

LTC-Communicator: A Web-based Multilingual Help Desk

NIGEL GOFFE
Language Technology Centre Ltd.

Software vendors and other service organisations operating in increasingly international markets face a twofold challenge: first, their products have to be localised to meet the specific linguistic and cultural requirements of each target country; and then there is the problem of providing quality support to diverse customers, particularly where the end-users do not speak the same language as the vendor's own support group.

Localisation (producing new versions of screens, help text and other documentation, as well as promotional and marketing material), while certainly not cheap, is at least relatively well defined and understood, with many language service companies providing expertise and tools. The problem of user support in multiple languages is much more complex, with few off-the-shelf solutions available.

The LTC-Communicator, a software product from the Language Technology Centre Ltd., now offers an innovative and cost-effective response to this growing need.

INTRODUCTION

Until now, there were basically three possible strategies for multilingual support:

- Maintaining a separate support centre in each target country; this is very expensive, and only really feasible for a large organisation with an established international presence;
- Running a centralised help desk employing multilingual support staff; also potentially expensive (people with the right technical and linguistic qualifications are not easy to find, and user numbers

- may not justify experts in every language to be covered); a variant of this is the distributed 'virtual help desk' with remote support staff connected to a central database;
- Using telephone interpreting services to offer multilingual support from a central location; expensive on a case-by-case basis, but perhaps justified where the volume of multilingual support is small and the demand sporadic.

The LTC-Communicator now offers a fourth possibility:

- Using translation software to incorporate multilingual capability into the help desk software itself, and integrating with web-based workflow and knowledge base facilities - reducing the dependence on human experts to manage communication.

BUSINESS CASE

LTC-Communicator allows a software vendor (or any service organisation) to run a centralised support centre without the expense of training and equipping local support desks in different countries, or employing multilingual staff; the end-user and the support engineer do not need to share a common language.

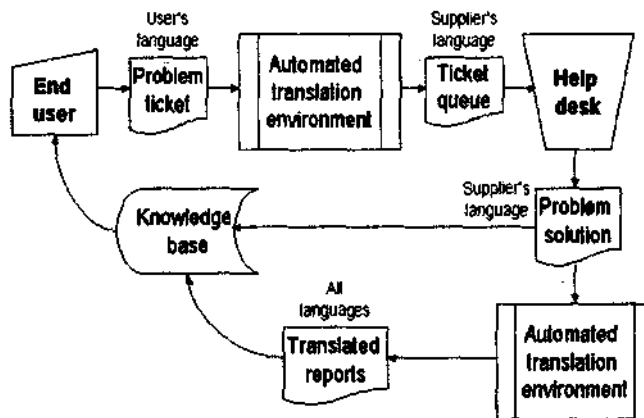
For example, for an English-speaking help desk supporting German users:

- The user enters a trouble ticket in German, via the software company's support portal;
- The trouble ticket is routed through the LTC-Communicator translation components;
- The request is displayed to the support engineer in English;
- The engineer prepares the solution, also in English, to be automatically routed back through the translation environment;
- The user can then view the solution (or status information) in German.

An attraction of this solution is that it can integrate with any existing web-based help desk application that is capable of communicating via XML and can be configured to provide flexible workflow; incoming messages in the company's 'native' language are passed directly into the normal help desk workflow, while 'foreign' messages are routed through the LTC-Communicator automated translation environment. Similarly, response messages from the help desk may be directed straight to the requesting user or translated first, as applicable.

LTC-Communicator can also be integrated with established knowledge base functionality, supporting user 'self-help' with multilingual query facilities against existing documentation and FAQs, and continuously enhancing the contents of the knowledge base with new queries and solutions.

Fig 1: *The overall workflow*



TECHNICAL ARCHITECTURE

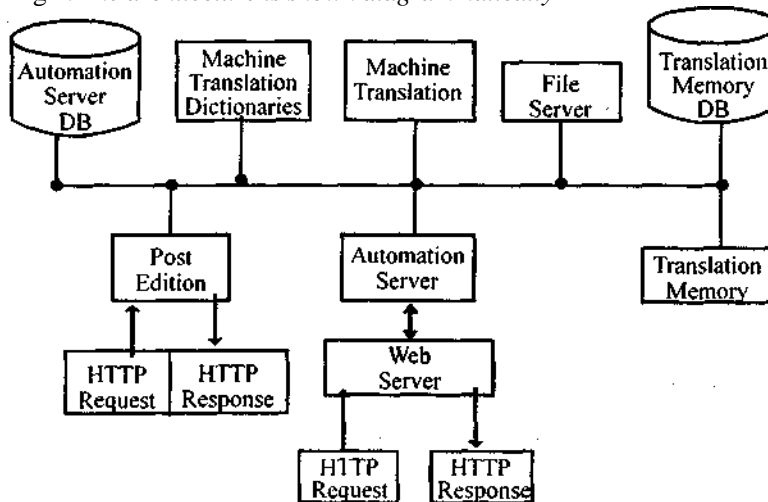
At the technical level, LTC-Communicator itself comprises three components:

- A translation memory populated with the localised software, bilingual versions of documentation, online help, and frequently asked questions;
- An interface to a machine translation program (e.g. SYSTRAN) including a custom dictionary holding key terms relevant to a given customer;
- An automation server responsible for workflow management, which routes messages through the translation modules and an optional post-editing service.

Together, these function as a 'black box', with input and output in XML format; the only real changes that have to be made to an existing help desk application to integrate with LTC-Communicator are:

- To generate (and read) XML messages in the specified format;
- To tag the messages with a language identifier to indicate whether they need translating, and if so, to route them correctly through the language software.

Fig 2: The architecture is shown diagrammatically



In principle, the system can handle any language combination, limited only by the capabilities of the machine translation software available; between them, the existing MT systems cover all the major European languages, and increasing numbers of non-European languages also.

Once the system is set up, it is relatively easy to add further languages, with the automation server ensuring that each incoming message is directed to the appropriate MT module within the translation environment. A single help desk can then support users in many different countries at little additional cost.

USE OF MACHINE TRANSLATION

Machine translation is particularly well-suited to help desk requirements, where:

- Output needs to be intelligible, and not generally of 'publishable' quality;

- The subject area is usually quite restricted (e.g. a range of software products), and custom dictionaries delivered with the MT software can be tailored to the specific 'domain';
- Existing material can be re-used; databases populated with bilingual documents, localised software, existing FAQs etc. are used by the translation memory;
- Similarly, each 'new' solution is potentially re-usable; it can be added to existing FAQs, or to a multilingual knowledge base, enabling users to resolve some queries without calling the help desk at all.

Moreover, the LTC-Communicator architecture allows results to be optimised in a number of ways:

- Input can be structured by providing pick lists for key items such as application/version, type of fault, severity etc., producing more matches against the translation memory, and minimising the amount of free text to be translated;
- Style guidelines can be applied to produce more consistent input, enforcing standard use of terminology and avoiding complex or ambiguous constructions that the MT software may not resolve correctly;
- Spell checkers can also be used to improve input quality, further increasing potential 'hit rate' against the TM;
- Where output of 'publishable' quality is required (e.g. to populate a multilingual knowledge base, or to be incorporated in formal documentation), a post-editing option is available; post-editing is also recommended when the system is first installed, to help 'tune' the TM and MT dictionaries;
- Post-edited output is automatically fed back into the TM.

BACKGROUND

LTC-Communicator grew out of a research project ('IMPACT') funded by the European Union as part of the 'Information Society' programme, and managed by LTC in partnership with CAS Software in Karlsruhe, Germany. This culminated in a pilot phase where the multilingual capability was integrated with the existing (German-speaking) CAS help desk, supporting users of the CAS groupware product 'genesisWorld'; system integrators in user organisations in Italy and Greece provided the input (in

English). The LTC-Communicator component itself was run on LTC's own servers in the UK, and LTC also provided post-editing services where required.

The results from machine translation were evaluated by the CAS support engineers (looking at incoming trouble tickets translated into German), and by end-users (dealing with responses translated into English); the emphasis was on intelligibility rather than strict grammatical correctness. LTC also reviewed the linguistic output, and used the results to extend the Systran custom dictionaries with more CAS-specific terminology and usage; this produced a significant improvement in quality over the course of the pilot phase, to the point where extensive post-editing was no longer required for the MT output to be usable.

Having established the effectiveness of the automated translation approach, LTC-Communicator is now being actively marketed by CAS as an extension to their new CAS-Helpdesk product, and by LTC as an add-on to other helpdesk products where a supplier or system integrator needs to support a multilingual customer base.

IMPLEMENTATION

Initially, LTC-Communicator will be offered as a hosted service only, with LTC running the machine translation environment, as in the IMPACT pilot. There is a fixed annual 'subscription' dependent on the number of language combinations required, and a small handling charge per message processed. Again, post-editing is offered as an option, and is recommended in the early stages - at least until domain-specific custom dictionaries have been built up.

For a customer handling 500 multilingual transactions per year, all in the same language combination, the projected annual cost without post-editing is £8,500; this assumes an average message length of 50-100 words (with each transaction comprising two messages -the trouble ticket itself, and the solution). For 1500 such transactions per year, the projected annual cost rises to £13,500. Each additional language pair adds £600 to the annual charge (to cover license costs); post-editing costs from £2-3 per message, again depending on the number of words.

The more transactions processed, the lower the average cost per transaction - and compared to the cost of multilingual human resources or distributed helpdesk facilities, the annual costs are modest indeed.

In principle, customers could also run the system in-house on their own hardware; however, the high initial costs of MT software licenses make this unattractive to any but the largest organisations processing 3000

or more transactions (message pairs) per year. Even then, the pricing models offered by MT vendors mean that relative savings only appear after 4-5 years, and customers also have the potential headache of operating and supporting specialist software outside their normal area of expertise.

CONCLUSION

The tremendous growth of the Internet means that more and more applications are becoming truly global in reach, and companies have to operate in an increasingly international - and multilingual - environment. The demand for instant translation (on demand) threatens to outstrip the capacity of human translators, who are an expensive resource in any case.

Machine translation is part of the answer, but it does have limitations when used on its own. LTC-Communicator offers an innovative combination of technologies to overcome these limitations while also minimising the reliance on human translators.

Although the focus so far has been on help desk systems, the LTC-Communicator would fit into any workflow application with a need to support users in more than one language; it could also be used to add multilingual capability to technical search engines and knowledge base products, both of which play an increasingly central role in the corporate IT environment.

ABOUT LTC

The Language Technology Centre is a limited company based in the UK, specialising in building multilingual websites, software localisation, consultancy in language technology, technical translation and software development. Clients include telecommunications companies, software developers, mechanical engineering businesses and European institutions. You can visit us at: www.langtech.co.uk

NIGEL GOFFE
LANGUAGE TECHNOLOGY CENTRE LTD
5-7 KINGSTON HILL
KINGSTON UPON THAMES
SURREY Kt2 7pw
EMAIL: NGoffe@langtech.co.uk