WC '97 P1 - Searching Quickly

Time limit: 1.0s **Memory limit:** 16M

1997 Woburn Computer Programming Challenge

Searching and sorting are part of the theory and practice of computer science. For example, binary search provides a good example of an easy-to-understand algorithm with sublinear complexity. Quicksort is an efficient $\mathcal{O}(n \log n)$ [average case] comparison sort. KWIC-indexing is an indexing method that permits efficient "human search" of, for example, a list of titles.

Given a list of titles and a list of "words to ignore," you are to write a program that generates a KWIC (Key Word In Context) index of the titles. In a KWIC-index, a title is listed once for each keyword that occurs in the title. The KWIC-index is alphabetized by keyword. Any word that is not one of the "words to ignore" is a potential keyword. For example, if the words to ignore are `the`, `of`, `and`, `as`, and `a` and the list of titles is

- descent of man
- the ascent of man
- the old man and the sea
- a portrait of the artist as a young man

A KWIC-index of these titles might be given by

a portrait of the	ARTIST	as a young man
the	ASCENT	of man
	DESCENT	of man
descent of	MAN	
the ascent of	MAN	
the old	MAN	and the sea
a portrait of the artist as a young	MAN	
the	OLD	man and the sea
a	PORTRAIT	of the artist as a young man
the old man and the	SEA	
a portrait of the artist as a	YOUNG	man

Input Specification

The first few lines of the input are a series of "words to ignore." Each is on a line by itself and is no more than 20 characters in length.
The list of words to ignore is terminated by the string `::` on a line by itself.
The next lines are the titles to be incorporated into the KWIC-index. Each contains no more than 15 words, all in lowercase, separated by single spaces. No word is longer than 20 characters, and no title is longer than 250 characters.

There will be no more than 50 words to ignore, no more than 100 titles, and no more than 200 keywords in the titles. No characters other than `a` to `z` and white spaces will appear in the input.

Output Specification

The output should be a KWIC-index of the titles, with each title appearing once for each keyword in the title, and with the KWIC-index alphabetized by keyword. If a word appears more than once in a title, each instance is a potential keyword. The keyword should appear in all uppercase letters; all other words should be lowercase letters. Titles in the KWIC-index with the same keyword should appear in the same order as they appeared in the input file. In the case where multiple instances of a word are keywords in the same title, the keywords should be capitalized in left-to-right order. The titles in the KWIC-index should NOT be justified by keyword, and must be listed left-justified.

Sample Input

```
is
the
of
and
as
a
but
::
descent of man
the ascent of man
the old man and the sea
a portrait of the artist as a young man
a man is a man but bubblesort is a dog
```

Sample Output

```
a portrait of the ARTIST as a young man
the ASCENT of man
a man is a man but BUBBLESORT is a dog
DESCENT of man
a man is a man but bubblesort is a DOG
descent of MAN
the ascent of MAN
the old MAN and the sea
a portrait of the artist as a young MAN
a MAN is a man but bubblesort is a dog
a man is a MAN but bubblesort is a dog
the OLD man and the sea
a PORTRAIT of the artist as a young man
the old man and the SEA
a portrait of the artist as a YOUNG man
```