COMP 7402 - Assignment 1

Tehrani Parsa, Peyman - A00922386

January 24, 2019

Task 1

Usage

You'll need golang to compile the code using the "make" or use the provided binarys in /bin

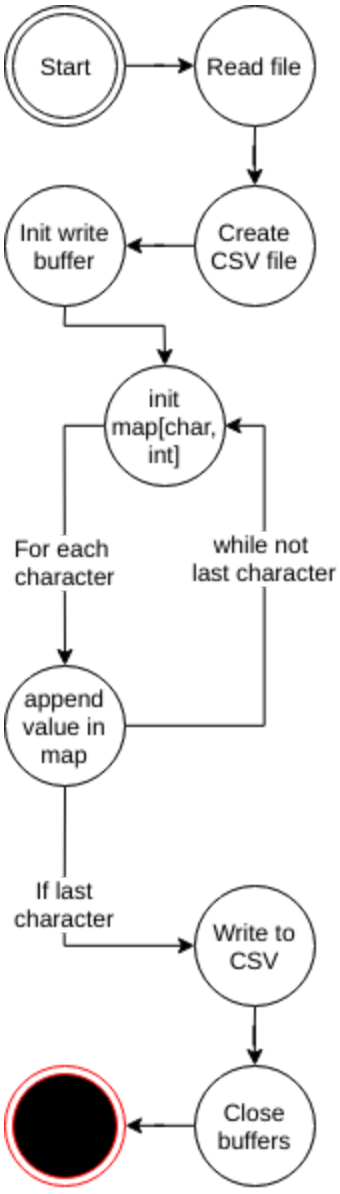
To run: Asn1 <input file> <output file>

Design

Pseudocode

```
Bstr = ReadFile Args[1]
File = Create(Args[2] + ".csv")
writer = New Writer for file
Make Map[string, int]
For every character in Bstr
     If key exists in Map append value by 1
     Else Create Key with value of 0
Writer write to File for every key,value in map
file.Close()
writer.Flush()
```

State diagram



Task 2

Distribution	Conditional probability	Answer
P(M=e)=0.1260302198	$P(M=e \mid c_i)$=0.1260302198*(1/26)	0.004847316145
P(M=t)=0.09920738046	$P(M=t \mid c_i)$=0.09920738046*(1/26)	0.003815668479

P(M=a)=0.08161939412	$P(M=a \mid c_i)$=0.08161939412*(1/26)	0.003139207466
P(M=i)=0.06975794476	$P(M=i \mid c_i)$=0.06975794476*(1/26)	0.002682997875
P(M=o)=0.07561441936	$P(M=o \mid c_i)$=0.07561441936*(1/26)	0.002908246899
P(M=n)=0.06515444015	$P(M=n \mid c_i)$=0.06515444015*(1/26)	0.002505940006

Appendix

Data set

character	found in set	distribution
e	13579	0.1260302198
t	10689	0.09920738046
a	8794	0.08161939412
o	8147	0.07561441936
i	7516	0.06975794476
h	7374	0.06844000594
n	7020	0.06515444015
s	6502	0.06034674785
r	5440	0.05049005049
d	4934	0.04579373329
l	4718	0.04378898129
u	3469	0.03219668845
w	2676	0.02483664984
g	2531	0.02349086724
c	2400	0.02227502228
y	2262	0.02099420849
m	2107	0.01955561331
f	2001	0.01857179982
p	1524	0.01414463914
b	1475	0.01368985744
k	1158	0.01074769825
v	847	0.007861226611
q	209	0.00193978319
x	148	0.001373626374
j	146	0.001355063855
z	78	0.0007239382239

Distribution chart