```
 1 # CSE250A HW1
 2 # Jiaxu Zhu
 3 import string
 4
 5 # Function to derive beat guess and its probability
 6 def bestGuess(wordNum, words, wordCount, correctGuess, wrongGuess):
 7     tag = []
 8     guess = [0] * 26
 9     guessChar = string.uppercase
10     sumCount = 0
11     # find words whose P(W|E) > 0
12     for i in range(wordNum):
13         word = words[i]
14         check = True
15         for j in range(5):
16             if (correctGuess[j] == '*' and word[j] in correctGuess) or \
17                     (correctGuess[j] != '*' and correctGuess[j] != word[j]):
18                 check = False
19                 break
20             if word[j] in wrongGuess:
21                 check = False
22                 break
23         if not check:
24             tag.append(0)
25         else:
26             tag.append(1)
27             sumCount += wordCount[i]
28     # sum P(L|W)
29     for i in range(wordNum):
30         if tag[i] == 0:
31             continue
32         p = float(wordCount[i]) / sumCount
33         for l in range(26):
34             if not guessChar[l] in correctGuess and not guessChar[l] in wrongGuess and
    guessChar[l] in words[i]:
35                 guess[l] += p
36     # get best guess with max probability
37     maxP = max(guess)
38     best = chr(65+guess.index(maxP))
39     return [best, maxP]
40 # pre process
41 inputFile = open('hw1_word_counts_05.txt')
42 words = []
43 wordCount = []
44
45 for line in inputFile.readlines():
46     data = line.strip('\n').split(' ')
47     words.append(data[0])
48     wordCount.append(int(data[1]))
49 # sort words according to count number
50 wordNum = len(words)
51 sortIndex = sorted(range(wordNum), key=lambda k: wordCount[k], reverse=True)
52
53 for i in range(0, 8):
54     index = sortIndex[i]
55     print words[index], wordCount[index]
56
57 for i in range(wordNum - 8, wordNum):
58     index = sortIndex[i]
59     print words[index], wordCount[index]
```

```
60
61  # compute 1.7 b)
62  print bestGuess(wordNum, words, wordCount, '*****', [])
63  print bestGuess(wordNum, words, wordCount, '*****', ['E', 'O'])
64  print bestGuess(wordNum, words, wordCount, 'D**I*', [])
65  print bestGuess(wordNum, words, wordCount, 'D**I*', ['A'])
66  print bestGuess(wordNum, words, wordCount, '*U***', ['A', 'I', 'E', 'O', 'S'])
67
```