



MoBiMouse, the world's first "no-click" dictionary program

by Bob Clark

Budapest-based MorphoLogic has recently released a new version of MoBiMouse, a.k.a. the translating mouse, which offers translations of words and expressions found anywhere on the screen, without clicking the mouse. The user only has to move the mouse pointer over the word in question and the translation is displayed in a tool tip-like bubble within a second. If the mouse is moved again, the translation disappears promptly, so work is not disturbed by another program requiring a whole window with all its bells and whistles.

The purpose of developing MoBiMouse was to make foreign text understanding as quick and easy as possible. The available translator and dictionary programs have been designed mainly for people who want to translate long texts. MorphoLogic has developed other tools for translators, but for those who need instant help while browsing the Web or reading e-mail received from abroad, MoBiMouse is a perfect solution.

MoBiMouse works on the entire Windows desktop (that is, Windows 95/98, Windows NT 4.0, and Windows 2000), by reading and processing words and expressions in the area where the mouse pointer is. The program recognises any piece of text written with any font installed in Windows. MoBiMouse is especially useful when browsing the Web where users often encounter pages written in foreign languages, whether they are German, Spanish, French, English, Hungarian or even Japanese. MoBiMouse helps the user read these texts quickly, without interrupting the current activity. MoBiMouse is able to recognise text that might be obscured by a colourful or patterned background. So, even if our eyes are unable to read pieces of text because of the background or a texture, MoBiMouse will still pick it up. Its processing speed (1000 characters/second) also makes it an important tool for Internet users.

Innovations in MoBiMouse

MoBiMouse has several innovative features that make it special among dictionary programs of its kind. MoBiMouse continuously monitors the motion of the mouse pointer and tries to look for text in the neighbourhood. Its mechanism appears to behave similar to an OCR program, but actually it is quite different. MoBiMouse actually performs linguistic analysis of the recognised pieces of text. Using a morphological analyser also developed by

MorphoLogic, called Humor, it determines the lexical stems (basic forms) of word in the recognised expression. It is the basic form that the dictionary module looks up in the dictionary. Meanings are then displayed in a bubble-shaped pop-up window on the screen.

As sketched above, the most innovative feature of MoBiMouse is that the user can use it without any interaction. It all looks very simple: users do not need to press a button or click the mouse. They just leave the mouse pointer on the word or expression they wish to understand. MoBiMouse then presents a like-like bubble with the translation of the text in the language they understand. Using tool tips to display additional information is not new, but this particular like does not belong to any application window on the screen. It is very interesting to note, however, that any sort of information can be shown in the bubble, not only dictionary look-up results.

Though MoBiMouse is very easy to use, the technology in the background is rather sophisticated. The system is a combination of three main modules: a shape identifier, a linguistic stemmer and a dictionary look-up engine. These modules incorporate numerous innovative aspects. One is the technology, mouScan, that MoBiMouse uses to select a monochrome shape from a set of previously learned sets of shapes. This technology is based neither on OCR, nor on system call manipulation. The technology has nothing to do with traditional character recognition methods. It is, however, more general, and suits character identification very well. It is particularly useful when the set of shapes to identify is rather large, e.g. in the case of Chinese and Japanese characters. Thanks to the fast identification algorithm, more than one single pattern is identified at a time. This makes it easier for a user to identify a full line or a full sentence at once, even when the traditional input devices are totally impossible to use. For instance, the user of a European version of Windows is unlikely to be able to request the meaning of an ideographical character, a so-called kanji, on a web page written with Japanese characters because they may not have the means of entering such characters. When applied to a set of 6,000 different shapes (dictionary of kanjis used in Japanese), MoBiMouse identified 1,000 kanjis per second on a standard 200 MHz Pentium-class machine. In this case, MoBiMouse is used as a special input method. The underlying mouScan technology is easily applied

to applications other than dictionaries as well, for example, to display an entirely different sort of information (encyclopaedia, large worksheets etc.).

Another interesting aspect is the ability to combine the shape identifier with any external module. MoBiMouse is based on a well-defined API between the output of the identifier and the application that processes the recognised pieces of text. Using this API, any application can easily be linked to the mouScan module, not only dictionary look-up engines.

Unlike many similar systems on the market, MoBiMouse does not rely on word lists, but on an algorithmic solution that makes it possible to handle millions or billions of word forms of highly inflectional languages, such as German, Hungarian, etc.

MoBiMouse can automatically connect the user to the developer/lexicographer team. While MoBiMouse is in use, unknown words are automatically and continuously collected into a log file. From time to time, this log is automatically sent to the developer team, unless the user refused to permit the system to do so. This is an option to be selected during the installation process.

MorphoLogic Ltd. was established in 1991 by three former researchers of academic laboratories and still owned by them. MorphoLogic has the distinction of being the only organisation in Hungary performing R&D solely in the field of Human Language Technologies. Worldwide leaders of software technology, such as Microsoft, Lotus, IBM, Xerox, Inso, etc., have licensed MorphoLogic's language technology applications (spelling checkers, grammar checkers, hyphenators, thesauri and dictionaries) for various languages for their applications to be localised for the CEE market. MorphoLogic was the first company to develop a single-engine-based language technology tool for the most important CEE languages, Hungarian, Polish, Czech and Romanian. Recent MorphoLogic technology covers not only Central Europe, but also most of the languages of the EU and even some Far-East languages. The close ties that MorphoLogic continues to maintain with academic labs in Budapest have resulted in a number of very profitable language technology products, and the sale of these products not only provides funding for the company's profit-making activities, but its basic research as well.

Over the past several years, MorphoLogic's multilingual applications have come into the limelight. The largest client-server version of the MoBiDic translation support system is used by the Hungarian Government's network interconnecting all the ministries and the Prime Minister's Office. Another important on-going application of MorphoLogic's translation support and terminology

management tools is the project for the Translation of EU Legislation Texts into Hungarian (1998-2000). More than a hundred professional translators will translate 40,000 pages.

MoBiMouse belongs to the MoBiDic product range.

MoBiDic (MorphoLogic Bilingual Dictionaries) is an intelligent translation tool accessing an unlimited number of dictionaries simultaneously. Look-up for a word or a phrase in MoBiDic starts by finding the stem (the basic lexical form of the word) in question, and this information is sent to the dictionary look-up module. As a client/server application, MoBiDic supports networked installations where the dictionaries reside on a central server that allows for the use of shared terminology databases. MoBiDic servers run on Windows NT, Linux, Solaris and Novell operating systems. MoBiMouse works under Windows 9x/NT/2000, both as stand-alone product and a special client to the MoBiDic translation support system

In 1999, MoBiMouse won the European IST-Prize, which is given every year by the European Council of Applied Sciences and Engineering to novel products with high information technology content and evident market potential.

MoBiMouse is now commercially available with Hungarian/English and English/Hungarian dictionaries (the price is €30), but dictionaries are being prepared for other CEE and EU language pairs, as well. The development of MoBiMouse for Far-East languages has also begun. MoBiMouse has been (and is to be) sold as an end-user product, as OEM software for mouse, and as an easy-to-sell tool for publishers who need advanced technology. Companies who would like to localise their products, while keeping costs down, have got a fast, new and smart solution for localisation. MoBiMouse can be used both along with and instead of localisation. If the dictionary contains all the words and expressions used in the resources of a product in question, MoBiMouse offers translation equivalents of the menu points and help files without being "inside" the application itself. ■

The MoBiDic product line contains:

MoBiDic stand-alone versions,

MoBiDic client/server versions (for Windows NT, Linux, Solaris and Novell operating systems),

MoBiWeb intranet/internet versions (to be used with any internet browser),

MoBiMouse (both as an intelligent client to the MoBiDic server and as stand-alone versions for Windows 95/98/NT/2000).

Detailed technical information may be found on an on-line HTML-demo at the MoBiMouse home page (www.mobimouse.com) and on the home page of the developer company, MorphoLogic (www.morphologic.hu).

MorphoLogic

Késmárki u. 8, Budapest 1118, Hungary

Tel/fax: +36 (1) 3614721, +36 (1) 3614722

E-mail: info@morphologic.hu