

示例程序

```
import os
import numpy as np

f=open("50W.txt", mode = "r", encoding = "utf-8")
fstr = f.read()
flist = fstr.split("\n")
for i in range(len(flist)):
    flist[i].strip()
f.close()

def func1(paramC, paramDic):
    if paramC in paramDic:
        paramDic[paramC]+=1
    else:
        paramDic[paramC]=1

def func2(paramDic, paramNum):
    li = sorted(paramDic.items(), key=lambda item:item[1], reverse=True)
    _str=""
    for i in range(paramNum):
        _str=_str+li[i][0]+","+str(li[i][1])
        if i!=paramNum-1:
            _str+=";"
        else:
            _str+="\n"
    return _str

def func3(paramDic, paramNum):
    arr = np.zeros(20000)
    li = sorted(paramDic.items(), key=lambda item:item[1], reverse=True)
    for i in range(paramNum):
        if arr[li[i][1]]==0:
            arr[li[i][1]]+=1
        else:
            print(li[i][1])
            return

# processing
dict_lastname = {}
dict_lastcha={}
dict_lasttwocha={}
dict_3begintwo={}
dict_pattern={}
for item in flist:
    func1(item[0], dict_lastname)
    func1(item[-1], dict_lastcha)

    func1(item[-1], dict_lasttwocha)
    func1(item[-2], dict_lasttwocha)
```

```
if len(item)>=3:
    func1(item[0:2], dict_3begintwo)

for i in range(len(item)-1):
    if item[i]==item[i+1]:
        func1(item[i:i+2], dict_pattern)

# func3(dict_lastname, 100)
# func3(dict_lastcha, 100)
# func3(dict_lasttwocha, 100)
# func3(dict_3begintwo, 20)
# func3(dict_pattern, 20)
result_str = func2(dict_lastname, 100)+func2(dict_lastcha, 100)+func2(dict_lasttwocha, 100)

# save
if not os.path.exists("result.txt"):
    f = open("result.txt", "w", encoding="utf-8")
    f.write(result_str)
    f.close()
```

主要错误总结:

输出格式不规范, 不能打印出来再手动拷贝到txt, 不要输出多余的空格、句号, 注意区分全角分号和半角分号。请尽量使用utf8编码。提交的作业中尽量使用相对路径读写文件。请勿double后缀名, 不要出现x.txt.txt、x.py.py。如需压缩提交的多个作业文件, 尽量别使用rar, 请使用zip格式压缩, 这种格式兼容性较好。

部分同学与标答有出入, 可能是index范围不对。若使用range, 思考下index的范围是多少, 若使用L[i-1], 思考下i为0的情况。