**Installation**

Unzip the repository to a folder of your chosing. Don’t use ‘program files’ as the program needs to update a history file in the Data folder. The executable is words.exe. The Data folder contains configuration and history files along with various word lists. The Words folder contains the source code which can be compiled using Visual Studio Express (2017, 2019 or whatever).

**Notes on Use**

Fuzzy word matching is based on the concept of Shannon edit distance which is the number of mutations that are required to change one sequence into another.

A mutation is a change of letter, a deletion or an insertion. So, with one insertion, 'cart' would match 'carat', 'caret' ... 'scart', 't-cart' etc, with 1 deletion - 'art', 'car' and 'cat' and with 1 substitution - 'Bart','Cara', 'cant',...,'wart'.

The same process can be applied to anagrams - you supply a string of letters and tick the anagram box. Now, all anagrams that have the required number of mutations will be listed. So for 'cart' and with 1 deletion you would get 'act','art','car' etc, for insertion - 'actor', 'artic','bract' etc and substitution - 'acer', 'acre'...'tzar', 'wart'.

It is possible to combine mutations so if the number required is, say 2, then it is possible to find all words matching any combination of 2 mutation types.

Any input letter can be protected from mutation by enclosing in square brackets. Substitution for a letter can be limited by including more than one letter between the []s.

It is also possible to select all mutations up to the input number which makes puzzles like the Times Polygon very easy! This can be used to do things like finding the longest word in the current dictionary that has all different letters (dermatoglyphics in UKACD17.txt) - just enter the alphabet, check anagram, deletion and select do all mutations. Enter a number such as 15 and run the search. This would find all words consisting of 11 or more different letters

The list generated can be filtered using a regular expression eg ^c[ao][¬aeiou] would match all words beginning with ca or co followed by any letter other than a vowel and ^x.\*m$ would match words starting with x and ending in m. There are also tick boxes for ^ and $ - if these are ticked then the associated character will be added by default to the front or back of the pattern as appropriate so there would be no need to type them in.

If there are no input letters then the regular expression will be applied to the whole word list and any other settings are ignored. For example, entering the pattern .b..e with the ^ and & boxes tick would match abase,abate,…,obese. If the ^ or $ is unticked then it would list all words ending or beginning with the pattern respectively. If both are unticked then all words containing the pattern anywhere would be listed.

The ‘Use diacritics’ check box indicates whether accents on characters are to be ignored. Changing the state of this box will result in the current dictionary being re-loaded.

There is a multi word feature for finding mutiple words formed from the supplied letters. There is no filtering by pattern for this. Min length is the shortest length word to be used. The digram filter is a file containing valid 1- or 2- letter words which are valid for use with this feature – otherwise any junk, such as abbreviations, will be deemed usable. There are filters set up for various languages although they may not be complete or contain entries that aren’t words. They can be edited to suit. Mutation settings are ignored for multi word searches. This type of search can take a long time to perform and can generate a large number of entries. Press cancel to abort the operation and list everything found up to that point.

Any text file containing a list of words can be used to match against. The program will work with French, Spanish, Italian and German lists as well with an option to ignore accented letters. The lists should be saved as ansi non-unicode text files.

You can edit a file containing a list of URLs to add your favourite dictionary sites. Double clicking on a word in the output list will display the web page for the word using the currently selected URL. Selecting a different URL will query it for the current word.

The Data folder contains the following :-

Defaults.txt - default settings.

lastsession.txt - history file.

urls - list of urls that can be used to look up words. This can be edited to add or remove entries.

The lists folder contains the dictionary files along with a folder of digram filters for various languages. The digram filters are used with the multi word feature to limit 1- or 2-letter words to those contained in the current list.