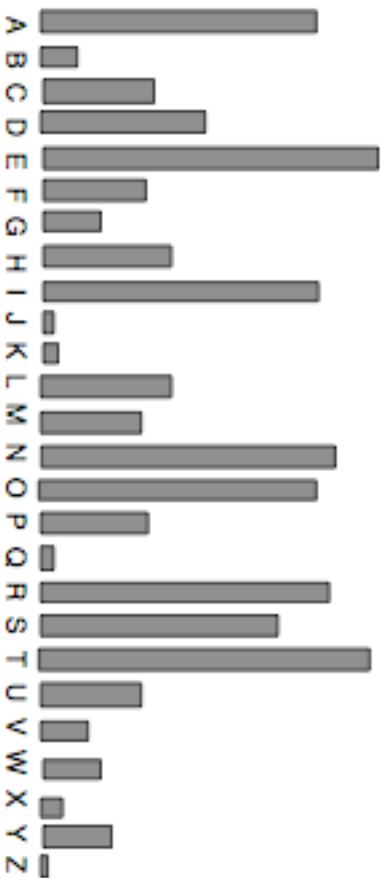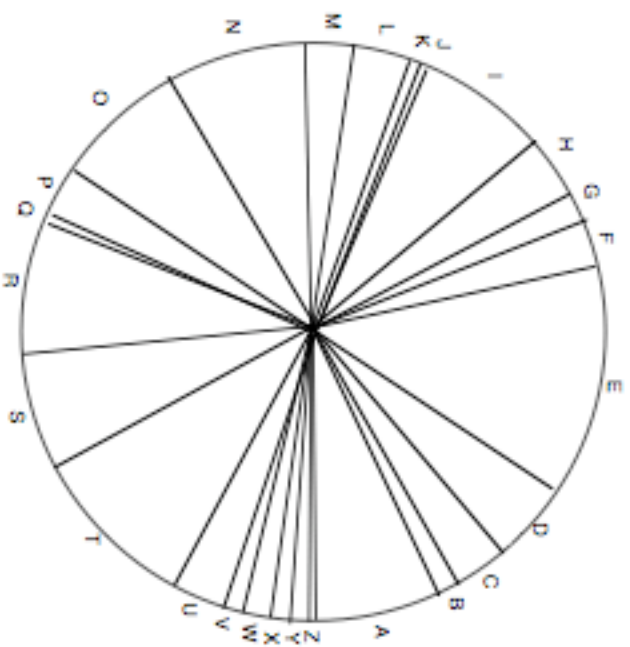# Some Monkey Words

The letters of the following words were chosen at random with equal probability.

```
zotwz kjptg uudrx mtagh gaqvl uhuaf oymcu iljaa jancm bxpbi cvcdo oceza wrgzg
wweuk phqvq rqtwq axqqm yxngh rsxch gsocp sssda jmjhl ykyml kdefv tmafc ncsiu
anunv hcwyy ueikr ifzta abjrq kiaez abjka hufwa fjenl ptpqm exjhx vzmeg uayxu
dkxid agsst gutif qzmyc lvood sobrh tkjoq vkghf rgxad frywd bkwql uvcpk qjdkk
mugak krwcb gpgsf dwsxu vafik sxiev lejjq pznet fjtqc jhnic vojrv ucfib nuzgb
xahvu lvwyq spvtf zdxob aigfa wgjeq wcqoe bbmbm fdrvx znhbo ovxqu lzznh vmdjj
kfkig jnugp xxgnf hhliw oveje mvujj hqucs plcuo aijqt hlgdj pvhhi jwyqo wbhbo
qcwtc ahbyt xbvrc kvcby etllo bgouc lwwer ewzpj yhysb grtba betir hekem mhjwl
zflze jtalc gxvhz rkaxp qastv ibuzn jvqvt vtxgw nvnlj muqwx nddwh gggrk jqgrp
xxahd tbvuw hgjya bnoqg yfwap gohru huluh xmhtu kbbxv toywf imblw shacr taaqv
xyamm ggbcp ykial aipgv osyqq iwwrw ivhpj ybugr rscvx zbdgy ndesm endwt tbgmd
yroti vvqas xbixy qzckh ayzaz bmsns macws skbqb orcva mxaah wdjax rpfwa tgepf
ckjgv msyiv nkemr lksbd fyrtv sgxlb ryqbd zkjed uornf fgcft harsx xonas pttlr
qrraj hwerm oanup pqkww wmtog sjblk hdgac etadi ljlql tfkla bfxug dwfow corfo
qnwjd tnaib sicil bukam jxhjx uyygr yhfxq goxor vgxsb ngnka pojfr mgoni rntkk
yjxua llggz yyuzf bijhf pfctu hnipg oczle glfli lozhs qcvmw ueuyl jnuxp xzpaz
otkyt jtaws ejigu xiqin txded zghgb jtovo oiaqa moiar yklzv kxmpe rtxwz qpvkz
chqxl mixcz kchjw xeljr hryqr kdbyp tirep whomf eoxwh iblae gbary qqhmi wcaam
vkyup zlkjx fwium pwske xfchv xxcap tauyz jqodr lehsc ssvos avjoh pbqif kpuch
bhmvv yojmr dhfvw gwruc xjywu fmlqe joqbi pwhci hcnjq xwktm sslpc ktwwd biwbl
hqtxz yismt tpovm vbdxk nibab ukxqo ahyzk vrwhb fgezq tkves veqlr pjiww dheql
zgyyt gjbjp npobo rcncd pxaqj rzolr ibtfe ddfri ahglo cjfcf kjqrp cacwt aipoe
jcndr nhnhi efugo blkhi ahvjj rxoxq mbcnn wlvqd ryoel oprtj grotg gnien jvruf
atedb piprx bywxl wluwv shxfx xmxxm aodlg mbfxk otpix rcfgo ylffg jyzvx wyvbe
exovx fgrqz umawo uyfua mzruo muols rpfyf fwkto tgdpn erwbi zqsva eyyxf mleas
trwhd nmosb kknqo nfiwm jaxhq woxeq sqipi xemfj cpxcq yhkvb zyelo mnaky rovgt
yzrek grdzk qvkog pgxpx gghhz lclgc osdrd biedw eujdz onvuz dcuoc vothj rtodx
dygqw ezhkf vghih vzyny aswum ulwcd jxsgf aibfp iwmwq qepaw npazz pwbzj aqpqo
jrfys ccegi anrge yrzzw hkwvs csrmo kpcov bvsqk cmhlv kapoo mszzc zzoql vpzxp
wsiyx oiljc joriz nnrfm prifj nnkaa zmdym kahue nicek zlxbo kelin wtvxd hekse
```

Letter		Count
A	→	73
B	→	9
C	→	30
D	→	44
E	→	130
F	→	28
G	→	16
H	→	35
I	→	74
J	→	2
K	→	3
L	→	35
M	→	25
N	→	78
O	→	74
P	→	27
Q	→	3
R	→	77
S	→	63
T	→	93
U	→	27
V	→	13
W	→	16
X	→	5
Y	→	19
Z	→	1

Letter		Count
E	→	130
T	→	93
N	→	78
R	→	77
I	→	74
O	→	74
A	→	73
S	→	63
D	→	44
H	→	35
L	→	35
C	→	30
F	→	28
P	→	27
U	→	27
M	→	25
Y	→	19
G	→	16
W	→	16
V	→	13
B	→	9
X	→	5
K	→	3
Q	→	3
J	→	2
Z	→	1

Some English Monkey Words

A monkey types on a keyboard but this time the frequencies of letters that he types are the same as one would find in average English text.

```
ccgai teeir euerd ommal iltom eeoys uneep risrf ocoee ifspo maxpu eerci upomr
samiu eeitr eeopo cuoic uelnr eoeew usleo mestr etfok thiee ocart tssss rseaa
anior oisda fteoa doals oiatl uyvre tiiit hifhq loean pvoom nuala appes aicha
capts ttwnh rsios fohln efsmh hnfee wtdwo ptnip toipi eretn jrpsh suodo rtuee
eosrf ueats tsedt aalet tplcf eanin tetdt evaos lefee dffsm eotfu seber keega
ehafe raaes uhtan nnspr dogoh esncn oooed hrrnu umrky hrhbe rttsn daofo otiqi
deepe fnaso atnnr heaes sinfa ogoeh oxint tdehr wsest reknd hcpte iaogg tdfro
ttsiu nrnle jpeof xyatt asdis aopon aoyrh ahsos dytic unswa efdrt eodrd toiro
efeni seief hnvao itieu ffdht pottr exewo darnh teeos yilem taros rclof octns
ltdte liatr altsi eescr tsiru oavre dtiay esait kttha fpdqt iteas ooocf rtmee
rwtez isaru ptask uneea cbsoi sedsr tocre nitda utous ehtia dnifi ratla defls
yiuib aldcn neyed dqsse ogtsi wiexu erpas rsoyr ocpdi oitdu opsju neesh sneea
htnol smrtf ionfd ntcoo himam arhyv olodp eanei roaie clued fecrm haeei opnar
hilau tprrt etaen htnfs naahi tiavf dorat idcmp nlyic eeiir nmgih cuuie sopeo
lifdg rowni lftif eodfe ceese haeii pfnrp yenin aaeoi drnll nawtl anite eoeor
oreel ehtrt wrree ortcd rnnan seiij yeodf qetwl instb ynrcx rdcth ftrne oeath
eospr qclea ntyep emifo raotx eaodx eklai rntpn rcaiw euvao caiio udlvt eghte
swvel ltase axpth ypati hotfn xoevt hotor oomcr yhlof eossc maego sahdt waopo
reeza elrdn itgen sauwu otdpe meeor wtnot dnvtt leuen eeeod tsdml nendf lnhta
chdeo rrnli kmolx nfwul ptste ntdwe ptehn pnosn dnoet hnols adtes asnen heoey
dofsr oyprg foina sosie nceoe suxlp eirpe tarac nyaot oedfa oaivf edgfn orgrt
nmrer noaes xeisi dttex mteiw ptwot afkwi deiee nludn iiinn nitlr tripb elret
eedrl ngsan tacse foudh oinnu xiiep footn rdlep ryvtn nlstn fneit ewemr eryae
pssnu ctmye ncenm aosnm ounnn ouswn tmref nnnol msplt meptc ebyke smeni inecl
edoen rnctr airwn mbere oanty sefcp eekdo steat eahir ootaf tfdel oslor oerop
enmst nssda gesrf nmtie staem uhfil erpun retel pdoea evwmr dnuvi aetme fwdeo
avlag syera nsnth eidns mitcc refst tfede lpsrw nswtn snere iloaa ioesy isprr
yuacm eoene rdrum rrtmy ubmol taemr ansst xeccs rrltn osnne nnern onsir mnite
stwss utrst iudeo dnnau tyllt etnoo ereho nslyl elpee sryen yctsm camee uheuw
```

TABLE OF ENGLISH BILETTER CONDITIONAL PROBABILITIES

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	11	193	388	469	20	100	233	20	480	20	103	1052	281	1878	8	222	0	1180	1001	1574	137	212	57	26	312	23
B	932	16	0	0	3220	0	20	0	605	57	0	1243	49	965	662	0	0	229	49	16	727	0	0	0	1165	0
C	1202	0	196	26	1707	0	0	1277	761	0	324	369	15	11	733	0	4	426	87	893	601	0	0	0	94	0
D	1044	20	26	218	3778	7	132	7	1803	33	0	178	53	73	1228	0	0	495	893	213	434	243	0	0	264	0
E	660	36	433	1195	438	142	21	158	456	5	36	340	9	1382	40	192	0	1231	404	48	215	152	205	152	121	4
F	838	0	0	0	1283	924	0	0	1608	0	0	299	9	0	2788	0	0	1215	496	462	48	0	0	0	43	0
G	1078	0	0	18	2393	0	177	1281	839	0	0	27	0	0	1140	0	0	1325	256	247	512	0	0	0	53	0
H	1770	5	14	8	5624	0	0	0	1168	0	5	16	16	38	786	0	0	153	233	27	0	0	0	0	41	0
I	380	82	459	437	16	129	280	2	0	0	50	567	297	2497	893	0	8	342	1135	1194	0	0	11	23	2	79
J	1259	0	0	0	1818	0	0	0	350	0	0	0	0	0	3147	0	0	70	0	0	3356	0	0	0	0	0
K	395	28	0	0	5283	28	198	0	1582	0	0	198	28	198	0	0	0	85	28	28	0	0	0	0	113	0
L	1342	19	28	736	1918	105	0	0	1521	0	79	1413	82	4	778	41	0	34	389	254	269	0	11	0	819	0
M	1823	337	26	0	2976	10	0	0	1345	0	0	10	654	42	1246	722	0	26	244	5	337	5	11	0	192	0
N	550	4	621	1681	1212	102	1391	665	665	9	66	104	706	194	528	4	7	11	751	1641	124	68	18	2	157	4
O	85	101	162	231	37	82	25	92	416	14	78	416	706	2191	222	292	0	1531	357	396	947	334	345	12	41	4
P	1358	0	6	0	1747	0	0	237	812	0	0	812	73	6	1511	581	0	2305	180	287	457	0	0	0	17	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10000	0	0	0	0	0
R	1026	33	282	17	2795	31	175	17	1181	0	205	164	303	55	1114	0	0	212	655	192	142	17	0	306	306	0
S	604	12	284	27	1795	24	0	561	1177	0	91	145	112	706	386	0	9	27	836	2484	579	0	39	0	81	0
T	619	3	36	2	1417	7	0	3511	1406	0	0	101	44	15	1228	3	0	479	418	213	195	5	88	0	203	5
U	344	415	491	243	434	52	382	10	258	0	14	1097	329	1518	19	386	0	1460	1221	1255	29	14	0	10	14	5
V	749	0	23	0	6013	0	0	0	2568	0	0	0	12	0	530	0	0	23	0	0	12	12	0	0	58	0
W	2290	8	0	32	1942	0	0	1422	2104	0	0	0	0	357	1292	0	0	106	366	16	0	0	0	0	24	0
X	672	0	1119	0	1269	0	0	0	0	0	0	0	75	0	75	3507	0	0	0	1716	0	0	0	373	0	0
Y	586	34	103	69	2898	0	0	75	172	34	0	379	75	172	2208	310	0	310	0	172	138	0	103	0	69	34
Z	2278	0	0	0	4557	0	0	2152	691	0	127	127	0	506	506	0	0	0	0	0	127	0	0	0	0	253

To compute this table a typical example of english text was chosen. Then the first row of the table was obtained by recording, for each of 10.000 occurrences of the letter A, the letter that immediately followed it. Thus the entry 469 in the column indexed by D means that in this sample of 10.000 occurrences of the letter A the letter D was observed to immediately follow A exatly 469 times. The same procedure was repeated for each of the letters of the alphabet. We see that in 10.000 occurrences of Q, the letter U followed it all 10.000 times (not surprisingl!).

To get the conditional probability P[next letter=Y | prec lett = X] you simply look at the entry in row X and column Y and divide by 10.000.

Thus

P[next letter=E | preceeding letter = R] = 2795/10.000= .27

Some Bi-letter English Monkey Words

A monkey types on a keyboard, the first letter appears with frequency in an English word, the second through fourth letters are chosen with a probability using a biletter frequency table given that the previous letter has been chosen.

```
thint oucon dingi these ryere veega manen elers dengi hedda hesnc rmiss isonk
ticic utasu thede thore ighab benan dilet sinya nthea rived there xched tenor
omobo diome emeth esast nyazz nscec rikir asthe everr blari disul hengi rinea
denth dmewo lralt rapol ldbre dites menti hesce osece oriom thear lloas ponoo
thesi ntiai beare veaki seere awhes suert theco ldent lenda sizal ingai irove
searm ereye atere antrn sunga endea plthe overu amese cyebl nynds hsere rtopl
pathe oares rstht stsut fennt enthe vangh ounen asile ninti xpess thegh speni
cigio tsore rendd halla erise siath erere sstra stsou theta ffant ingam onsty
reria costi denoa ezogr eewar gicau indma erofe tosth hetye teror uleem nicad
checo presa ilila derer fontr lasst ibuta atsth lirru maled doror expus grahe
nlsew rlune unche tourk shedr idens istis cenge sofic ruthe errgh rerga nftst
odata taler anthe renim lymer eulld cende ngsta piopo ctour sholl ndowa sengs
didos ndyin nthed toffi segha aveal hesee engom nenhi lesiv rksti susel wante
ofrra enero edemi mpast armop herel evede theys theco epotr lempi nsted rsper
omema ngeat satin pulli orunt erang peese ntiso istit ended resed rteem lldas
cathe evest nssll raicu entth errys ceeca theny tiong hecle seske inogh tlufo
pstav itofo giems feden eeaya erous cousc ikeng sulde scers derth renio penth
stuan shend calic whiso ndics eroro mprer spapo erint tluge tithe narda orkre
ecrad eromi ordst erofi iomim atofe trmin whall famer xcost stale sldev edowa
ourio dedut enerm narot nvily ndeys ndise sanmm egati ychis thamb dansh lomem
rmoni praut sesiv oumut erthe ethod enepa erkan poida lentl rorch fresi thess
inthe ctans ppomi halye owero sitye oneth oryei oulor whali deist eathe inicy
esfon gesti rleme ndesw wheit olalo eedit titha tanat isona haned ngrer tonmo
tryst ulyor tewed conce herrd lexar isoma savio dyent ainow senor cirsc utest
finin acewe weryo redee sofor nstta sousi kstco emese delin hedle demev ppain
nyesh sithe eatea witue ededi eathe wicel nther kesin latex foken ecoup londi
thofo esive arkeg ngrer deply anthe sermo peour rontl ayerc redul ncoti ullor
overt ttror tipom abyri aslld orent eyoch mersa heaye ameat ronef sedei utidc
usper rares issth ysand amast nicec ereth ithan stson aysoi damia nsais shani
dsoth trket sstrb rolly onthe retec egron stomy thena ooiti rvesi esagh putro
```

Some Tri-letter English Monkey Words

A monkey types on a keyboard, the first letter appears with frequency in an English word, the second letter is chosen with a probability using a biletter frequency table given that the first letter has been chosen, the third through fifth letters are chosen using a triletter frequency table given that the previous two letters have already been chosen.

```
dafcf atoho eledc oldeb osaye illjp ussya insem icing gooke ititi abook quesi
andes thind teonf rsett theof eacbc lalvi ningt oresp egyho dimet iblen seean
lingw seass putiw regrp diero demar ncrar scorr nadfd serem herse thasc amrmo
xtran eenki bysit ontau aygrp newha wevea texas endge instc eseen simet reing
ngalh ddeaf reefa altre renit odewa ayiss hilai rasub vdfae mbnou nconb relev
nthep xphis escmu preal ddows gauda hstor fuysu froco raway ionsa orxuy orutt
ereis cifyi emben prabs thell liame untal thepr ffaec overt repan areha escof
iafcc eusta thema ecukc nftha redit andpa butio edefe eveep hapto astia thshe
ommex ecaaa theed ntoth atusu hefer comet eveun rside eetth ncomb rtuda bafro
thict ofthe eprie estra plyst hatmy ypreo shrma olsom mmikm esdan fomme eandr
fortu eedth lthst sarav esloo ndren aysub yafbo mency didfe theha cofao quica
ethin martf wnfef tilli thing iaand neice ffanj etsin rdetm ouwro ament quesc
puton orhic enced merni thats ptrix opedu nsfun thisp ouion eensh aunto tassb
ngooc stred ngwil aysul andec trike othis hardi ltlip eduas parik ichop pandr
thers adlim renes udans itsro ionsw pardr viden eebmy conth oasea inmnm lasub
ndefo redek hopha stati solea uitha fiewh maitb egeto dupfq outeb quitr edheg
ensun atows nanai thaps cafaf nthel iemer sepif enway olbyt tenok ontow thsta
ester erpos eckim omsag justu avole ngfov raver yinda thess revie estca matto
rictu frosa iaots neisa thiso rould rewro sitze lyous fingo ercec nginm foesw
ingth yiweh inkin toant romes wando melaa ldber escon nmike exhav bedto ntake
orevu necfc fffca ugooc areyo itake hethe ryork neeha ethin suatw plewi tocki
metpo eryel gnslo ivesl isisy iarts tecfc vetry ofthn alged datna emaya itsta
thego lintu ition lyssi ntemu ticei redat urvex rnary imeno ntwin ellie atiom
eendd ernee ytomp ienre siloc leget asfor yelll iendt ypros rtmor usthe ionee
stalb rtwas negat posar ifwco ngssi youta ricss titha ojulh shrou ymike uened
trepu bymmu tohen nsfae ctthi montd andwi ranto tardt corfr ervar esbet acour
iting ectqu lphat chnin mikei wlips nffpr ngead tyouw fozth hingi ourli ertom
enopo niing rsmei educa oters yetty titma sthuf sapet grati helli septo enwhe
enede edtei setwh thexa ngoin ysess ermst temag rtomm nause daryo urdin ngsfd
```

2. Suppose we have a computer program which generates words from the alphabet A,B,C,D according to the following procedure:

Pick the first letter according to the single frequency table given below then constructs each additional letter using the table of conditional biletter frequencies given below.

a) Calculate the probability that the program produces the word "DACB"
b) Determine the 2 letter word that has the highest probability.

Single letter table

table	
A	10
B	9
C	12
D	9

Biletter table

	A	B	C	D
A	0	3	1	0
B	2	1	4	0
C	0	2	0	3
D	2	1	0	1