В результате сравнения определено, что метод преобразования Фурье-Меллина предпочтительнее для извлечения инвариантных признаков чем метод геометрических моментов, что дает и лучшее качество распознавания. В свою очередь, для метода преобразования Фурье-Меллина для сокращения размерности входного вектора удалось применить метод покомпонентного анализа (PCA) и было определено, что 8 наибольших собственных значений дают лучшие результаты распознавания чем вектор $|\text{FFT}|$ размерности 64.

Литература:

**Machine translation error analysis**

Краснова Татьяна Ивановна, старший преподаватель;
Ванюшин Иван Сергеевич, студент
Национальный исследовательский Томский политехнический университет

**Krasnova T.I.**
National Research Tomsk Polytechnic University, instructor

**Vanushin I.S.**
National Research Tomsk Polytechnic University, student

**Introduction**

In the age of globalization when the world tends to erase the boundaries for global benefit, understanding of foreign languages gives new perspectives for a world citizen. People of different cultures get closer and distance between countries is swept away with the rise of new communication technologies. The only thing that can disturb this convergence of cultures is language. In multilingual society the knowledge of one or two foreign languages is not enough now. Under such conditions machine translation (MT) goes through a rebirth. Information technologies and the Internet made a tremendous impact on translation. We can call it a digital revolution in translation. Alongside with reliable professional applications there is a rapid proliferation of automated online translation services and translation applications for smartphones.

MT is often criticized for poor quality output that demands manual post-editing to bring it up to high-quality standard. This low-quality translation can be used only to get indication of the content of the original text. Sometimes this ‘indicative translation’ is enough, especially when you don’t care about the details and need only the main idea of the text, for instance it is good enough for the translation of web pages. But most often this problem is seen as MT weakness.

**Historical background**

In our article we will define MT as a process of text translation from one natural language to another, using software. It can’t be seen as a simple substitution of words as it is a very complicated process which main purpose is the realization of high-quality translation of the text in natural language to its equivalent in the translated language.

The concept of MT is quite old. It begins with the ideas of G.W. Leibniz about the possibility of the mechanical translation through philosophically-mathematical interlingua (1646—1716), C. Babbage about the possibility of implementation of translation done by machine (1836—1848) and with the invention of Russian scientist P.P. Smirnov-Troyansky, who offered in 1933 a mechanical translator which automatically selected word equivalents for the units of the input language. On these ideas the theory of machine translation was based. The founder of this theory was W. Weaver. In 1947 for the first time he proved theoretically the fundamental possibility of MT systems creation. The foundation of the theory was the fact that any natural language is a code system and the automated translation process may be limited to the decoding process.

Until the late 1980s, MT was largely dominated by rule-based systems where grammar and syntax rules were combined with cross-language dictionaries. In the 1990s, the shift was to experimenting with sets of parallel texts. In statistical based MT, algorithms analyze large collections of previous translations or parallel corpora to estimate the statistical probabilities of words or phrases in one language ending up in another. A model is then constructed on the basis of these probabilities and used to evaluate new text. By implication, these systems perform best on the types of texts on which they have been trained [1].
Yandex MT approach

Nowadays there are a lot of machine translation systems that can be classified on different grounds. The most popular applications are Google Translate Toolkit, Babylon Translator, PROMPT, Yandex, Systran and so on. They all have different algorithms; let's look, for example, how Yandex translation system works.

The main distinctive feature of this system is that it is statistical. It means that its translation methodology is based not on language rules (the system even doesn't know them) but on statistics. To learn a language, the system compares hundreds of thousands of parallel texts which contain the same information, but in different languages. It may take, for example, large texts from multilingual versions of organizations' websites. Initially, the system finds parallel texts at documents addresses, often these addresses differ only by notes, for example, «en» or «ru» for the English version and «ru» for the Russian one. For every studied text the system builds a list of unique signs. These could be rarely used words, numbers or special symbols found in the text in a certain sequence. When the system gains a sufficient number of signs from texts, it begins to look for parallel texts comparing with their help the characteristics of the new texts and already studied. To meet current translation quality standards, the system should learn the hundreds of millions of phrases in different languages. It requires very large resources: a lot of space on HDDs, lots of RAM and so on. That is why the existing machine translation systems are in such limited number.

In Yandex machine translation system there are three main parts: the translation model, language model, and a decoder. The translation model is a table, in which all words and phrases the system knows in one language lists all possible translations into another language and contains the possibilities of these transfers (for each pair of languages there is their own table). This model is created in three steps: firstly we select parallel documents, then in them — the pairs of sentences, and then a pair of words or phrases [2]. After that the decoder performs a translation. For each sentence of the original text, it finds all transfer options, combining together phrases from the translation models, and sorts them in the descending order of probability. The decoder estimates all variants of the output combinations using the language model. As a result, the decoder selects a sentence with the best combination of probability (in terms of translation model) and frequency of use (in terms of language model).

Current problems in MT

There is no doubt that MT is still imperfect and there are a lot of problems that arise during the translation process. All human translators know translation is not simply a matter of finding the target words that correspond to the words in the source text, and then getting the target grammar right. In fact it involves selecting the correct sense of each individual word, and recognizing the relationship between the words, as expressed by the syntax of the source text [3]. This task is quite difficult for a computer programme.

We will have a closer look at these problems and try to consider them by translating the same phrase in such MT systems as Google Translate, Yandex. Translate and PROMPT.

1. Lexical problems
   Word usage of translators often conflicts with the database of words known by translator.
   Source: Scuba, wetsuit, swimin are necessary for divers.
   Google Translate: Scuba, wetsuit, swimis necessary for divers.
   Yandex. Translate: Scuba, wetsuit, fins are necessary for divers.
   PROMT: Scuba, wetsuit, fins required for divers.

2. Word conjunction and polysemy
   Multi-meaning words are real problem for machine translation for one simple reason: sometimes it is really difficult to choose one or another. People usually use the context of the phrase, but meaning of phrase, which is cut off from text or speech, becomes undefined for translator.
   Source: My bow is more beautiful than your bow!
   Google Translate: My beautiful bow your bow!
   Yandex. Translate: My bow your beautiful bow!
   PROMT: My onions are more beautiful than your onions!

3. Syntactic problems
   Source: Don't be angry with him.
   Google Translate: Do not hold a grudge against him.
   Yandex. Translate: Don't be angry at him.
   PROMT: Don't harbor malice against it.

4. Problems at the level of production and transmission
   Source: Listen, if the stars are lit, it means there is someone who needs it.
   Google Translate: Listen, because if the stars are lit, it means someone needs?
   Yandex. Translate: Listen, if the stars are lit, it means someone need?
   PROMT: Listen, after all if stars light, it means to somebody it is necessary?

These examples demonstrate that MT systems can't translate with a hundred per cent accuracy. Thus the problem of accuracy remains central for MT systems developers.
Conclusion

Machine translation has a long history but is still relatively immature technology. For the past decade researchers and developers have been trying to determine the efficacy of existing MT systems and to find solutions for optimizing these MT systems. The progress in the field of MT depends on systematic evaluation and quality control. Every new system works better than the previous one. There are still certain limitations in applications but MT accuracy increases every year.

References:


Increasing contact center efficiency through automatic dialers

Maksutova Akmaral Toregaliievna, master
International Information Technologies University (Kazakhstan, Almaty)

The article proposes a new approach (algorithm) of quickly and efficiently calling models of customers in organizations which using CRM system for customer service through telephone channels.

The large organizations as Banks, consulting companies use CRM systems for customer service through the telephone channels. And it’s complicated to make the communication effectively without lost, technical breaks, and thus save the time.

There are several kind of drivers and systems offer the solution to make communication automatically, and one of the best decisions is offered by CISCO, as called CISCO Dialer or Cisco Unified Contact Center Enterprise. But having the automatic tool for call is not enough to organize the effective calling when the phone number for calling is not available, when the client has several phone numbers, when the call is broken, or there is need to exclude the phone in a real time and etc.

Automatic dialers increase contact center efficiency because they eliminate the possibility of reaching a wrong number, save time, and make contact center agents more productive. By automatically dialing and screening for busy signals, no answer, and answering machines, dialers ensure that agents do not waste time on non-value-added mechanical and routine tasks. Only when the dialer reaches a live contact will the solution transfer the call to the next available agent.

The Cisco Outbound Option application provides outbound dialing functionality along with the existing inbound capabilities of Cisco Unified Contact Center Enterprise. This application enables the contact center to dial customer contacts and direct contacted customers to agents or IVRs. With the Cisco Outbound Option, you can configure a contact center for automated outbound activities.

Outbound Option features include:

- Unified CCE Compatible Dialer
- Campaign Management
- Unified Contact Center Software Management of Skill Groups
- Outbound Option Dialing Modes
- Outbound Option activity reports
- Callbacks
- Call Progress Analysis (CPA)
- Transfer to IVR
- Sequential Dialing
- Cisco IP Contact Center Agent Re-skilling
- Abandoned and Retry Call Settings
- Campaign Prefix Digits for Dialed Numbers
- Outbound Option Support on Cisco Unified Contact Center Hosted
- Outbound ECC Variables Support in Siebel 7.5.3 and Later

Outbound Option supports the following dialing modes:

- Preview
- Direct Preview
- Progressive
- Predictive

Depending on the Outbound Option campaign settings, a callback can be scheduled as a personal callback or a regular callback.

The transfer to IVR feature provides Outbound Option with another outbound mode. This mode causes the Dialer to transfer every customer call associated with a specific skill group to a service control-based IVR instead of an agent. This feature allows a contact center to run unassisted outbound campaigns using pre-recorded messages in the Cisco Unified IP-IVR and Cisco Unified Customer Voice Portal products.