

THỐNG KÊ TỪ KHÁC NHAU TRONG FILE NHỊ PHÂN

Bài làm tốt nhất

Cho file nhị phân DATA.in có một ArrayList<String> được ghi vào file theo kiểu writeObject().

Người ta định nghĩa một từ là dãy ký tự không chứa khoảng trắng, dấu tab và dấu xuống dòng, đồng thời không tính các dấu câu như dấu phẩy (,) dấu chấm (.) dấu chấm hỏi (?) dấu chấm cảm (!) dấu hai chấm (:) dấu chấm phẩy (;) dấu ngoặc đơn, dấu gạch ngang (-), dấu gạch chéo (/).

Hãy đọc dữ liệu từ file và liệt kê các từ khác nhau xuất hiện theo thứ tự số lần xuất hiện giảm dần.

Chú ý:

- Khi thống kê thì không phân biệt chữ hoa, chữ thường.
- Bỏ qua các dấu câu đã liệt kê ở trên khi tách từ

Input - file nhị phân DATA.in

Chứa duy nhất một ArrayList<String>

Output

Ghi ra danh sách các từ (ở dạng in thường) theo thứ tự số lần xuất hiện giảm dần.

Trong trường hợp số lần xuất hiện bằng nhau thì liệt kê theo thứ tự từ điển tăng dần.

Ví dụ

Input
File nhị phân DATA.in theo mô tả trong đề bài.
Output
Ghi ra theo yêu cầu đề bài, mỗi từ trên một dòng đi kèm số lần xuất hiện. Ví dụ: tuong 99 huong 11

lap 10

trinh 8

doi 5