

Bài G. Chuẩn hoá câu văn

Mã bài: CHUANHOA

Thời gian: 1,0 giây Bộ nhớ: 256 MB Dữ liệu: CHUANHOA.in Kết quả: CHUANHOA.out

Đề bài

Khi gõ vội, học sinh hay để thừa dấu cách: thừa ở đầu dòng, thừa ở cuối dòng, hoặc gõ liền mấy dấu cách giữa hai từ. Phần mềm chấm bài của trường cần đưa mọi câu về dạng chuẩn trước khi so sánh.

Câu ở dạng chuẩn là câu không có dấu cách ở đầu, không có dấu cách ở cuối, và giữa hai từ liên tiếp có đúng một dấu cách. Ở đây một *từ* là một dãy ký tự liên tiếp khác dấu cách và không kéo dài thêm được nữa, nghĩa là ngay trước và ngay sau nó hoặc là dấu cách, hoặc là hết câu.

Cho một câu, hãy cho biết câu đó có bao nhiêu từ và viết lại câu ở dạng chuẩn.

Dữ liệu

Vào từ tệp văn bản CHUANHOA.in:

Một dòng duy nhất chứa câu S . Câu có thể chứa dấu cách ở bất kỳ đâu, kể cả đầu dòng và cuối dòng.

Kết quả

Ghi ra tệp văn bản CHUANHOA.out:

- Dòng thứ nhất ghi số từ của câu.
- Dòng thứ hai ghi câu ở dạng chuẩn.

Ràng buộc

- $1 \leq |S| \leq 10^5$;
- mọi ký tự của S đều là ký tự ASCII in được, tức có mã từ 32 tới 126;
- S chứa ít nhất một ký tự khác dấu cách, nên câu luôn có ít nhất một từ.

Chấm điểm

Bài gồm 2 subtask với tổng điểm 6. Thí sinh được trọn điểm của một subtask khi và chỉ khi chương trình cho kết quả đúng ở *tất cả* các test của subtask đó.

Subtask	Điểm	Ràng buộc bổ sung
1	2,5	$ S \leq 1000$
2	3,5	không có ràng buộc nào thêm

Ví dụ

Ví dụ 1

Dữ liệu (CHUANHOA.in)

Hoc Tin hoc that thu vi

Kết quả (CHUANHOA.out)

6
Hoc Tin hoc that thu vi

Ví dụ 2

Dữ liệu (CHUANHOA.in)

lop 11A4, to 3

Kết quả (CHUANHOA.out)

4
lop 11A4, to 3

Giải thích ví dụ

Ở Ví dụ 1 câu vào có ba dấu cách ở đầu, hai dấu cách ở cuối, và vài chỗ giữa các từ bị gõ thừa. Hình dưới đánh dấu các ô là dấu cách bằng ô tô xám.

				H	o	c				T	i	n					h	o	c	...
--	--	--	--	---	---	---	--	--	--	---	---	---	--	--	--	--	---	---	---	-----

Sáu từ thu được là Hoc, Tin, hoc, that, thu, vi. Nối lại bằng đúng một dấu cách ta được câu chuẩn.

Ở Ví dụ 2 chỉ có một chỗ thừa dấu cách, sau dấu phẩy. Dấu phẩy dính liền với 11A4 nên 11A4, được tính là một từ; đề không tách dấu câu ra khỏi từ.

- HẾT -