Using MicroSPELL

version 2.0 - released June 27, 1992

Spelling Checker and Corrector
written by Daniel Lawrence

MicroSPELL 2.0 is a compact, efficient, fast spelling checker for use both as a stand alone program, and with MicroEMACS v3.11 and above on many different computers. Currently supported computers include IBM-PCs under MSDOS and OS/2, ATARI ST, Amigas, and UNIX machines of various types. MicroSPELL is supplied both as IBM-PC executables, and as C source code suitable for compilation by most C compilers.

MicroSPELL contains tools for marking misspellings, scanning for replacements, single word suggestions, and user and main dictionary maintenance. MicroSPELL is normally used in conjunction with MicroEMACS 3.11 in order to scan text files, stepping through misspellings and making suggestions and corrections.

Spell Correction

The SPELL program is responsible for scanning files and marking suspect words. The SCAN.CMD macro is used by MicroEMACS to then step through suspect words, making suggestions, and allowing corrections. This combination can be called in two ways. If you have a text file in MicroEMACS, call up the Word processing page (<F8>-W) and use the SHIFT-<F4> key to activate a spell scan on the current buffer. Outside of MicroEMACS, use the SPELL command as follows:

spell {<options>} <file> {<file>.........<file>}

The currently available options are:

-d	debugging mode
-e	use MicroEMACS to scan errors
-u <fname>	use a user word list
-w	output a word list instead of a location list

Example

Let’s assume that you want to spell check a text file called text.txt which is on the **B** drive of your machine. To invoke spell type:

spell –e b:text.txt

Be patient. It will take SPELL a short while to complete its setup procedure. If would like to know more about what SPELL is doing at this point, you can set the debugging switch (**-d**} when you first invoke the program. SPELL will then give you detailed information about what it is doing.

When SPELL has finished checking your file against its internal dictionary, it will give you a brief report of how many suspect words it found, and then, since we used the **–e** switch, it will invoke MicroEMACS in order to scan through the suspect words. You will see the screen divided into three windows. The top window is the SPELL functions and status window. The middle window will display suggested replacements during the spell scan. The bottom most window will contain a portion of the document being scanned, so that you can see the suspect words in context.

There are a number of elements in the uppermost window that you need to be aware of. On the top line, the “File” field lists which file the scan is currently on. This is useful when many different files are spell checked together in one run (many file names can be put on the command line at once). The first number is the current file, and the second number is the total number of files.

The second field on the top line is the “Word” filed. This describes the current suspect word and the total number of suspect words in the current file. This will give you an idea of how far throughout the file you have progressed during the scan.

One at a time, each suspect word in the file is presented to you for you to decide on its disposition. The suspect word is displayed, and the question “Option:” appears in the Dialog Window. The options you have include:

(S)kip	Skip this word, leaving it alone, and go on to the next.
(I)gnore	Ignore this word and all other occurences of this word during this scan.
(C)hange	Change this word. You will be asked to type a replacement word. This word will replace the original and MicroSPELL will remember the replacement.
(D)efault replace	If you have already replaced an identical occurrence of this suspect word, MicroSPELL will present the last replacement as a "Default Replacement". Picking this option automatically replaces the suspect word with this default word.
(G)lobal change	When you first encounter a suspect word, if you wish it to be replaced in the entire document, use this option. You will be prompted for the replacement word.
(A)dd word	To a user dictionary. This word will then be remembered for the remainder of the current scan, and also for all future scans when that user dictionary is invoked.
(U)pper case	Add a word to a user dictionary in its current case. This is good for names, proper nouns and acronyms.
(Q)uit	Finish the current scan, saving the file with the already completed changes. MicroSPELL will ask if you want to save the results of the scan, or discard them.

Also, in the center window may be a list of possible replacement words. An arrow to the right of one of these words indicated which word is currently selected. The cursor keys can be used to move the arrow through the list of words. The <RETURN> key causes the suspect word to be replaced with the currently selected replacement word.

Using the Mouse

During the spell check run, the mouse can be clicked on any option to use that option. Clicking on one of the replacement words causes that word to replace the suspect word.

Merging a user word list into the Main Dictionary

While normally you would probably keep your user word lists separate from the main dictionary, if you discover a common english word or words which appear missing from the main dictionary, the DMERGE program is used.

dmerge <output dictionary> <user file> {<user file>..}

One or more user word lists will be merged with the currently active dictionary resulting in a new word list containing both the main dictionary and the given user word lists. Then you would go on to use . . .

Creating New Dictionaries

CDICT is the program that takes an alphabetized file of words, one to a line, and creates a compressed dictionary which MicroSPELL can use.

CDICT <text dictionary> <compressed dictionary>

Support:

If you wish to register this program, and be entitled to phone support, please send $25 to the address given below. If you do not yet have MicroSPELL 2.0, specify which disk format you would like it sent on. Currently I can make IBM-PC 5 1/4“ and 3 1/2”. Other formats can be made available by special arrangement.

If you make what you think are changes that are useful to many, send me the updates, and as time permits, I will incorporate the ones I understand, and agree with into the master sources.

USmail: Daniel Lawrence
617 New York St
Lafayette, IN 47901

UUCP: dynamo.ecn.purdue.edu!mdbs!dan
ARPA: mdbs!dan@dynamo.ecn.purdue.edu
FIDO: The Programmer’s Room 1:201/10
(317) 742-5533
ATT: (317) 742-5153