Has auto-translation software finally stopped being so useless?

October 4 2016, by Vidyasagar Potdar

Credit: AI-generated image (disclaimer)

If you've ever put a phrase into an online translator and then laughed at the garbled results, your fun might be coming to an end. Google claimed last week to have eradicated 80% of the errors made by its translation software.

Translating text from one language into another is a simple proposition but a fiendishly complicated problem. Of course, it has traditionally been a job for human translators, but over the past half-century or so automated machine translation has become an important sub-field of artificial intelligence.

Auto-translation systems, including [Google Translate](), were already pretty good at translating single words or even short sentences. But people are well aware of the limitations of this technology when it comes to translating longer, more complex passages, and hence are cautious about relying on them for important tasks.

Machine translation systems work by analysing an input text from one language and creating an equivalent representation in the target language. This can be as simple as word substitution, but such a system cannot guarantee high-quality output. That's because it is difficult to program a computer to understand the text as humans do, and then to translate it to another language while keeping the meaning and semantics intact.

This is partly because different languages originated at different times and have different evolutionary histories. This gives each language a set of unique subtleties that can be difficult for humans to learn, let alone a computer attempting to go from simple word substitutions to intelligible sentences.

Not all words in one language have a direct equivalent in another, so several words might be needed to convey the meaning of a single word in the original language (a classic example being the German *schadenfreude*). The grammatical structure can also be different. Not all languages use the same subject-verb-object format found in most English phrases.

Auto-translation software can also struggle with words that have

different definitions depending on their context. This means that the program will need to analyse the entire sentence or paragraph as a whole to deduce what it means.

Clearly, understanding the broader meaning is critical for producing useful translations. But teaching a machine to derive the subtle meanings of language is no easy task, as the often comical results of older translation software make clear.

Human translators rely on knowledge, experience and common sense, but we don't really know precisely what is going on when the brain synthesises language. If we don't know how it really works, how do we go about teaching a computer to do it?
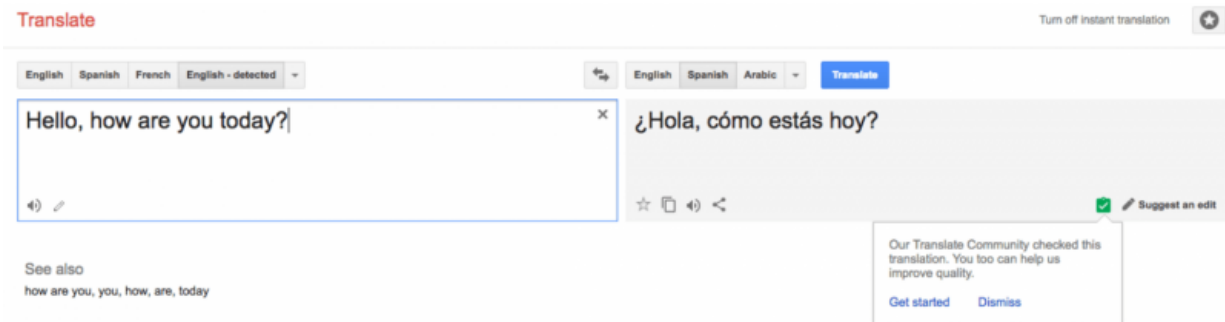
Google Translate: good for short phrases, and about to get better at long ones too. Credit: Google Translate

The machine learning approach

As described above, the real challenge lies in moving beyond individual words or short phrases to translating large pieces of text such as entire

websites or novels.

At the simpler end of the spectrum the technology already does a pretty good job. If you're travelling in a foreign country you can use an augmented reality app such as [Word Lens](#) to decipher street signs in real time. Simple tourist phrases are easily conjured up using programs that have simple language rules hard-coded into their programming.

But say you want to read a novel, or browse a foreign-language website, or translate a PowerPoint presentation in real time at a conference. This needs a new approach – one that recognises and reproduces the flow and meaning of the whole.

Google's new approach involves what it calls "Neural Machine Translation (NMT)". It relies on an artificial neural network which attempts to simulate the human brain's approach to translation. Crucially, it can "learn" as it becomes more experienced, gradually improving its accuracy as it translates more text.

As NMT algorithms do not rely on human logic (that is, hand-coded algorithms), they can modify themselves as they go. In theory, they should be able to find ways to translate text that the human coders might not have conceived when designing the system.

The future

Reaching 100% accuracy will not be easy, but we can expect tech companies like Google to devote a lot of energy to trying. It is likely to be an evolutionary process, not a one-off breakthrough, and it will take huge amounts of time, data and processing power to improve the results until they are effectively flawless.

The latest development nevertheless represents a huge step forward –

finally propelling machine translation to a standard that is acceptable for most tasks. For now, if you need 100% accuracy you will still need to hire a human translator, but with every day that passes computers are honing their skills.

This raises the question of how seamlessly auto-translation will become a part of our everyday experience in the future. In time, we may browse websites that automatically open up in our preferred language based on our profile, or listen to lectures in whatever language we choose, or engage in real-time discussions with people speaking a different language without having a human translator listening in. The opportunities are limitless.

If we can improve the accuracy to almost 100%, language barriers will begin to disappear. We would belong to one global village, where anyone can share their knowledge and expertise with anyone else.

In a world where computers are multilingual, will anyone need to bother learning another language? It's too early to say. But just as mapping software has all but eradicated the feeling of being lost in a strange place, we're heading for a world where you can be anywhere on the planet and never be lost for words.

This article was originally published on The Conversation. *Read the* original article.

Source: The Conversation