

From websites to apps, and now from apps to chatbots? leveraging the next interaction channel for the global single market with machine translation

António Branco
University of Lisbon

LIGHTNING TALK ABSTRACT

Abstract:

Triggered by the vision for the next decade, concerning web-based interaction and his company, recently shared by the CEO of a large social network, in the present talk we discuss how machine translation is crucial for this web-based key interaction technology and how much it helps to leverage this technology for the global single market.

1 From apps to chatbots

In the recent annual conference for their developers, the CEO of a very large social network shared his vision on what he believes is the key disruptive trend in ICT for the next ten to twenty years. He shared how this vision is shaping his company's roadmap for the next decade and why the company will be pursuing, fostering and benefiting from that identified game changing trend.

In a nutshell, that vision can be rendered as follows: with the advent of personal computers, every company needed to have a presence in the digital world by means of a website; later on with the advent of smartphones, this evolved into every company putting much effort in using apps as their privileged channel to reach out the digital world; the next disruptive step is for the companies to focus on intelligent chatbots to interact with their customers.

2 A hidden barrier

Much before chatbots have been put under the spotlight, many companies have already been doing their business by providing online helpdesk services through a chat channel. If you were the CEO of one of these companies, you would know that a most important hidden barrier in this business is the barrier of multilingualism.

When running your service, a most important and invaluable asset is the database that gathers the accumulated questions from the customers and respective answers along the years. However, when moving to another market where the customers use a different natural language, the value of that asset almost vanishes to zero before your eyes: the wealth of knowledge accumulated with the customers in the previous market cannot be reused, not at least immediately, to support the helpdesk service in another market running on another natural language.

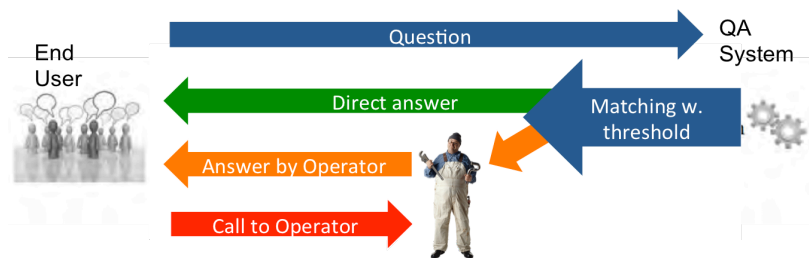
3 A real usage scenario

The scenario just described is the real usage scenario with which an on going European research project on quality machine translation is seeking to cope with, the QTLearn project (qtlearn.eu). This scenario is made available to the project by an SME that is a member of the project consortium, the Higher Functions company.

The process of providing support to end-users involves remote written interaction via chat channels through a call centre. The support from the call centre operators, and a fortiori this process of problem solving, is made efficient by a Question Answering (QA) application that supports the operators in preparing the replies to clients.

Each query for help is matched against a memory of previous QAs, and drawing from that repository, a list of possible replies is displayed, ranked by relevance according to internal heuristics of this support system. If the top reply scores over a threshold, it gets automatically returned to the client. If the reply does not score over the threshold, the operator is presented with the list of possible answers delivered by the system and he can either pick the most appropriate reply or modify one of the replies or write a completely new reply. In both the last two cases, the new reply is used to further feed the QA memory.

This workflow is illustrated in the following diagram:

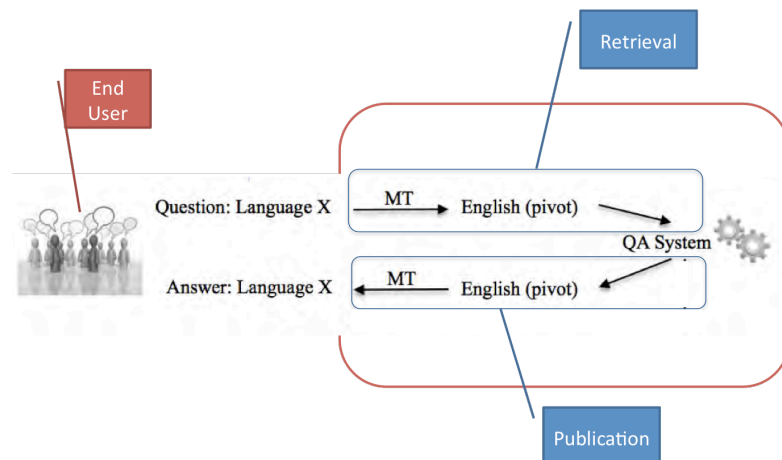


4 Machine translation put to work

A major goal of the project is to assess if and how much can machine translation help in porting the online service to another market with another language.

In order to cope with multilingual markets, to minimize costs, and to optimize both human and technological resources, a crucial component to introduce in the system is high quality machine translation. For instance, a user entering a query in Bulgarian will benefit from the fact that, if similar query had been entered by a user in Dutch and associated to a correct reply in the internal database, then he can benefit from that previous interaction, even though it was performed in a different language. The Dutch query was translated into the pivot language English, and associated with an reply, also in the pivot language. The Bulgarian query can be automatically obtained from there, by translating it with quality MT.

The gist of this workflow is depicted in the diagram below:



When a request for help gets translated, the search for its reply benefits from the entire Q&A memory of previous interactions, input in all languages. Importantly from the business point of view, an economy of scale can thus be obtained in terms both of the technological infrastructures needed to be deployed and of the size of the team of human operators to recruit.

4 How much can machine translation help

In a monolingual scenario, the indicator for the cost of *not having* a QA database in support of the online helpdesk service is the number of phone calls to the human operators that *cannot be spared* by resorting to (at least some) answers being delivered automatically by the system. Labour costs with human operators is a most substantial part of the operation costs and minimizing them is the key factor here.

Likewise, in a multilingual scenario, when porting you business to market with another natural language the indicator of the cost of *not having* a machine translation integrated in the QA system is the number of phone calls to the human operators that *cannot be spared* in the day 1 of the operation in the new market.

In order to asses how much machine translation can thus help in porting chat systems to multilingual markets, the project set up an evaluation methodology to

estimate the number of phone calls to human operators spared once machine translation is embedded in the QA system.

According to the results of that assessment exercise, with the support of machine translation the overall chance of dispensing a human operator, and thus of dispensing with language specific costs of porting the system to a new market with a new language at day 1 of its operation, is on average up to 20%.

5 Leveraging the next revolution

These results are important indicators for different stakeholders. For companies in the business of online (global) helpdesks via chat channels, they give a first objective indication of how much can be saved by resorting to state of the art machine translation technology.

For science policy decision makers, they show once again and more than ever how much important and impactful for the multilingual digital single market is to take decisive measures that support the progress of research and innovation in the area of machine translation and multilingual technologies, thus supporting the emergence of new cross-border businesses and companies.

6 Final remarks

The present submission is a short abstract of a lightning talk presented at META-FORUM 2016. Complete information on the methodology to estimate probability of phone calls to human operators, etc are provided at the QTLeap project website qt Leap.eu.

Acknowledgments

This research was supported by the grant of the European Commission FP7-ICT-2013-10-610516, to the project QTLeap.

This is a collaborative project and the present paper is inspired on its results, which would not have been possible without the contribution of the project participants, including the heads of the different teams in the project consortium, namely Hans Uszkoreit (DFKI-German Research Centre on Artificial Intelligence), Jan Hajič (Charles University in Prague), Kiril Simov (Bulgarian Academy of Sciences), Eneko Agirre (University of the Basque Country), Markus Egg (Humboldt University in Berlin), Gertjan van Noord (University of Groningen) and Filipe Barrancos (Higher Functions), and including also, in what concerns the results on the estimation of the contribution of machine translation in particular, Rosa del Gaudio and Aljoscha Burchardt together with several other colleagues from all teams that contributed for collecting this estimation.