Homework 05
Wenyu Yang

Questions specific to this assignment:

```
○ > fact
  correct: fict

  > poame
  correct: poeme

  > uneque
  correct: unique
```

Please give an example of a word which was correctly spelled by the user, but which was incorrectly "corrected" by the algorithm. Why did this happen?
> fact
correct: fict

In the case of "fact" → "fict", the algorithm made an incorrect correction to an already correct word. The error occurred because the algorithm likely put too much emphasis on pattern matching with common word stems like "fiction" and "fictitious", while failing to properly validate against a dictionary or consider the high frequency of the word "fact" in common usage. This reveals a fundamental flaw in the algorithm's weighting system, where it prioritized potential word relationships over actual word validity.

Please give an example of a word which was incorrectly spelled by the user, but which was still incorrectly "corrected" by the algorithm. Why did this happen?
> poame
correct: poeme

For "poame" → "poeme", both the user's input and the algorithm's correction were incorrect, as the intended word was "poem". The algorithm failed here because while it recognized the basic structure of the word, it suggested a variant that exists in French but not in English. This suggests the algorithm lacks proper language context awareness and may be conflating multiple language dictionaries. Additionally, it appears to prioritize minimal edit distance over linguistic correctness, leading to a technically close but semantically wrong suggestion.

Please give an example of a word which was incorrectly spelled by the user, and was correctly corrected by the algorithm. Why was this one correctly corrected, while the previous two were not?

> uneque
correct: unique

The correction of "uneque" → "unique" succeeded where the others failed because of a perfect confluence of favorable factors. The misspelling involved only one letter difference, following a common phonetic substitution pattern (e/i confusion) that the algorithm likely encountered frequently in its training data. Furthermore, "unique" is a high-frequency English word with few similarly spelled alternatives, making it an unambiguous target for correction. The algorithm's pattern matching and frequency weighting worked as intended in this case, demonstrating how spelling correction works best when dealing with common words and typical spelling errors.

The usual questions:
How long did this assignment take you? (1 sentence)
I worked on this assignment for 3 days.

Whom did you work with, and how? (1 sentence each)
I worked on this assignment by myself, reviewing lecture materials.

Discussing the assignment with others is encouraged, as long as you don't share the code.
Which resources did you use? (1 sentence each)
I discussed with my classmate and I primarily used the course slides as a resource for this assignment.

For each, please list the URL and a brief description of how it was useful.
https://docs.google.com/presentation/d/1fGzSMYKyx4Vr-BEVJzXKnjANyK3AuCKyptepqXSAV9U/edit#slide=id.g30882c6eb14_0_484

A few sentences about:
What was the most difficult part of the assignment?
The most difficult part of the assignment was crafting complete, analytical explanations of why the algorithm made different decisions in each case. It required understanding both the technical aspects being able to explain these concepts clearly while maintaining a logical flow.

What was the most rewarding part of the assignment?
The most rewarding part was seeing how each analysis built upon the previous ones to create a comprehensive understanding of the spell-checking algorithm's behavior. Breaking down why the algorithm succeeded in one case while failing in others helped reveal the complex interplay of factors involved in spell-checking.

What did you learn doing the assignment?
I learned the importance of systematic analysis and clear communication of technical concepts. I gained deeper insight into how spell-checking algorithms balance multiple competing factors - like edit distance, word frequency, and language patterns.

Constructive and actionable suggestions for improving assignments, office hours, and class time are always welcome.
To enhance class time, more interactive analysis exercises would allow students to practice explaining algorithmic behavior in real-time before tackling assignments independently.