**Problem 1:**

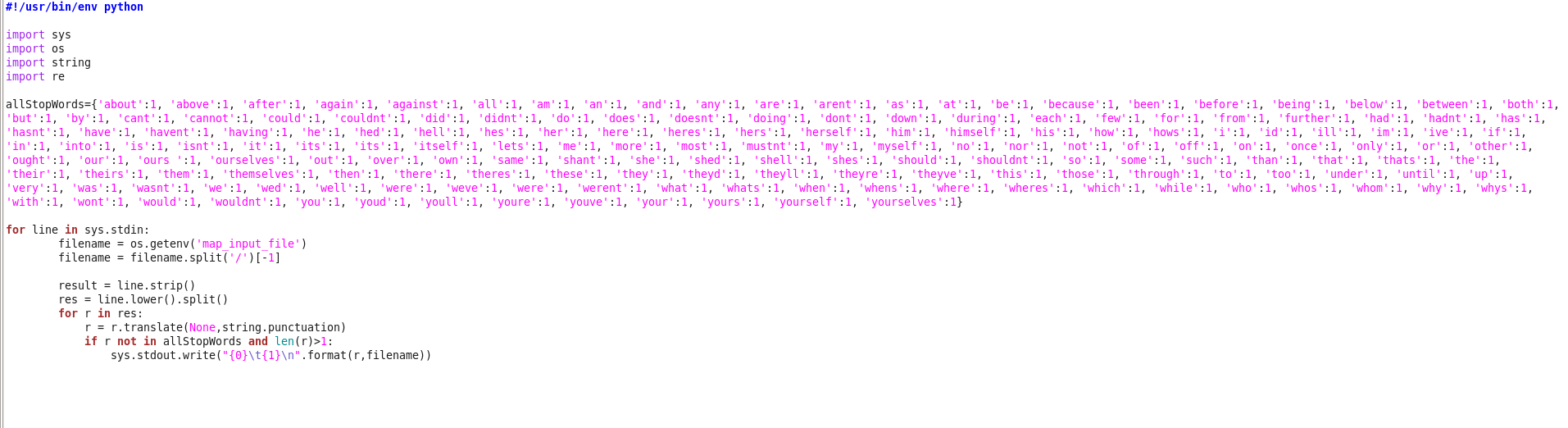
**1.Write a map-reduce program that creates an inverted list of words and the files that contain them. That is, for each word, you will display a list of files that contain that word. A sample output is given below:**

**‘anger’ [‘histories’, ‘tragedies’]**

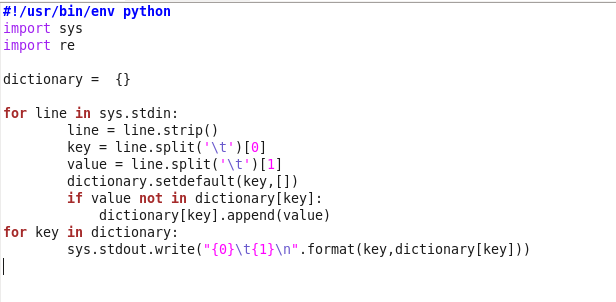
**‘laugh’ [‘comedies’, ‘poems’, ‘histories’]**

**…….**

**Mapper program (mapper.py)**

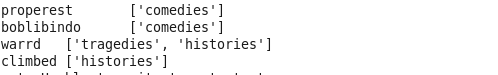


**Reducer program(reducer.py)**



**Output of first 50 lines:**





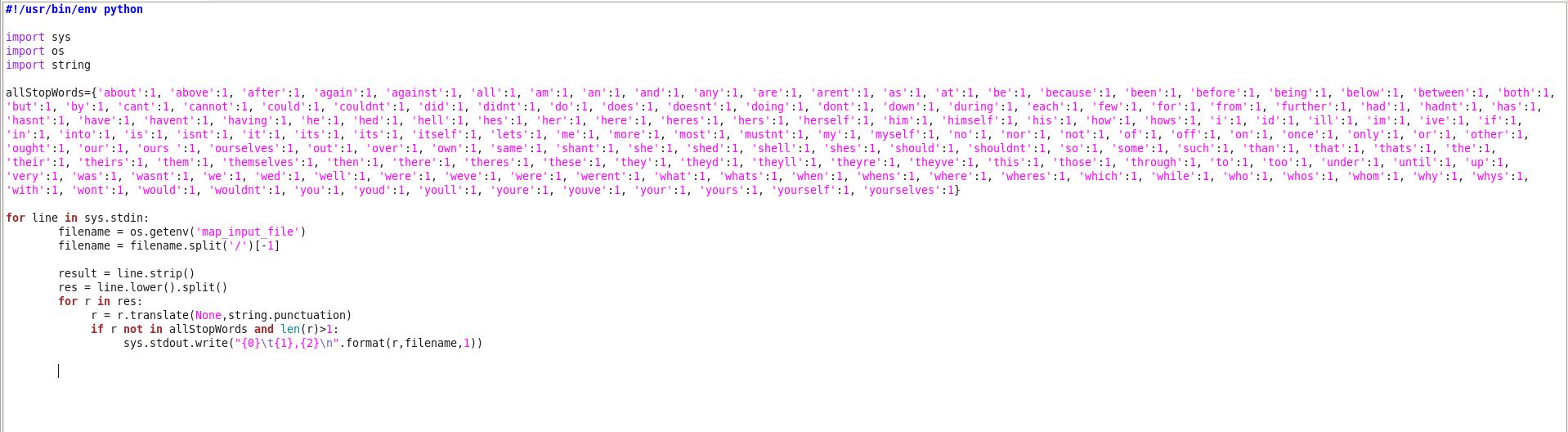
**2.Modify your map-reduce program (in 1) to display the number of times the word occurs in each file. A sample output is shown below:**

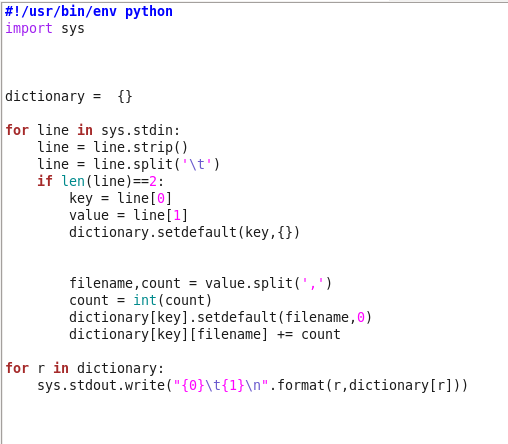
**‘anger’ {‘histories’: 3, ‘tragedies’: 8}**

**‘laugh’ {‘comedies’: 7, ‘poems’: 2, ‘histories’: 15}**

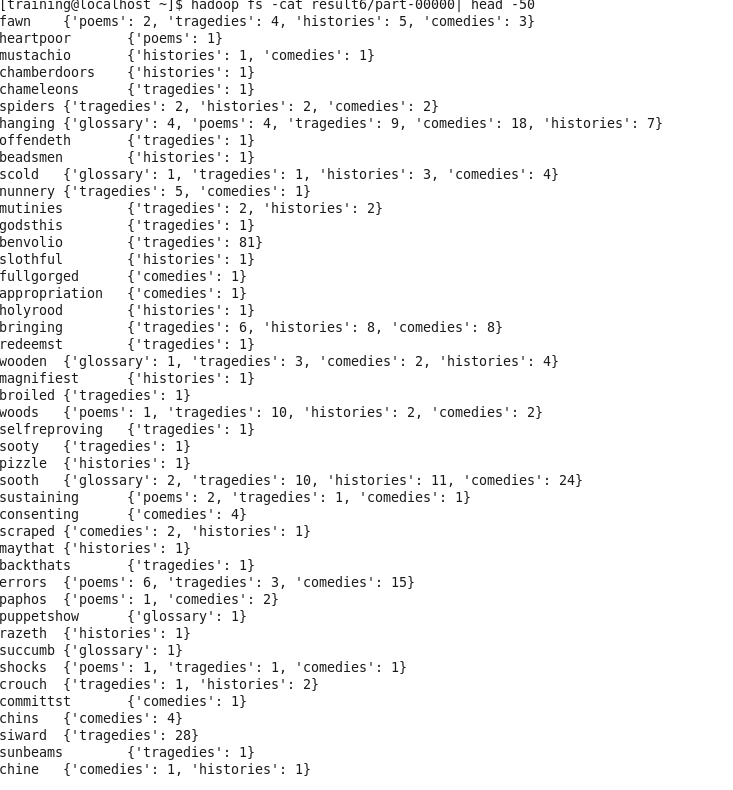
**………**

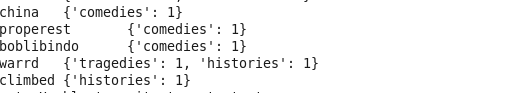
**Mapper program (mapper1.py)**

**Reducer program (Reducer1.py)**



**Output of first 50 lines:**





**Problem 2:**

**Each file in this archive contains entries that look like:**

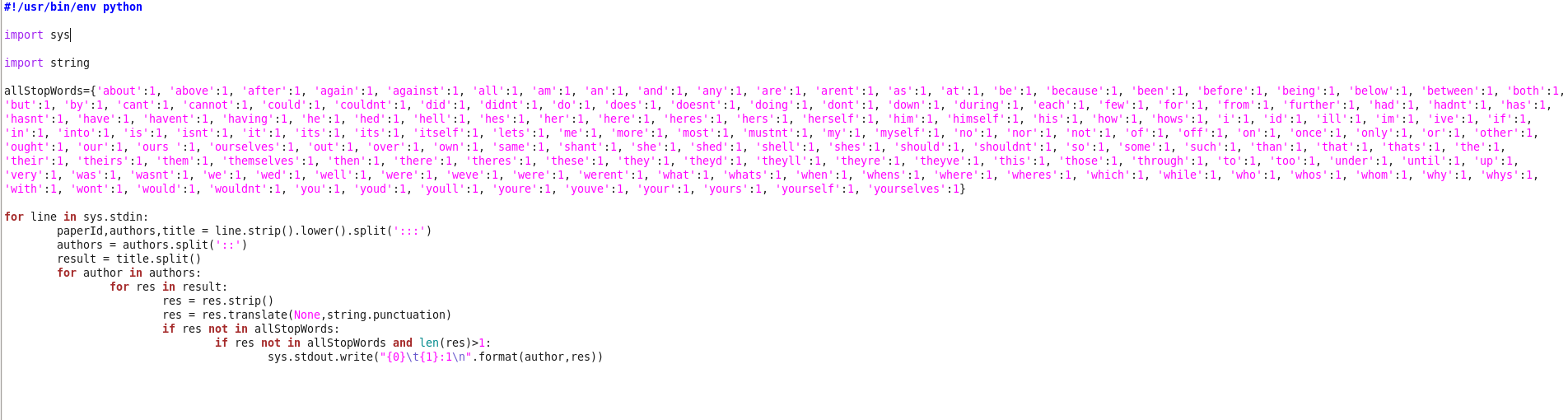
**journals/cl/SantoNR90:::Michele Di Santo::Libero Nigro::Wilma Russo:::Programmer-Defined Control Abstractions in Modula-2.**

**that represent bibliographic information about publications, formatted as follows:**

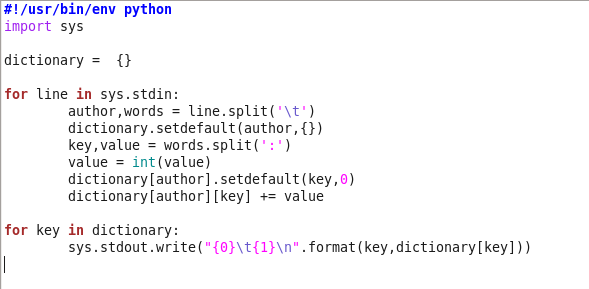
**paper-id:::author1::author2::…. ::authorN:::title**

**Task is to compute how many times every term occurs across titles, for *each* author**

**Mapper program (mapper2.py)**



**Reducer program (Reducer2.py)**



**Output of first 50 lines:**

