CASE AND KEYBOARD INDEPENDENT WORD COMPARE

D. A. Hamilton, S. S. Hobbs and C. R. Nunn

The same character keyed in different keyboards may be represented by a different code. In fact, the conversion of a character from one keyboard to another is not only not unique, it is also not necessarily one-to-one. For example, ö may be keyed as one character or as three characters -- " o RBS ¨ ", where RBS is a required backspace. In addition, the conversion must be case independent. The conversion of a character to upper or lower case is language dependent. For example, in National French, it is inappropriate to convert lower case accented characters to upper case with accent. Thus, "é" converts to "E".

The following solution to the above problem is disclosed:

1) Define word delimiter codes which include punctuation. Thus, " SP w o r d SP " and " (w o r d ." will compare equal.
2) Convert each character to a unique code (font code). Thus, "é" is defined to be "F1" whether it is keyed as a single stroke or as a three-stroke construct.
3) Convert each character to upper case (within this unique code set). This process is table-driven so that a language appropriate table can be chosen when the program is assembled.
4) Compare the words.

This solution offers the advantages of recognizing case changes for characters with diacritics and handling language dependent problems.