CISS451: Cryptography and Computer Security
Substitution examples (old Assignment 3)

The goal is to decrypt the following ciphertext (i.e., you need to compute the plaintext) and also to discover the key used. The substitution cipher is used.

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewm
hmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehq
ermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtx
lajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatlj
mwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqq
evjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlh
jggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamt
liammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqt
xltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmloji
akexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvat
kkmxnmh
```

SOLUTION.

The top few 1–gram frequencies of the ciphertext are

```
1gram: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30 ...
```

The gap between the frequency of `m` and `t` is extremely large. Therefore we suspect that part of the encryption is `e->m`. The rest, at least up to `x` are most probably from `t, a, o, i, n, s, h, r`:

$$\{\texttt{t, a, o, i, n, s, h, r}\} \to \{\texttt{t, q, h, j, e, i, a, x}\}$$

We can try different possible assignments on the above 8 letters to 8 letters, but that's $8! = 40320$ which is too big. At this point we have

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
```

```
-e--------------------e--------------e-----------e-e----------------------------e--e-e-----e--------

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
--------e-e------e-----e---e-e----------e---e-------e-----------e-------------e----e------e---------

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
-----------------e-----------e---e-------------e----------------e-e--e----e-----e---------e----e----

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
--ee-e--ee------------e---------e-----------------e-e--e----e----------------e-----e---e--ee----e---

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
---e--e-----------e---e-----e---e----e-----------------e--------------------e----e---------e----e---

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
------------------e---------------e----e----e------------------e------e----e--------e-----------e-e-

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
-------------------------------e----------e---e-e----ee-e--e-e---e-------e---e-------------e----e---

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
e--------------e------e--------e---e----------e--------e-----e-----e--e-e--e---e-------------------e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
--e-----e--e----e---e------e---------------e-e---------------------------e--------------------------

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e---e-------------e---------e--------------e--------e--e-

ciphertext
2-grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
         tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3-grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
         feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
plaintext
1-grams: e t a o i n s h r
2-grams: th he in er an re ed on es st en at to nt ha nd
         ou ea ng as or ti is et it ar te se hi of
3-grams: the ing and her ere ent tha nth was eth for dth
```

We now look at 2-grams and 3-grams.

The common 2-grams are `th, he, in, er, an, re, ed, on, es, st`. Since we are assuming `e->m`, `mh` is either from `er` or `ed` or `es`. Note that `er` and `re` are common 2-grams. We also note that `mh` and `hm` are high frequency 2-grams in the ciphertext. Note further that `ere` is a common 3-gram and `mhm` is also a common 3-gram in the ciphertext. We suspect that `h->r`. Therefore we now have

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
-e------r-----------r-e------r-------e-----------ere---------------------r------e--ere-----e------r-

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
--------e-e------e-r---e---e-e---------re--re---r---e-------r---e--------r----er---er-----er--------

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
-----------------e-----------e--re-------------e---------------re-e--e----e-----e--r------e--r-e----

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
--ee-er-ee------------e--------re-----------------e-er-e---re----------------e-----e---er-ee---re---

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
```

```
--re--e-----------er--e-----e---e----e-r---------------e----------r---------e--r-er---r----e----er--

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
------------------e---------------e----e----e----r-------------er-----e----e---r----e-----r-----ere-

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
-------r------------r----------er---------e--re-e----ee-e--ere---e-r-----er-re-------------er---e-r-

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er--r----------er-----er-------e--re----------er-------e-----e---rre--ere--e---er--r---------------e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
-re-----e--e----e---e--r---e--r------------ere-----------------r---------e-------r--r---r-----------

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e---e------r-----re---------e----r-----r--e---r-----e--er

ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h
```

Now we look for `the`. We have `the -> ??m`. The most commonly occurring ciphertext of this form is `iam`, `ewm`, `mhm`, `twm`. `mhm` is from `ere` which we already know so this is useless. So we are left with `iam`, `ewm`, `twm`. The frequency between `iam` and `ewm` is a huge 30% drop. So hopefully `the -> iam`. This means `t->i` and `h->a`. We have to take note of this since here we are creating two substitutions and the confidence is not as high. If we end up in a deadend, we will have to backtrack to this point. With the above two new substitutions, we have

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
-e--t---r-t-t-------r-e------r---t---e----------here---t-----------------r------e-tere-----e-t----r-

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
---t----e-e------e-r---e---e-e---------re--re---r---et-----tr---e--t--t--r---her---er-----er------t-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
--t----t---------e-----t-----e--re-t-----------e---------------re-e--e----e-----e--r----t-e--r-e-t--

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
-heeher-ee------------e--------ret--t---t----t----eherhe-t-re-t-----t-------he-----e--her-ee-t-re---

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
--re-te-----------er--e---t-e---e----e-r--------------te--t-------r-------t-e--r-er---r--t-e--tter--

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
------------tt---he---------------e----e----e----r------------her-----e-tt-e---r----e-----r----tere-

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
-------r--t---------rh--------her--------he--re-e--t-ee-e--ere---e-r-----er-re-t-----t--t-her---e-rh

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
```

```
er--r-t-------her----ter-------e--re----------er-------e-t---e---rre--eret-e-t-er--rt--------------e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
-re----het-et---e--he--rt--e--r-----------here---t-------------rt--------e-------r--r---r-------h---

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e---e------r-t---re-------t-e--t-r-----r--e---r-----e--er
ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h, t->i, h->a
```

`ent` is also a common plaintext trigram. This is encrypted as `m?i`. The only one that fits is `mqi` but the frequency of this is only 4 – so this is probably wrong.

Another high frequency plaintext 3-gram is `tha`. This would encrypt as `ia?`. We notice that `iam` has a high frequency. So perhaps `a->m`.

Now let's look at pairs of digrams.

`es,st` is a high frequency digram. This is encrypted as `m?,?i`. The only possibility is `mq,qi`. So we suspect `s->q`. This is what we have now:

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hes-t---r-t-t----h--r-eh-----r---t---e-h--h--s---ere---th----s---s-------r--s-s-e-tere-h-sse-ts---r-

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
---t----e-eh-s---e-r---e---e-e--s----s-re--re---rs--eth---str---e--t--t--rs---er---er-----ers-----t-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
-sth-s-t--sh-ss--e-h--ht-----es-re-th----h-ss--e---h-s----s----rese--eh-she----se--r---sthe--r-est-h

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
--ee-ersee------h-----e------s-reth-th-st--h-th---e-er-e-t-re-t-----t-----h--es----e---er-ee-t-re---

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
--reste-----sh----ersheh--the---e----e-r--h--h--ss----te--th---ss-r-------the--r-er---r--the--tters-

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
------s----stt-h--e-s-s-----------es---esh--e----r-------------erh-s-hesttheh--r--s-e-----r---stere-

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
-------r--t---------r------s---erh-s--ss--e--rehe--thee-es-ere---e-r-----er-re-t-----t--ts-er---e-r-

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er--r-t--------er---ster----h--es-re----sh----ers-----hest---e---rre--erethe-ther--rts--h---h--h---e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
-re-----ethet---es--e--rt--e--r---sh---s---ere---th---------h--rth-s-----e-------r--r---r-----------
```

```
mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e---e------rst---ress-----the--t-r---s-r--ess-r-h---e--er

ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
        plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h, t->i, h->a, s->q
```

Note that `ti,is` is a common plaintext 2-gram. When encrypted, this is `i?,?q`. Unfortunately we can't find this pattern.

We now have enough substitutions to consider multiple cases of pairs of digrams.

Consider the common plaintext digram `aj,jq`. With what we have at this point, the encryption is `h?,?s -> aj,jq`. The possibilities for `h?,?s` are

- `he,es`: Therefore `e->j`, but `e` is already encrypted as `m`.
- `ha,as`: Therefore `a->j`.
- `hi,is`: Therefore `i->j`.

So we have `a->j` or `i->j`. Before we make a choice, let's consider more digrams.

Consider `h?,e?`. `h?,e?` might be encrypted as `at,mt`. Possibilities for `h?,e?`

- `hi,ei`: But `ei` is not common.
- `ha,ea`: Therefore `a->t`.

Therefore `a->t`.

Consider `h?,?r`. `h?,?r` might be encrypted to `at,th`. The only possibilities for `h?,?r` are

- `ha,ar`: Therefore `a->t`.
- `hi,ir`: But `ir` is not common.

Therefore `a->t`.

Consider `?s,e?`. `?s,e?` might be encrypted as `tq,mt`. The only possibility for `?s,e?` is `as,ea` which implies `a->t`.

`?s,e?` might be encrypted as `tq,th`. The only possibility for `?s,e?` is `as,ea` which impplies `a->t`.

`e?,?r` might be encrypted to `mt,th`. The possibilities for `e?,?r` are

- `ed,rd`: But `rd` is not common.
- `es,sr`: But `sr` and `os` not common.
- `en,nr`: But `nr` is not common.
- `ea,ar`: Therefore `a->t`.
- `et,tr`: But `tr` is not common.

All in all, this case implies `a->t`.

From all the above cases, it seems that `a->t` and `i->j`. (The argument for `a->t` is stronger.) We now have

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesati-ar-tati---hair-ehi--a-r-a-ta--e-hi-h-as---ere--ith----s-a-sa---ia-ra-sasie-tere-hisseats---r-

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
---t--a-e-ehisa--eara--e-a-e-e-as-i-as-re-are---rs--ethi--stra--e--t--t--rs---er---eri--a-ers--a-it-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisit-ashissi-e-hi-ht-----es-reatha-a-hissi-ea--hisi---si---rese--ehishea--ase--r---sthe-ar-estih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
a-ee-ersee-----ah--a--ei--ia-s-rethathist--hatha-ie-er-e-t-re-t----it-----ha-es-i--e---er-ee-tire--a

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
--reste-----sh----ersheha-the-a-ea---ear--hi-hiass--iate-itha-ass-ria-----the--r-er---ri-the-atters-

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
--a--asa---stt-ha-eas-s-i-i-------es-a-esha-e-a--ri---i--------erhis-hestthehair-as-e---iar--astere-

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
----i--r--ti-a------r-i---is---erhis-assi-e--rehea-thee-es-ere---e-ra----er-reat--a--t--ts-er---ear-

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er--riti-a-a---er--aster---ah--es-rea---sh----ersa--a-hest-i-ea-arre--erethe-ther-arts--hi--hi-ha--e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
are-a---etheta--esa-e--rt--e--r---sha--s---ere--ith------a--hairthisa--a-e----i--r-ari--r----i----i-

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e-a-e-----irsti--ressi----the--t-ri--s-r--ess-r-ha--e--er

ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
        plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h, t->i, h->a, s->q, a->t, i->j
```

At this point we can already see (possibly) “`he sat`” at line one and “`that his`” at line 4 and “`his -hest the hair`” at line 4 – perhaps “`chest`” is the second word?

Next, we try `tx,jx`. `a?,i?` might be encrypted as `tx,jx`. The possibilities for `a?,i?` are

- `an,in`: Therefore `n->x`.
- `ar,ir`: But `ir` is not common.

We get

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesatinar-tatin--hair-ehin-a-r-a-ta--e-hi-h-as---ere--ith----s-a-san--ia-ra-sasientere-hisseats--nr-

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
-n-t--a-e-ehisa--earan-e-a-e-e-as-i-as-re-are---rs--ethin-stran-e--tn-t--rs---er---erin-a-ers-na-it-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisit-ashissi-e-hi-ht----nes-reatha-a-hissi-ean-hisi---sin--resen-ehishea--asen-r---sthe-ar-estih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
a-ee-erseen---nah--an-ein-ia-s-rethathist--hatha-ie-er-ent-re-t---nit-----ha-es-i--e---er-eentire--a

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
n-reste--n--sh----ersheha-the-a-ean--ear--hi-hiass--iate-ithanass-rian----the--r-er---ri-the-atters-

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
--a--asa---stt-ha-eas-s-i-i-n-----es-a-esha-e-an-ri---in----n--erhis-hestthehair-as-e---iar--astere-

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
---nin-r-ntina--n---r-in--is---erhis-assi-e--rehea-thee-es-ere---e-ra--n-er-reat--a--t--ts-er---ear-

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er--riti-a-an--er--aster---ah--es-rea---sh----ersan-a-hest-i-ea-arre--erethe-ther-arts--hi--hi-ha--e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
are-a---etheta--esa-e--rt--en-r---shan-s---ere--ith--n---a--hairthisan-a-e----in-r-arin-r----in---i-

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e-a-e-----irsti--ressi-n--then-t-ri--s-r--ess-r-ha--en-er

ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
        plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h, t->i, h->a, s->q, a->t, i->j, n->x
```

The beginning of the plaintext now reads “`he sat in a`.” We now look at `xl,ml` and `ex,eh`.

`n?,e?` might be encrypted as `xl,ml`. The possibilities for `n?,e?` are

- `nt,et`: But `t` is already assigned.
- `nd,ed`: Therefore implies `d->l`.
- `ng,eg`: But `eg` is not common.

`?n,?r` might be encrypted as `ex,eh`. The possibilities for `?n,?r` are

- `in,ir`: But `i` is already assigned.
- `an,ar`: But `a` is already assigned.
- `on,or`: This implies `o->e`.
- `en,er`: But `e` is already assigned.

Adding `d->l` and `o->e`, we get

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesatinarotatin--hair-ehinda-roadta--e-hi-h-as-o-ered-ith-oo-s-a-sanddia-ra-sasienteredhisseats--nro

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
-ndto-a-e-ehisa--earan-e-ade-e-as-i-as-re-ared-orso-ethin-stran-e--tnot-orsoo-er-o-erin-a-ersona-it-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisit-ashissi-e-hi-htoo-ones-reatha-a-hissi-eandhisi--osin--resen-ehishead-asenor-o-sthe-ar-estih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
a-ee-erseen--onah--an-ein-ia-s-rethathisto-hathadie-er-ent-redtodonit-o--dha-es-i--edo-er-eentire--a

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
ndrestedon--sho--dershehadthe-a-eand-eard-hi-hiasso-iate-ithanass-rian----the-or-er--oridthe-atterso

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
--a--asa--osttoha-eas-s-i-iono----es-adesha-edandri---in-do-no-erhis-hestthehair-as-e---iar--astered

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
do-nin-rontina-on---r-in--is-o-erhis-assi-e-oreheadthee-es-ere---e-ra--nder-reat--a--t--ts-er---ear-

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er--riti-a-and-er--aster---ah--es-reado-sho--dersanda-hest-i-ea-arre--eretheother-artso-hi--hi-ha--e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
areda-o-etheta--esa-e-ort-oenor-o-shands-o-ered-ith-on---a--hairthisanda-e--o-in-roarin-r----in--oi-

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e-ade-----irsti--ressiono-thenotorio-s-ro-essor-ha--en-er

ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
        plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h, t->i, h->a, s->q, a->t, i->j, n->x, d->l, o->e
```

The beginning reads "`he sat in a rotatin--hair-ehinda-road...`" which is very likely "`he sat in a rotating chair-ehinda-road...`", giving us `g->n` and `c->v`. This gives us

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesatinarotatingchair-ehinda-roadta--e-hich-asco-ered-ith-oo-s-a-sanddiagra-sasienteredhisseats--nro

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
-ndto-ace-ehisa--earance-ade-egas-i-as-re-ared-orso-ethingstrange--tnot-orsoo-er-o-eringa-ersona-it-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisit-ashissi-e-hichtoo-ones-reatha-a-hissi-eandhisi--osing-resencehishead-asenor-o-sthe-argestih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
a-ee-erseen--onah--an-eingia-s-rethathisto-hathadie-er-ent-redtodonit-o--dha-es-i--edo-er-eentire--a

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
ndrestedon--sho--dershehadthe-aceand-eard-hichiassociate-ithanass-rian----the-or-er--oridthe-atterso

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
--ac-asa--osttoha-eas-s-iciono----es-adesha-edandri---ingdo-no-erhischestthehair-as-ec--iar--astered

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
do-nin-rontina-ongc-r-ing-is-o-erhis-assi-e-oreheadthee-es-ere---egra--ndergreat--ac-t--ts-er-c-ear-

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er-critica-and-er--aster---ah-ges-reado-sho--dersandachest-i-ea-arre--eretheother-artso-hi--hicha--e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
areda-o-etheta--esa-e-ort-oenor-o-shandsco-ered-ith-ong--ac-hairthisanda-e--o-ingroaringr----ing-oic

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
e-ade-----irsti--ressiono-thenotorio-s-ro-essorcha--enger

ciphertext
1grams: m:116 t:95 q:75 h:74 j:68 e:65 i:59 a:59 x:53 l:40 k:34 g:30
2grams: mh:25 ia:20 hm:20 aj:19 wm:18 at:16 am:16 tq:15 mq:15 xn:14
        tx:14 mt:14 jx:14 xl:12 qi:12 jq:12 th:11 ml:10 ex:10 eh:10
3grams: wmh:11 ajq:11 jxn:10 iam:10 txl:9 hml:7 ewm:7 mhm:6 otq:5 mhq:5
        feh: 5 twm: 4 qqj: 4 oaj: 4 mth:4 mqi:4 mia:4 jva:4 imh:4 iat:4
        plaintext
1grams: e t a o i n s h r
2grams: th he in er an re ed on es st en at to nt ha nd
        ou ea ng as or ti is et it ar te se hi of
3grams: the ing and her ere ent tha nth was eth for dth

e->m, r->h, t->i, h->a, s->q, a->t, i->j, n->x, d->l, o->e, g->n, c->v
```

Near the middle of the first line, "`-hich-asco-ered-ith`" is probably "`which-asco-eredwith`" giving us `w->o`.

On the second line "`orso-ethingstrange-`" is probably "`or something strange-`", giving us `m->r`.

On the third line "`sthe-argestiha-ee-erseen`" is probably "`s the largest i have ever seen`." This gives us `l->k` and `v->w`.

At this point we have

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesatinarotatingchair-ehinda-roadta-lewhichwascoveredwith-oo-sma-sanddiagramsasienteredhisseats--nro

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
-ndto-acemehisa--earancemademegas-iwas-re-ared-orsomethingstrange--tnot-orsoover-oweringa-ersonalit-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisitwashissi-ewhichtoo-ones-reathawa-hissi-eandhisim-osing-resencehisheadwasenormo-sthelargestih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
aveeverseen--onah-man-eingiams-rethathisto-hathadievervent-redtodonitwo-ldhavesli--edovermeentirel-a

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
ndrestedonm-sho-ldershehadthe-aceand-eardwhichiassociatewithanass-rian--llthe-ormer-loridthelatterso

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
-lac-asalmosttohaveas-s-iciono--l-es-adesha-edandri--lingdownoverhischestthehairwas-ec-liar-lastered

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
downin-rontinalongc-rvingwis-overhismassive-oreheadthee-eswere-l-egra--ndergreat-lac-t--tsver-clearv

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er-criticalandver-master--lah-ges-reado-sho-ldersandachestli-ea-arrelweretheother-artso-himwhicha--e

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
areda-ovetheta-lesave-ortwoenormo-shandscoveredwithlong-lac-hairthisanda-ellowingroaringr-m-lingvoic

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
emade--m--irstim-ressiono-thenotorio-s-ro-essorchallenger

e->m, r->h, t->i, h->a, s->q, a->t, i->j, n->x, d->l, o->e, g->n, c->v, w->o, m->r,
l->k, v->w
```

At the third line "`hisim-osing-resencehisheadwasenormo-sthelargestihaveeverseen`" is probably "`his imposing presence his head was enormous the largest i have ever seen`" giving us `p->g` and `u->b`.

At line 7, "`hismassive-orehead`" is "`his massive forehead`" giving us `f->f`.

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesatinarotatingchair-ehinda-roadta-lewhichwascoveredwith-oo-smapsanddiagramsasienteredhisseatspunro

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
undtofacemehisappearancemademegaspiwaspreparedforsomethingstrange-utnotforsooverpoweringapersonalit-

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisitwashissi-ewhichtoo-ones-reathawa-hissi-eandhisimposingpresencehisheadwasenormousthelargestih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
aveeverseenuponahuman-eingiamsurethathistophathadieverventuredtodonitwouldhaveslippedovermeentirel-a

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
ndrestedonm-shouldershehadthefaceand-eardwhichiassociatewithanass-rian-ulltheformerfloridthelatterso

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
-lac-asalmosttohaveasuspicionof-luespadeshapedandripplingdownoverhischestthehairwaspeculiarplastered

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
downinfrontinalongcurvingwispoverhismassiveforeheadthee-eswere-luegra-undergreat-lac-tuftsver-clearv
```

```
mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
er-criticalandver-masterfulahugespreadofshouldersandachestli-ea-arrelweretheotherpartsofhimwhichappe

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
areda-ovetheta-lesavefortwoenormoushandscoveredwithlong-lac-hairthisanda-ellowingroaringrum-lingvoic

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
emadeupm-firstimpressionofthenotoriousprofessorchallenger

e->m, r->h, t->i, h->a, s->q, a->t, i->j, n->x, d->l, o->e, g->n, c->v, w->o, m->r,
l->k, v->w, p->g, u->b, f->f
```

The beginning “`hesatinarotatingchair-ehinda-roadta-le`” is “`he sat in a rotating chair behind abroad table`” giving us `b->u`.

At line 5, “`restedonm-shoulder`” is “`rested on my shoulder`” giving us `y->s`.

At line 8, “`shouldersandachestli-ea-arrel`” is “`shoulders and a chest like a barrel`, giving us `k->c`.

At line 3, “`itwashissi-ewhichtookonesbreathaway`” is “`it was his size which took ones breath away`” giving us `z->p`.

We now have

```
amqtijxtheitijxnvatjhumajxltuhetlitukmoajvaotqvewmhmlojiaueecqrtgqtxlljtnhtrqtqjmximhmlajqqmtiqgbxhe
hesatinarotatingchairbehindabroadtablewhichwascoveredwithbooksmapsanddiagramsasienteredhisseatspunro

bxlieftvmrmajqtggmthtxvmrtlmrmntqgjotqghmgthmlfehqermiajxnqihtxnmubixeifehqeewmhgeomhjxntgmhqextkjis
undtofacemehisappearancemademegaspiwaspreparedforsomethingstrangebutnotforsooverpoweringapersonality

tqiajqjiotqajqqjpmoajvaieecexmquhmtiatotsajqqjpmtxlajqjrgeqjxnghmqmxvmajqamtlotqmxehrebqiamkthnmqija
asthisitwashissizewhichtookonesbreathawayhissizeandhisimposingpresencehisheadwasenormousthelargestih

twmmwmhqmmxbgextabrtxumjxnjtrqbhmiatiajqiegatiatljmwmhwmxibhmlielexjioebklatwmqkjggmlewmhrmmxijhmkst
aveeverseenuponahumanbeingiamsurethathistophathadieverventuredtodonitwouldhaveslippedovermeentirelya

xlhmqimlexrsqaebklmhqamatliamftvmtxlumthloajvajtqqevjtimojiatxtqqshjtxubkkiamfehrmhfkehjliamktiimhqe
ndrestedonmyshouldershehadthefaceandbeardwhichiassociatewithanassyrianbulltheformerfloridthelatterso

uktvctqtkreqiieatwmtqbqgjvjexefukbmqgtlmqatgmltxlhjggkjxnleoxewmhajqvamqiiamatjhotqgmvbkjthgktqimhml
blackasalmosttohaveasuspicionofbluespadeshapedandripplingdownoverhischestthehairwaspeculiarplastered

leoxjxfhexijxtkexnvbhwjxnojqgewmhajqrtqqjwmfehmamtliammsmqomhmukbmnhtsbxlmhnhmtiuktvcibfiqwmhsvkmthw
downinfrontinalongcurvingwispoverhismassiveforeheadtheeyeswerebluegrayundergreatblacktuftsveryclearv

mhsvhjijvtktxlwmhsrtqimhfbktabnmqghmtlefqaebklmhqtxltvamqikjcmtuthhmkomhmiameiamhgthiqefajroajvatggm
erycriticalandverymasterfulahugespreadofshouldersandachestlikeabarrelweretheotherpartsofhimwhichappe

thmltuewmiamitukmqtwmfehioemxehrebqatxlqvewmhmlojiakexnuktvcatjhiajqtxltumkkeojxnhethjxnhbrukjxnwejv
aredabovethetablesavefortwoenormoushandscoveredwithlongblackhairthisandabellowingroaringrumblingvoic

mrtlmbgrsfjhqijrghmqqjexefiamxeiehjebqghefmqqehvatkkmxnmh
emadeupmyfirstimpressionofthenotoriousprofessorchallenger

e->m, r->h, t->i, h->a, s->q, a->t, i->j, n->x, d->l, o->e, g->n, c->v, w->o, m->r,
l->k, v->w, p->g, u->b, f->f, b->b, y->s, k->c, z->p
```

```
q->q, x->x, j->j
```

Note that `q,x,j` were not used in the plaintext. We have added `q->q, x->x, j->j` to the substitution key. The following is the plaintext with spaces inserted (puncuations not restored):

> he sat in a rotating chair behind a broad table which was covered with books maps and diagrams as i entered his seat spun round to face me his appearance made me gasp i was prepared for something strange but not for so overpowering a personality as this it was his size which took ones breath away his size and his imposing presence his head was enormous the largest i have ever seen upon a human being i am sure that his tophat had i ever ventured to don it would have slipped over me entirely and rested on my shoulders he had the face and beard which i associate with an assyrian bull the former florid the latter so black as almost to have a suspicion of blue spade shaped and rippling down over his chest the hair was peculiar plastered down in front in a long curving wisp over his massive forehead the eyes were blue gray under great black tufts very clear very critical and very masterful a huge spread of shoulders and a chest like a barrel were the other parts of him which appeared above the table save for two enormous hands covered with long black hair this and a bellowing roaring rumbling voice made up my first impression of the notorious professor challenger

□

The following programs are helpful:

1. Code to print the top 1-grams, 2-grams, 3-grams. The 1-grams will help determine the character that `e` is encrypted to. The trigrams might help determine what `the` is encrypted to.
2. Given a character `c`, code that computes character(s) `d` such that `cd` and `dc` occurs most frequently. This will help determine what `r` is encrypted to, based on the fact that the character `x` that appears most commonly before and after `e` is `r`.
3. Given a collection of common 2-grams (in plaintext), a partially specified substitution, compute pairs of commonly ocurring digrams of the form `xy,xz` or `xy,zx` or `xy,yz` or `xy,zy` (i.e., there are three distinct characters in the pairs of digrams) where two of the characters have already been decrypted and the remaining one has not and has not been assigned to a plaintext character.
4. Instead of the above where there are two digrams, listing 4-grams where the decryption of 3 are known and one is unknown is also useful.

Here are some exercises for you:

Exercise 0.1. Find the key and the plaintext of the following ciphertext encrypted by teh substitution cipher:

```
psxdltuxtcwauuxvifgtzwacsppstpcvxqtdgcvxqtdgrvxbxd
pwzxdgxgadtdipsxvtvlspstpwiijxgihpinxvtblxdxbiwajx
tralphvxihpifteiijteihpaptwzpstpxnxvzidxbevxtpsqtb
ptjxdtqtztdgpsxzbaurwzqtwjxgfivqtvgbawxdptdgbptvad
ctbsivptnxdhxiflzrvxbbxbwxgqagxdadctbapqxdppitutve
wxpxvvtlxpstpwtzevitgtdgqsapxadpsxbhdwacsppsxlsawg
vxdewadjadcwxtdxgpsxavtvubidpsxevitgfwtpetwhbpvtgx
tdgctmxgauuxgatpxwzexwiqpsxuqtbtwtjxyhbpwajxtwtjxa
dpsxexthpaxbifaptwztwtjxqapsbqtdbtdgtdabwtdgtdgqxx
radcqawwiqbexzidgapqxvxcvxxdbwirxbgippxgqapscvinxb
ifpvxxbtdgtuagpsxpvxxbcwxtuxgpsxqsapxwauebifbptphx
btctadbptwappwxsawwpipsxwxfpqtbtvihdgqsapxehawgadc
qapsrawwtvbtdgpipsxvacsptqtpxvftwwltuxphuewadcgiqd
tuidcuibbzbpidxbpibrwtbsadpipsxwtjxbpxrbfxgfviupsx
pxvvtlxpipsxqtpxvtdgipsxvbpxrbpipsxcvxxdwtqdbexbag
xaptqtztlvibbpsxcvtbbzbwirxbgxxvqxvxfxxgadctdgadps
xgabptdlxqsxvxpsxcvinxbifpvxxbpsaljxdxgadpiqstpwii
jxgtwuibptfivxbpqxvxxdivuihbbstrxbifcvxzbpidxwajxd
ipsadcpstppsxlsawgvxdstgxnxvbxxdexfivx
```

Exercise 0.2. Here's another one that is harder:

```
oznftyomrrtqlnlzftqlqlmxemftlnyoozjtzyvqlzfzvgmnzm
drsjmrrlxtqlxyjtynmvxqlhlmnfjmkmvxpyhvlglvhqlvdznj
qzvpemrlomjtynfhqzjqqlxylfzvtlnezttlvtrstqnyipqyit
tqlxmsynmttlvxzvpmjnzjaltemtjqoynmrrhlavyhqlhlmnft
qlezvdlx
```

□

Exercise 0.3. And another:

```
xsdwoddnuskapbgxaayrpcuevbsxjjuskyaqddpdbxbvrolrhb
qsyqmsdnvxumxsnxsxgxfuskbdahjrbbdbbursrhgxssdmnxme
dnxmrqsnevdedsexsnjxqbdnusevdumgunbebrgdcvxevdxedn
cuevvubvxbed
```

□

Once you have solved a few substitution ciphers by hand, you are ready to write a program to automate your process. Instead of a perfect solver, your program should aim to print a list of possible decryptions, ordered by likelihood (by being the original plaintext).

You can learn more about breaking substitution ciphers by doing a google search. Clearly the process of breaking a substitution cipher involves trying substitutions. This leads to search algorithms. Starting with the traditional backtracking search, you will be led to other heuristic search algorithms such as local seach algorithms, genetic algorithms, etc. Studying substitution ciphers will also lead to studies of probability theory and specifically markov chains. A search on google will reveal many research papers (many of which are very recent) on substitution cipher, AI search algorithms, markov chains, etc.