

Problem 237: Searching for a Sign

Difficulty: Medium

Author: Kelly Reust, Denver, Colorado, United States

Originally Published: Code Quest 2024

Problem Background

If you've ever used a search engine before - and honestly, who hasn't used Google on a near daily basis - you're familiar with keyword searching. Given a word or a short phrase, a search engine is able to scan through large amounts of text in order to find sentences that include that keyword. Most search engines help the user to determine how useful a search result is by showing the keyword's context - the text surrounding the keyword that provides additional information as to what is being discussed.

Problem Description

You're working with Lockheed Martin's Enterprise Information Technology division to help create a new search engine for company policies, news articles, and other internal information. To try to improve the performance of the search engine, your team will conduct several searches on common search terms in advance, then save the results of those searches to be presented to users whenever they wish to do a search using those same terms.

Your search engine will need to provide context with each of its search results, highlighting the search keyword within that context. Your team has decided that the best way to provide this context is to show up to three words before the keyword and up to five words after the keyword, all within the same sentence. The keyword itself should be highlighted by placing asterisks on either side of it.

For example, if someone were to use the search engine on the previous paragraph to search for the word "engine," they should see this portion of the first sentence:

*Your search *engine* will need to provide context...*

Only two words precede the word "engine" in this sentence, so we include both of those words. The word "engine" is highlighted, and followed by the next five words in the sentence. If the user were to search that paragraph for the word "sentence," they would see this result instead:

*...within the same *sentence.**

Here, we provide the three words before "sentence." Since "sentence" is the last word in the sentence, we don't print anything after it. In both cases, we include an ellipsis (three periods: ...) to indicate that there is more to the sentence in the corresponding direction.

In the event that the keyword appears more than once in a sentence, treat each appearance as a separate search result, with its own context (which may include a non-highlighted keyword). For example, searching that same paragraph for the word “search” yields three results, all from the first sentence:

*Your *search* engine will need to provide...*
*...each of its *search* results, highlighting the search keyword...*
*...results, highlighting the *search* keyword within that context.*

A “word” should be considered to be any collection of letters and punctuation separated from other words by spaces or the start or end of a line. A sentence ends with a period (.), question mark (?), or exclamation point (!). Remember to ignore punctuation when searching for keywords, however; note the second example above, where the highlighted word “sentence” includes the period at the end of the sentence.

Sample Input

The first line of your program’s input, received from the standard input channel, will contain a positive integer representing the number of test cases. Each test case will include:

- A line containing a positive integer, R, representing the number of lines of reference text
- R lines containing the reference text to be searched against. Lines may contain up to 80 characters each, including letters, spaces, and punctuation symbols. Sentences may be spread across multiple lines of text.
- A line containing a single keyword to search for within the reference text. Keywords contain lowercase letters only. Searches are not case-sensitive.

1
4

This is a sample sentence within a sample paragraph. All of this is further held within a sample input. Below this is the section about the sample output. Remember that your program must work on this sample input and output as well as the secret, “official” input and output.
sample

Sample Output

For each test case, your program must print the search results for the given keyword, one result per line, in the order in which they appear in the reference text. Results should contain an appropriate amount of context, as described above, and should highlight the keyword with asterisks (*). Text should otherwise remain as shown in the reference text.

This is a *sample* sentence within a sample paragraph.
...sentence within a *sample* paragraph.
...held within a *sample* input.
...section about the *sample* output.
...work on this *sample* input and output as well...