Stream Programming - 2019
Laboratory 4

Task Implement the extended version of Mirsa-Gries algorithm in Scala. The standard algorithm finds such elements in the stream which occur more than n/k times where n denotes the length of the stream and k the number of counters. The extended version of the program should work for different object types. In the program, the classes like: *cars*, *books* and *employees* should be defined. Then, the stream of such object can be provided. The algorithm should work for all object types separately. For example, it should find the object of type *car* which appear more than 5 times.

Hint 1 The Misra-Gries summary is a simple algorithm that solves the frequent items problem (see here). It can be viewed as a generalization of Majority to track multiple frequent items. Instead of keeping a single counter and item from the input, the Misra-Gries summary stores $k-1$ (item,counter) pairs. The natural generalization of Majority is to compare each new item against the stored items T, and increment the corresponding counter if it is among them. Else, if there is some counter with count zero, it is allocated to the new item, and the counter set to 1. If all $k-1$ counters are allocated to distinct items, then all are decremented by 1. A grouping argument is used to argue that any item which occurs more than n/k times must be stored by the algorithm when it terminates.

Hint 2 At first, you should probably count the number of different object types in the stream. The object types and their attributes should be hard implemented in the program.