Languages in Cyberspace

The Role of Machine Translation in the Global Market

by Wee Li Kwang Angela

The Internet has brought about the blurring of the geographical boundaries of the world, making it easier for businesses in different parts of the world to communicate despite physical barriers. However, businesses are still finding difficulty in reaching out to the global market. There are many reasons for this and one major factor is the language divide.

The world is multilingual but many enterprises are not leveraging on the multilingual opportunities that it presents. For example, many English-language businesses are only targeting the English-speaking market, the Chinese-speaking businesses are targeting the Chinese-speaking. For a business to target different language markets, it has to have multiple business models, each targeting one language market. Imagine a scenario of the business actually targeting the world market with just a single concerted effort, regardless of language constraints.

Figure 1

Machine Translation—The Language Bridge

Machine Translation (MT) is a technology that enables real-time on-the-fly translation of textual information from one source language to another target language. MT, a branch of Artificial Intelligence, emulates the human capability to translate information based on an accumulated pool of linguistic knowledge of both the source and target language, a bank of vocabularies, as well as semantic knowledge of the world.

MT has come a long way. Since the seventeenth century, there were suggestions on the use of mechanical means to overcome language barriers but they failed to address the issue of syntactic differences. The thought of using actual machines to perform translation was not to be until mechanical calculators were invented in the

The Internet has increased the demand for easier and better communication across borders and MT is poised to become the essential technology to bridge this digital language divide in the years to come.

nineteenth and twentieth centuries. In the 1970s and 1980s companies began commercializing the technology. The Internet has increased the demand for easier and better communication across borders and MT is poised to become the essential technology to bridge this digital language divide in the years to come.

Skepticism of MT still exists despite its long history. The main reasons are attributed to MT not attaining 100% accuracy and translators not being able to allay the irrational fear of MT totally replacing the human role in translation. In spite of all the skepticism, MT has its practical strengths in this globalized economy where there will be explosive demand for translation that is beyond the capability of human translators. MT has strengths that could add value to the role of human translators, especially when much of this translation demand requires fast turnaround and only a gist of the original information is required.

MT should best be deployed where its strengths can be leveraged. The technology is best suited for deployment with online applications that leverage its real-time characteristics. MT provides real-time translations with an immediate turnaround within seconds that no human translator can produce. This is critical for time-sensitive information.

The technology is a cost-effective tool for businesses to gain immediate presence in multiple markets and access to a global multilingual market base without much rework of their existing systems. This allows businesses to save resources to focus on their business in one language. It is also a convenient tool for Internet users to freely surf multilingual sites on the Internet, thus allowing the millions of businesses and consumers across the world to interact beyond language barriers.

MT's high processing capacity can also be used as a productivity tool to generate large volumes of first draft translations for post-editing. Another plus factor is that the technology can be further customized for higher accuracy with deep knowledge of the domain field. The accuracy of the technology can be incrementally improved with usage through time.

Practical Applications of Machine Translation

Machine Translation enables all real-time online communication, e-commerce and content creation in multiple languages, revolutionizing Web interaction from search, surfing, e-mail, chat, messaging, Web publishing and e-transaction. This enabling effort is all done at a much faster turnaround (instantaneous) and at a more affordable cost compared to human efforts.

• Localization & Globalization Usage

As companies globalize, they are automatically becoming more Internet savvy and more informed on the need for multilingual content management. The need to have a good multilingual content management workflow with control and monitoring of content authoring and publishing processes becomes important.

Limited by their own language constraints, searches are only done for information in their familiar languages. This search is further aided by search engines that inherit the same language constraints.

Although content authors and translators are using various tools to improve

productivity, they are still facing difficulty coping with the overwhelming amount of content as well as the rate at which it arrives. Hence, MT will become an effective aid to this overall content management workflow, helping to increase the translation throughput.

Businesses that provide localization and globalization services have more important needs for this multilingual content management workflow to provide better service to globalizing companies.

• Multilingual Cross-lingual Search

Traditionally, humans are deployed in all kinds of information retrieval and gathering. Limited by their own language constraints, searches are only done for information in their familiar languages. This search is further aided by search engines

MT can be effectively used for preliminary communication exchange before human interpreters or translators come into the picture.

that inherit the same language constraints. However, information is multilingual in nature. Relevant information pertinent to businesses may be available in foreign languages.
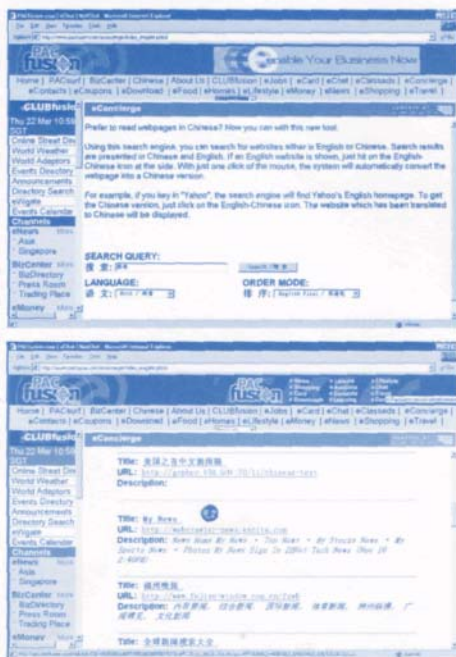

Figure 2

MT combined with free-text search technology empowers multilingual cross-lingual search, where a search conducted in one language can be done across other foreign languages. This multilingual search can either be for general Web information or for specialized domain databases. For example, when a search on "traditional medicine" is conducted, the abundant information about traditional medicine available in the Chinese language is also retrieved.

Figure 2 shows the multilingual search application deployed using EWGate's EWTranslate MT technology on the e-commerce portal site of Pacific Internet. It enables surfers on the site to search for multilingual information on the Web.

Figure 3 shows a multilingual search application using MT technology on a specialized patent database for the Singapore Intellectual Property Office's Patent Search site.

• Multilingual Site Presentation and Web Surfing

Not all companies can afford the resources to develop their company Web sites in more than one language to serve multiple language markets. The most cost-effective technology that can bring them to multiple markets in the shortest possible time frame is Machine Translation. The technology combined with surfing tools will enable the company Web site to present a gist of their company information in multiple languages, thus reaching out to other language markets instantaneously. With this, the business can focus on developing their business strategy in a single language instead of diverting resources to develop several language versions.

An example is the LISA Web site at www.lisa.org, where the site can be presented in English (original) and machine translated into Simplified Chinese, Traditional Chinese and Bahasa Melayu (Malay language).

• Multilingual Communication

Machine Translation has become a very useful tool in global communication. The world speaks more than 500 different languages, and, although an interpreter or a translator can facilitate communication across different languages, this quickly gets very expensive and difficult, especially with time zone differences and quick turnaround time demands.

MT can be effectively used for preliminary communication exchange before human interpreters or translators come into the picture. All types of communication such as email, chat, conference and SMS can be processed through MT. An added dimension of complexity for communication is
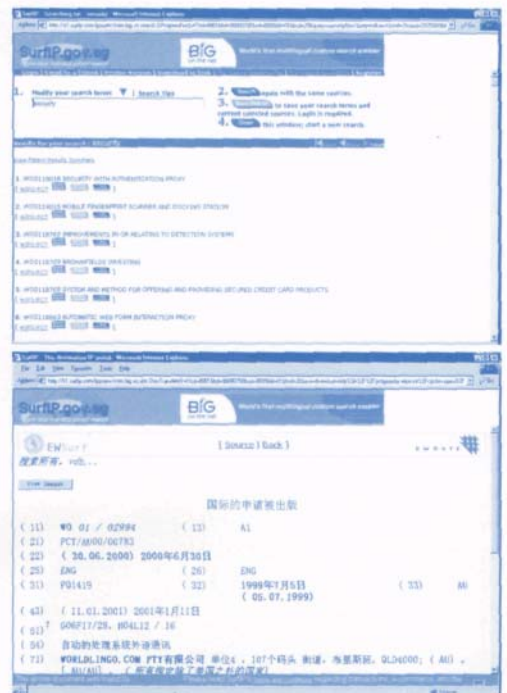

Figure 3

the requirement of a two-way bi-directional translation of source and target languages. Unless an MT system for both directions exists, interactive communication cannot be effectively carried out.

Figure 4 shows how a Web-based email utility in the ePedoman system uses MT technology for users to create emails in one language (either English or Bahasa Melayu) and have them translated to another language. Figure 5 shows how a multilingual chat uses MT technology for
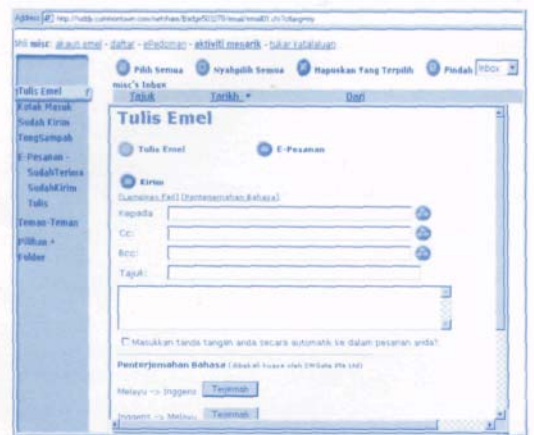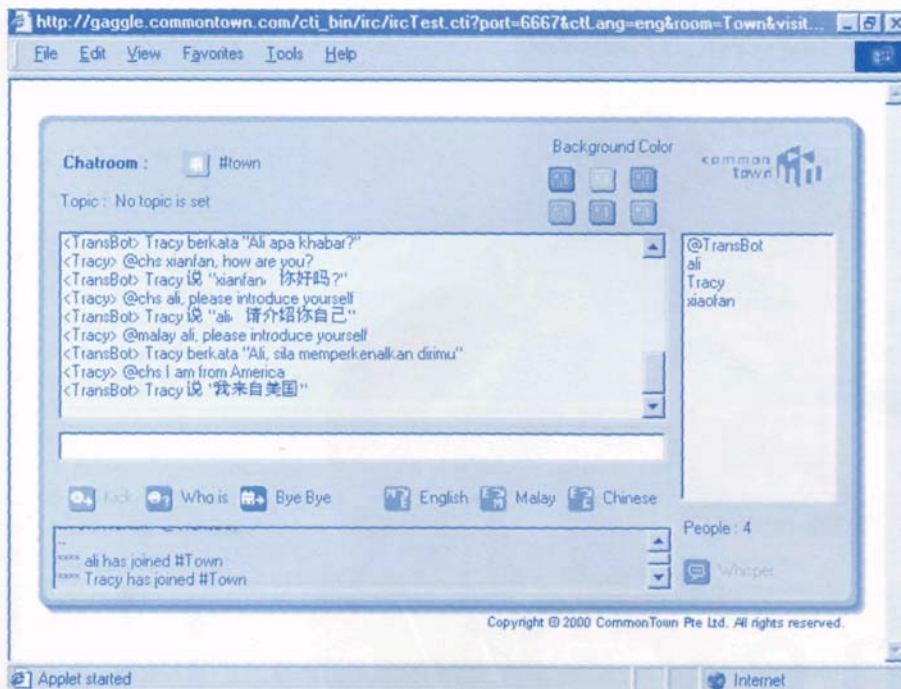

Figure 4

Figure 5

multilingual users to enter a single, truly multilingual chatroom.

MT has many other roles in global communication that are not covered here in this article. This includes empowering multilingual broadcasting of messages, multilingual call centers, as well as real-time speech translation.

Conclusion

MT is an increasingly important technology in the Internet era to further bridge the digital divide of the global market. Its greatest strength lies in its ability to translate real-time in seconds and provide a quick gist of the source information. MT will be extensively used in this global market despite its limitation of not attaining 100% accuracy at this stage. The increasing demand of translating large volume and time-sensitive information at immediate turnaround calls for the strength of this technology.

Wee Li Kwang Angela is CEO of EWGate, a technology venture from Kent Ridge Digital Labs (KRDL). For more information on EWGate, visit www.ewgate.com.