CALIFORNIA STATE UNIVERSITY, LOS ANGELES
PROGFEST 2013

Problem 1

THE RHYMING SOUNDEX

A group of songwriters are trying to make a new song and are having a hard time finding rhyming words at the end of each line. They need your help in classifying words that are similar in how they sound in the last syllable.

There exists such a phonetic algorithm for indexing words by sound called the Soundex algorithm. However, this algorithm normally results in a code that is based on the first syllable. What you need is to modify the algorithm to work backwards based on the last syllable.

In order to adapt the Soundex algorithm to your purpose, the modified algorithm is as follows:

1. If a word ends in a vowel (a,e,i,o,u) or semi-vowel (h,w,y), and an 'h' at the end.
2. Reverse the letters of the word.
3. Retain the first letter of the word and replace all occurrences of consonants (not a,e,i,o,u,h,w,y) with a dash '-'. If there are two or more adjacent consonants, they are treated as one letter and replace with just one dash.
4. After the first letter, remaining letters that are separated by dashes are treated as vowel groups. If a vowel group has more than two letters, drop the excess letters so that a vowel group may have either one or two letters.
5. Replace the vowel groups after the first letter with digits as follows:
    - a, aa, ea, ha, ia, ya, ye, yh, ei, yu □ 1
    - e, ae, ee, he, ie, hi, y □ 2
    - i, ai, hi, ii, oi, ui, yi □ 3
    - oa, ua, wa, oe, o, ao, eo, ho, io, wo, yo, iu □ 4
    - ue, we, wi, oo, uo, u, au, eu, hu, ou, uu □ 5
    - all others □ 6
6. Drop all dashes '-' so that the first letter and encoded digits remain.
7. If there is only one encoded digit, add a 0 at the end so that the final sounex code has 3 characters.

Example:
glacier □ riecalg □ rie-a- □ r2-1- □ r21 □ r21
filter □ retlif □ re-i- □ r2-3- □ r23
laser □ resal □ re-a- □ r2-1- □ r21 □ r21
race □ raceh □ hecar □ he-a- □ h2-1- □ h21 □ h21
queue □ queueh □ heueuq □ heueu- □ heu- □ h5 □ h50 □ h50
why □ whyh □ hyhw □ hyh □ h1 □ h10 □ h10

Your program must take in words and convert them into rhyming soundex codes and create list of rhyming words based on the first two characters of the code.
Input will consists of words separated by a space and output will be sorted lines consisting of the soundex code and a list of words (ordered alphabetically), separated by commas.

Sample Input
glacier action certification race laser filter space why queue lie surface

Sample Output
h1:lie,why
h5:queue
n1:action,certification
r2:filter,glacier,laser
y2:race,space