranslators.com.

# Reflections of a Human Translator on Machine Translation

## or

## Will MT Become the "Deus Ex Machina" Rendering Humans Obsolete in an Age When "Deus Est Machina?"

*by Steve Vlasta Vitek*

*There's a tremendous gulf between microprocessors and brains. People still don't know exactly how the brain works. Brains remain quite a mystery. ...... A machine can collect lots of data and make inferences and judgement. It becomes philosophical: What does it mean to understand?*

*You can create something that has the look of understanding and the feeling of understanding. But it is what lets us make a leap to understanding something and make insights into something else. This is one thing that electronics can't do.*

*Michael Slater, publisher of the Microprocessor Report (in Sebastopol, California).*

*E*very now and then Federal Express delivers a thick envelope to my home office with at least a dozen Japanese patents and a cover letter from a patent lawyer asking me to provide a price quote for translating the whole package. After the initial rush that comes from seeing big dollar signs when I am done counting the pages and multiplying them by my estimate for the number of words to arrive at the dollar

figure, I try to calm down as I prepare a fax to the lawyer. More often than not, the price is too high and the lawyer has to find a way to argue the technological aspects of a patent case without knowing all the details of what is in all those patents. There are, however, ways to get around the cost problem. Sometime I am asked to translate only the patent claims of all the patents, which will reveal some of the needed information at a fraction of the cost.

> although most of the technical terms used by a machine will be correct, it is up to the reader to make sense of those words haphazardly jumbled up together by a non-thinking machine.

There are also English summaries online, available for 5 dollars, describing in peculiar, but usually fairly understandable, English, written by native Japanese speakers, the gist of the design. Another alternative is to have the package translated by a machine at a fraction of the cost of human translation. One can achieve some kind of understanding of the design from the words thrown at you by a machine, especially if the reader of the translation is an expert in the field. It is much cheaper to use this option—the average cost for translating a machine-translated patent is about 60 dollars, while the average cost of human translation is at least ten times higher.

The problem is that the machine does not understand the meaning of the document at all. Therefore, although most of the technical terms used by a machine will be correct, it is up to the reader to make sense of those words haphazardly jumbled up together by a non-thinking machine. The following is a random example of commercial machine translation of a short section taken from a simple Japanese patent. The translation was obtained from an online search service offering among other things machine translation to its customers:

> "circle 1.. In case of mask which uses metal sheet. You explain making use of Figure 1. pattern a which corresponds to mark "A" in metal sheet 4 is formed, the metal sheet 4 must be formed with photograph etching and not. As for this pattern b because of notch type, bridge 11 in order to prevent the coming out portion of metal sheet become necessary. As a result, mark "A" which marking is done is not correct mark "A" always in object to be marked, it becomes mark where portion of bridge 11 lack. Because of this, it was a eyesore even in eye where portion which lacks existed in mark "A", saw, there was a possibility which the mark misperception is done."

In case you are wondering what the text above actually means, this is how this imperfect human translator would translate the same paragraph:

> *"(1) Figure 1 indicates a case when a metallic plate is used for a mask. In order to form pattern "a" with a corresponding mark "A" in metallic plate 4, the metallic plate must be formed with photoetching or a similar process, including a notch in the pattern, and bridge 11 must be formed to prevent partial detachment of the pattern from metallic plate 4. The result is that the marking substance will not necessarily form a precise mark "A" which can be used for marking, but rather, the mark will be formed with a deficient part containing the bridge part 11. That is because the bridge part is normally left in the pattern, although this not only creates a visual distraction, but it can also cause a mark recognition error."*

## A Picture Is Worth a Thousand Words

Even this translation, done by an experienced translator who translated thousands of similar patents from Japanese and other languages, may still not be completely clear unless the reader can see the accompanying Figure 1 and understands how marks are used during the manufacturing of electronic components. This is particularly true when one translates between two languages that are as dissimilar as Japanese and English. Unlike in European languages, Japanese nouns usually have no singular or plural, Japanese verbs, especially verbs used in patents, usually have no tense, and other grammatical features which will be normally always present in a European language, such as the subject, may be missing in Japanese or they will be replaced by a unique Japanese grammatical feature called "wadai" or "topic" which has largely adverbial characteristics from the viewpoint of Western grammar. It should be noted that Western grammatical concepts are not really applicable to Japanese, because many important concepts and aspects of the Japanese language do not exist in European languages and vice versa. For instance the all-important Japanese grammatical category of a "topic" or "wadai", or the Japanese grammatical category of "particles" or "joshi" simply have no equivalents in Western grammatical systems which were basically developed on the basis of a descriptive grammatical theory designed for Latin, French, English and other European languages. Given how difficult it is to explain all of these linguistic aspects even to a linguist, it must be very difficult to program all these grammatical differences into a piece of software.

However, one look at Figure 1 would explain to a human reader exactly what is meant in the paragraph above. Obviously, I always translate the text while looking at the figures, and I was only guessing the precise meaning of the Japanese text until I saw the figure. There is no way around it—we cannot translate that which we don't understand. The meaning is of paramount importance in the translation process. And unlike humans and chimpanzees, machines by definition don't understand the meaning of anything and never will. This is why machine translation that aims at accurate translation of the meaning of the original text is an exercise in futility, regardless of how many billions of dollars, yens and marks are spent in the pursuit of this elusive aim. MT will never really amount to anything more than a tool, a useful tool for translating words from one language into another, words that do not necessarily say anything about the meaning of the original text at all, except perhaps by accident. The meaning cannot be supplied by a machine—it has to be supplied by a human being. It is possible, perhaps even likely, that a patent lawyer will be able to supply the real meaning of the passage by reading the machine-translated words and looking at the figures. However, most of the time, the machine product will be crude and almost incomprehensible, even if it's a very simple descriptive passage. In my opinion, forcing patent lawyers to go through this process every time when they need to arrive at the real meaning of a sentence represents abuse of very intelligent humans by dumb machines. I would also argue that not even patent lawyers are paid enough to deserve being abused by unfeeling machines in this manner. There must come a point at which the patent lawyer's brain will refuse to play a silly game with a silicon translator, wherein the silicon translator supplies the words in English and the specialist tries to supply the real meaning of these words.

### A Cost-Effective Alternative Or Abuse of Humans By Machines?

At that point, the lawyer may realize that sixty dollars for a machine-translated patent is not such a bargain after all. This is, obviously, what I'm counting on. I think that, all things considered, MT is an excellent invention. It makes sense to pay sixty dollars for a patent translation when you're swamped with dozens of patents which *might* contain the information that you are looking for. And it is probably possible to determine *based on machine translation of words* which patents do not contain the relevant information and which patents may contain this information. At that point, a human translator will probably be asked to supply a human translation containing not only the proper technical terms, but also the real meaning of the Japanese or German, etc., text. It is probably wasteful to go ahead

and translate 24 patents at the average cost of six hundred dollars per patent in order to discover two or three patents that are in fact relevant and important for arguing a case, as a good accountant will be quick to point out. But it makes sense to translate all of them "on the cheap" by a machine and then to ask a human translator to translate one or two important patents as evidence of prior art design. And judging from my interaction with my clients, this is already happening. Machine translation, instead of being the scourge of our profession, is in fact an extremely fortunate development for translators. It broadens the scope of resources that are now available to our clients during a search for prior art (existing design). It is also likely that CAT (computer-aided translation), a close relative of MT, is suitable for translating repetitive passages which need constant updating with only minor changes in each update. It makes sense to quickly and inexpensively update existing translations of weather forecasts, the daily status of water levels on European rivers (which flow through a different country every few miles, as oblivious of national and linguistic borders as Vikings of old time), or perhaps even cellular phone communication manuals, etc., by a machine and then quickly edit them by an experienced human translator who understands the specific conditions and requirements. This kind of translation can be done with the help of CAT/MT provided *the input range is extremely limited and the input, the translation process, and the output is subject to a strict and professional human control*.

### Building a Tower Of Babel from Silicon Bricks

Which also leads me to believe that simple, routine kind of translation work that probably can be done by machines may soon disappear in the quicksand of machine translation, although even this kind of translation will still need human editing, so that the only kind of translation that may be left for me… *is the kind of translation that I like to do!* The kind of translation that requires an intimate knowledge of languages combined with analytical thinking and understanding of an infinite number of situations that can only be achieved by the human brain, with its billions of connections supplying the sparks needed in a mysterious thinking process. I don't think that we understand what really causes thinking to occur and how it all works any better than we did two or three thousand years ago. If we did, we would not be trying to build another Tower of Babel, this time around from silicon bricks. It is an expensive proposition and the payoff may not be really worth it. The expectation that machine translation will soon replace humans, which is still prevalent in the gullible, monolingual public, is—dare I say it—nothing but a new piece of evidence of human ignorance and arrogance. Regardless of the

speed with which computer chips can process information, and in spite of the fact that this speed has increased about 500 times in the last few years, the quantum leap that is a characteristic of human thinking, when we reach a new conclusion based on the information presented to us and based on our human experience, is something that we will never be able to program into a machine. If we could do that, machines could replace not only human translators, but also their clients—patent lawyers, medical doctors, judges, and engineers. To try to reduce human language, which is as complex as human thinking, to a series of zeroes and ones, is clearly an exercise in futility. In fact, it might be easier to input all applicable laws and technical designs into a computer and then use this computer instead of a patent office examiner to evaluate the relative merits of a technical design.

Commercial services offering MT online realize that their product is not really a translation, that is to say a rendition of the *meaning* of the original text into another language. They sometime use a different term for the MT product, such as "gisting" in the sense of a "machine-translated text", and contrast this "gisting" to "custom translation" (i.e. human translation), etc. Machine-translated texts usually also have a disclaimer on every page in which no guarantee is given that the "translation" is accurate and no responsibility is assumed for the machine-translated product. This protection against a lawsuit filed by an angry customer who might have expected a real translation for his money is definitely needed when every single sentence usually contains a number of gross errors. If a translation done by a machine is accurate, it can be accurate only coincidentally because the machine does not understand the concept of accuracy.

Still, it is clear that machine translation will play an important role in the new millennium, helping to bring down the communication barriers in the newly interconnected world. It is up to us, translators, to explain to the general public what machine translation is, what are its strengths and weaknesses, and what is its likely role in the future development of our civilization. We translators have an insight into this problem that other professionals can hardly be expected to have.


### Deus Est Machina!—The God Is a Machine!

About a year ago, I saw on C-Span (a public TV channel in United States covering political and business issues) Raymond Kurzweil, the author of the Kurzweil scanning method for character recognition by software, answering questions about the likely future trends in technological development. He was very optimistic about the future of machine translation. He was

convinced, he said, that machine translation will soon achieve the same kind of accuracy that is now achieved by optical scanners, which can convert printed pages into digital units containing the words printed on the page. I don't know whether he really believes what he was saying, or whether he was mostly interested in giving a little boost to his company's stock or promoting his new book. But his geeky audience was clearly pleased with his answer. That was what they expected. The public wants to believe that machines will soon replace humans and complicated texts containing the result of an extremely complicated thought process expressed in languages evolving continuously for millennia will be soon translated by slightly smarter machines with a faster microprocessor enabling to achieve an accuracy of, say, 95.5%. "Deus ex machina" will soon be replaced by "deus est machina". The Bible will be soon translated by a sheet-fed optical scanner instead of a team of biblical scholars and instead of a hundred years, the whole translation will take only a few hours! That will truly be the New Testament of our age!

And since this exciting technological development is just around the corner, or at the worst, no more than a few decades away, there is really no need to learn foreign languages. All we have to do is design a faster chip and hire a few good software programmers, as Raymond Kurzweil proposes in his book The Age of Spiritual Machines *(L'âge des machines conscientes)* when computers exceed human intelligence.

### And I Thought Silicon Breast Implants Were a Scary Concept!

Kurzweil probably does believe in what he is saying because he believes that human consciousness, a *conditio sine qua non* if we want to create artificial intelligence that would enable machine translation of the real meaning of any text, can and will be simulated by computers in the near future. He also says that humans and computers will merge so that human memories will be downloaded into a machine and mechanical neural implants will be installed in human brains. (And I thought that silicon breast implants were a scary concept!) There are, of course, other scientists and philosophers examining the issue of human consciousness and intelligence who come to the exactly opposite conclusion. For instance, in his recent book *The Mysterious Flame,* the philosopher Colin McGinn argues that evolution itself has so designed our minds that we cannot understand or explain intelligence. (Whether we call something evolution, God, or cosmic intelligence, all of these names are indicative of the same principle—everything happens for a reason and this reason can usually not be understood on the level of human consciousness. Unfortunately, we humans

are capable only of this relatively low level of consciousness, although we may be able to catch from time to time a glimpse of divine consciousness, or evolution if we want to call it that, usually without realizing what is going on).

All I can say is, good luck, Mr. Kurzweil, and more power to you! Thanks, among other things, to your superior machines whose intelligence will presumably soon exceed yours and mine, we human translators can look forward to a booming business in the exciting field of human translation for a few more centuries!