

# Лабораторна робота №3

## Криптоаналіз афінної біграмної підстановки

Виконали:

ФБ-23 Литвин Руслан

ФБ-23 Ващаєв Тимофій

Варіант 1

### Мета роботи

Набуття навичок частотного аналізу на прикладі розкриття моноалфавітної підстановки; опанування прийомами роботи в модулярній арифметиці.

### Порядок виконання роботи

1. Реалізувати підпрограми із необхідними математичними операціями: обчисленням оберненого елементу за модулем із використанням розширеного алгоритму Евкліда, розв'язуванням лінійних порівнянь. При розв'язуванні порівнянь потрібно коректно обробляти випадок із декількома розв'язками, повертаючи їх усі.
2. За допомогою програми обчислення частот біграм, яка написана в ході виконання комп'ютерного практикуму №1, знайти 5 найчастіших біграм запропонованого шифртексту (за варіантом).
3. Перебрати можливі варіанти співставлення частих біграм мови та частих біграм шифртексту (розглядаючи пари біграм із п'яти найчастіших). Для кожного співставлення знайти можливі кандидати на ключ  $(a, b)$  шляхом розв'язання системи (1).
4. Для кожного кандидата на ключ дешифрувати шифртекст. Якщо шифртекст не є змістовним текстом російською мовою, відкинути цього кандидата.
5. Повторювати дії 3-4 доти, доки дешифрований текст не буде змістовним.

### Хід роботи

Виконання першого пункту відбувалося у декілька основних кроків:

1. Обчислення оберненого елементу за модулем із використанням розширеного алгоритму Евкліда
2. Розв'язання лінійних порівнянь

Для обчислення оберненого елементу за модулем  $(a^{-1} \bmod b)$  було реалізовано функцію *invertElement*, яка приймає два значення  $(a \text{ і } b)$  та повертає найбільший спільний дільник (gcd), а також обернений елемент до  $a$  за модулем  $b$ , використовуючи розширений алгоритм Евкліда

```
def invertElement(a, b):
    if a == 0:
        return b, 1, 0

    gcd, prev_u, curr_u = invertElement(b % a, a)
    new_u = prev_u - (b // a) * curr_u

    return gcd, curr_u, new_u
```

Для розв'язання лінійних порівнянь ( $ax = b \bmod m$ ) було реалізовано функцію `solveEquation`, яка знаходить кількість можливих розв'язків заданого порівняння, а також, використовуючи функцію `invertElement`, обраховує та повертає їх усі (або `None`, якщо задане порівняння не має розв'язків).

```
def solveEquation(a, b, m):
    gcd, _, _ = invertElement(a, m)

    if gcd == 1:
        _, _, elem = invertElement(a, m)
        # res = b * (elem % m) % m
        return b * elem % m
    elif gcd > 1 and b % gcd == 0:
        _, _, elem = invertElement(int(a / gcd), int(m / gcd))
        first = int(b / gcd * elem % m)
        return [first + int(m / gcd) * i for i in range(gcd)]
```

Наступним завданням стало розшифрування наданого шифротексту.

Спочатку, використовуючи функцію `findBigrams`, яка була реалізована в ході виконання комп'ютерного практикуму №1, знаходимо 5 найчастіших біграм запропонованого шифротексту, а також 5 біграм російської мови.

Топ 5 біграм шифротексту	Топ 5 біграм російської мови
рн	ст
ьч	но
нк	то
цз	на
иа	ен

Далі перебираємо усі можливі варіанти співставлення значень найчастіших біграм шифротексту з біграмами російської мови.

(545, 509), (545, 860), (545, 413), (545, 689), (545, 248), (417, 509), (417, 860), (417, 413), (417, 689), (417, 248), (572, 509), (572, 860), (572, 413), (572, 689), (572, 248), (403, 509), (403, 860), (403, 413), (403, 689), (403, 248), (168, 509), (168, 860), (168, 413), (168, 689), (168, 248)

Тепер для кожної пари значень знаходимо можливі значення ключа  $a$  та  $b$  для подальшого дешифрування заданого шифротексту.

Параметр  $a$  знаходимо за допомогою формули:

$$Y^* - Y^{**} = a(X^* - X^{**}) \bmod m^2$$

Приклад коду для знаходження параметра  $a$ :

```
a = solveEquation(all_bigrams[i][0] - all_bigrams[j][0] % 31**2,  
all_bigrams[i][1] - all_bigrams[j][1] % 31**2, 31**2)
```

Параметр  $b$  знаходимо за допомогою формули:

$$b = (Y^* - aX^*) \bmod m^2$$

Приклад коду для знаходження параметра  $b$ :

```
b = (all_bigrams[i][1] - a * all_bigrams[i][0]) % 31**2
```

Потім, використовуючи знайденні значення  $a$  та  $b$ , дешифруємо заданий шифротекст за допомогою формули:

$$X_i = a^{-1}(Y_i - b) \bmod m^2$$

Приклад коду для дешифрування заданого шифротексту:

```
def decrypt(text, a, b):  
    _, _, elem_a = invertElement(a, 31**2)  
    line = ""  
    if elem_a is not None:  
        for i in range(0, len(text) - 1, 2):  
            val_y = alph.index(text[i]) * 31 + alph.index(text[i + 1])  
            res_x = (elem_a * (val_y - b)) % 31**2  
            line += alph[res_x // 31] + alph[res_x % 31]  
    print(f"a: {a}, b: {b} -> {line[:30]}...")
```

Наступним кроком стало розпізнавання змістовного тексту російською мовою, для якого було реалізовано розпізнавач з двома методами.

Перший метод ґрунтується на тому факті, що надлишковість природної мови доволі висока, а текст отриманий в результаті неправильно ключа має бути випадковим, а його надлишковість меншою. Тому цей метод обчислює надлишковість кожного розшифрованого тексту і видає той варіант, який має найбільшу надлишковість.

Обраховує надлишковість для кожного ключа, ключ представлений як кортеж, якому у відповідність присвоюється значення надлишковості і додається в словник.

```
red4key = dict()  
for obj in root:  
    f_symb = freq_symbols(obj["text"])  
    en = entropy(f_symb)  
    re = redundancy(en, alph_)  
    red4key[(obj["a"], obj["b"])] = re
```

Сортування по значенням ключа, надлишковості.

```
red4key = dict(sorted(red4key.items(), key=lambda item: item[1]))
```

Вивід в стовпчик

```
for k, v in red4key.items():  
    print(f"{k[0] :<5} {k[1] :<5} {v}")
```

Пошук ключа, який дав найбільшу надлишковість

```
max_key = max(red4key, key=red4key.get)
```

Вивід відповідного тексту, що відповідає знайденому ключу

```
for o in root:  
    if o["a"] == min_key[0] and o["b"] == min_key[1]:  
        print(o["text"])
```

В результаті отримуємо істинний текст, що було зашифровано

П'ять значень з найбільшою надлишковістю		
a	b	Надлишковість
13	151	0.10471366086993172
948	257	0.06938056480805155
117	74	0.057987678754593874
922	435	0.057557965454682725
844	67	0.050074753687583184

При аналізі результатів можна сказати, що хоч такий підхід і спрацював, але різниця між надлишковістю істинного тексту і найкращим хибним результатом не дуже велика. Хоча для показника надлишковості мабуть все ж помітна.

Другий метод використовує перевірку частот рідкісних біграм. Для цього було згенеровано всі можливі біграми з алфавіту. Та на деякому великому тексті підрахована кількість кожної біграми. Якщо біграма не зустрілась жодного разу, вона потрапила в список рідкісних біграм.

```
def bigrams(text, gap, cross):  
    bigrams = dict()  
  
    alph = alph_ + ' ' if gap else alph_  
  
    for i in alph:  
        for j in alph:  
            bigrams[i + j] = 0  
  
    for i in range(0, len(text) - 1, 1 if cross else 2):  
        item = text[i] + text[i + 1]  
        if item in bigrams.keys():  
            bigrams[item] += 1  
  
    new_b = dict()
```

```

for key, val in bigrams.items():
    if val == 0:
        new_b[key] = val

return new_b

```

Таким чином, список рідкісних біграм складається з 60 елементів:

*аы, бй, гй, дй, еы, жй, жиц, зиц, иы, йй, йэ, кц, оь, пй, сй, уь, уы, фжс, фй, фх, хй, хю, цжс, цй, цц, чг, чз, чй, чиц, чю, шжс, шз, шй, шц, шя, щб, щд, щжс, щз, щй, щс, щф, щх, щц, щцц, щы, щэ, щю, щя, ьы, ьа, ьб, ьы, ьэ, эц, эь, эы, юу, юь, юы, яы, яэ*

Такий підхід теж знайшов оригінальний ключ з такими результатами:

Пять значений з наименьшей кількостю рідкісних біграм		
а	б	Кількість рідкісних біграм
13	151	8
173	943	134
821	930	145
295	178	147
230	89	163

Проаналізувавши результати можна сказати, що даний підхід є більш чутливим, та працює краще. Адже в даному випадку вхідний текст має менше 10 рідкісних біграм, в той час як найкращий хибний результат має показник більше 100 одиниць. З цього можна зробити висновок, що такий алгоритм зможе краще розпізнавати природну мову.

Таким чином, знайдений ключ:

Ключ	(13, 151)
------	-----------

## Шифрований текст

лквдвдъышкрбызйакиабшачрнвязарчтчлчъкзтманэмнязйабштрпнхтрхрнзтжккысечамнмпы  
 вйвфяжтинфвйвйвсжнпчнмпуцзкыфвйвутсюцзкыкынмотзщбйьбшхолуычгкицепзкиань  
 уыфллфтыраючькиацзтыфэнкйяпезтнкжккысечамнмпжэпаычйдбцвсшчмтшслаиятасзбж  
 йьыбшывлтйэзщбцпцмщрифкздтеэккццзархрчосйпрйжклекакьяжюыщяояфскчбьязрчйз  
 чвгзжзычэявсшчтцлжочшызюшхачрнтмнкуфйзбчечвпчнотмнктхеотнчняцзбшрчычбчнкиц  
 гцлчъкевочфыщяцзреотйсфтбйщялчдечамнмпйарчтчццзтьярняыхашхаытызсепцяяюч  
 шзбшзтжмсячарнвзяозеарчэяицкятчрогцфэкыпэзтйпчазеэявахыдпдойдкрмпбцмвеэлжочр  
 чщтецрнбьяшкуэтыычлчокбцккузбнинепжвининачрнсджяццаяятчщтецрнбьяшквдиабцотия  
 ьаццйвычфткампьяэяддаьчшызюсяуядсяжутрхбцшчрнфэтзткзтцтеялчакиажажштзмнксяб  
 яешщтецрнбьяшкуэццеопнхояючбьястзырзгфлуфжмнкецьэтнкфячащжвжяымэвячатияцз  
 оеззднеэмэйкоевсщяыаяажвычцяучпяэязяшкинвдэакзюнзтмакырцсоушрнецчнкяуялжочоз  
 нкызаццнкяжсгмпчнвдепйдрчкеэярклнвцычпрычжкнпщюрчньаччквсеокяорнбччнйцнбш  
 зикзчшклзпеепаопниашчеквдзезэгцеккызаццнкшчрнхкнчъхвсфэиашцинэяьцзцычжтмэ  
 ывйвщтецрнбьяшктфбйьемтццзжеьытнщрпаозвзьнотпанхзайдкрмпбцсрпаццрущзлчшкле  
 еэхкжяццлтяыбчлуучвзпяэякящяцзклтвсбцяыыцлбцдйрцецкзвзвычяквсойюшххолуычнн  
 йвбнзеесвсоцпахышчгзючущядкщрпаозмеязябчмтмаззуыйюфэхьбшркбцуэдйуфрняынн

йвцяучрнкейпрцккутггцяжйухыксмпкырабцпабштхлтйвчябксогъракыбротхыачрнмнкршчу  
ярачыбязцрчфяяктфчнвдщтецрнбашкдфчжшюжачрнвязарчтгчучнплзраоьтпнкшчюйзтвйп  
цдзтофтфэцтнкэофтгчнщцккуфпяыщщряжеегщпцбцхкюзгзщырнэяччяыцзыэщрмпбцсрпарч  
тчбйхярняыжклжыьцснкшчэяутпамзгьпнсевсэзфяцзоэцтнвеэззвдчекеэгызнзтгчнпнивучп  
жкнкэблыибшхязрнпыьарчньччфьстланвеэиэмпрчвмкеэйкогхчтыыззэивьяньзяфякщтыэз  
чягшяжпсьжфтщюызкдзтзщачзяюшкзйзлафпэойзьялчуцднеэнпейвязярнбйеплюдфызякиа  
цзачрнвязаозеьхьрнфпечзэгмшчрнйахыбшнрчнммпмэхчйцбйвсчнммпмэьяючбьяярняыцеязо  
чйсхкфпхотнртмэчзкыквипйнктейесолйджкмэшчрзжйеспнмэйчяовытылуычмебцкяюцотн  
оыкиащзфтнгозаашятчфяжтгщтщвырчычбчтгжкрйупиажмыяшкмнйврбфяесоркееэллцеиа  
щзцяцзэмзщяебтцфвебзозяньюжючьвзжчсгьтчэыучрнепйаозделнйааьцяцзэкйэфтйсрнеце  
опнхоинхыэврцсбчзтманэмнязьящзйсиаычицнввдбцкыярнбютсюцзкыфпцеэярнкецзкы  
шчднжчюнийпозьящцнкйсепькжчокбцпцмнйаэккчюжяычягшнвдфкгнкмяфтпаюьукфвещыог  
збшучяпхкьоеинрцогэбфтпаюьтпнкэофяачщдвсеофтпаюьукфвмаолпаццнкяжыцсротвжуя  
ддыщзьяквякяюеобхзлзмзгштышспаэтивщцзексонвючшкиабшбйчззсеобйлзиротщзфтйсучф  
жэвдфяпзьеебччщяцзкодпшяюачйкщбечекиабшфяяцмнкыбэкгхчтыгшшчкгнккршчтчин  
шцияцзывьяючбятьююаькьзаучйзтысюиебчщзечучючкьквяднеэьачрнвязарчтчйдбйепл  
юрбучэтийшчрнвцебтцузйджчутеэьсаучочкиабшебхзбшфтнгозйюрбхобятчйцотасбйбчч  
яцегщечеоийюрбмэипкйчнезучлчмыбшхыздыяжкфэмпюжфтецжкнкецспнезнащзбштыфтф  
эотучиншцияцзовйдзеотечамнклзйяебччекфвийкинвдщыечикфвжяццзебчочьвеслеяздчюзю  
абйчыикфтщрчащяцзшсиаычицнввдвфтпаюьукфвйэинбьящзещецпйтзжятчхбцяычлуычфт  
лзньхярнбашкжкмафпзкфвчьхззгьутчняньзяньвсяююьытнотшрычйцсснмпйаццеяычрьхя  
рнечяыщзчнйвшхнвючшкиачяюйдбцьэтнкфякэцтзыхынмлзецккмвинзгчхрытнбцйдгмтщ  
цзрньырнсятчкывыгняжйзутйэлчцяцйцнйамврйпзквдзтмаьпнкэофяйтмпдфяечювузпбц  
йснуычфтинрцзтсрсяыйтсюжяюаяшявьфлфэбйьыичнафпзксоыярнгьтнрцтыяьрнэякпнкш  
чрнгсиаычицнввдевинзтсолчспейцаыячыбшйдзеэярнкецзрчжйупецйдгмтщцзтыфтещяты  
спецяжлчштзщецэтиылчтчкяяоечеклнжщдэпаычытчбнбйтзиклнязчнйвфэбйьыичжцхтзщ  
фпмавцеыичвззэлзбьзаццицхкпцкяхыозбятчызякиащзфяеьноччажсчащзьянвшхьягнлжццео  
флшххобятчьыдсьышзчягшшчрнфэнрчнмпйаццнкпнотсзлчрнссзмоежчыккюнкэбппкйфэуэ  
ебзоеыхынмицйдеэккотнчштплнкэотрчнмнммпмэчнйвдэмпкрнхжкыиюзрнечекицяыькеэиы  
юзрнучиншцияцзовиылчнькяуянпйсбцмнмпзкеэщйхчащднеэшдшызюуфачштвснюфязю  
уфзайдщытчычлждеэкрлрмпбцмвзаючькдфызякиащзачрнвязарчтгсжлжыяызыэтшийвыч  
ыввсхкрчызьярнбашкфтфссяыкыярнбашкчхйдрэягцширфшчучлжияшкрбнитятнрцшчрнгя  
тчлаэзмэщяшкиабшсеотбьяшущрчычышысепькейуплеязьярнсятчтажсеэщйхтщньфпчаыяч  
ыбшфтпаюьукфвееятчфяучысбхяпацытыызыкцзтьянввящыбчяыцпнйввяочьяхыцциуч  
юкмэвдючюжрьхярнечяыбшрийкщфяжтгщцейсвийпцсбшмпаычфткгнкыкряеыичвзрнпйкщт  
ыыззэкицбичжеиажчыккюнкэбмзяеязговьццеотгзякхучожечгзфтинрцбйзтрнзьфлшхф  
эычаэгмнкуффтчавяюзаоялсецгщлчькиащзрьцпфэцтбцккэоачрнвязарчтгзайяхялчькбйупб  
йфчыкпащзстзщиовьфэхьгшмзекчхююьытнотбщчучючцяццицтлфвычялкшяюаэкйпщрся  
лкицбчвыфябйщцмнмпзквдевийвюжючнвзщккзязщышкчхбйрнночягшрняыдкбцкяцяечик  
фвсбхятччянарчэясрмэтыфжхяшкйяиаючькнксяучяпкмплйяочрнзтжкшрмпбцсрпарчтчюеэ  
явсепнкэбфяжтгщднинепжвгщтытнвдкрычянийвдфмзынкшфяесйпхобнжшчфтыуычдезецн  
мяучтпмнфпйаеафэйсхкрнечжцяямицрнбчтчнасжнпоебчцеопнхофяжтгщачрнвязаозгк  
зщпцйпкяюиыйзбтедсяхынмпаэзхыызйдемусзщяхнфвеэтыычлчокбцккузбнжчуйупучьцотц  
ьянщммпуэфтцежскыназебечечсецкзйзхоуччяэяеагщтыщзяесзтвдйэузучнпйсрбчзньныачя  
куэтырнбчнксяжцпажэецотноыккрычднмнйвтыожаымэсогепоемзчйупйпщюйафэхнеэей  
джкицбчвырчычзжюцхырчнааьшыпащявьпнзеэяыязбшкыозрнотмусзщяхаэбычпабшкыт

нцммпрбчачязсьцотцсмннуычпеепшчьбъяяшкиабшпкмдщюевсзьмеязэзтыжцзеотлже  
еинеэнрычщывжккйэфяжзьянвшхфтцежсрчзнйвтыоjoyамэдфгефпоемзссиаыцицнввджкйс  
иахыычяктзфятыыяькоыечзнзтчхучычньбнзежкфэкксйщщккяжжагефпоеычссяжйзфтце  
жскыйзччцяикнкяжжаиаычэкуфиахыпнхофяаяяжы

## Розшифрованный текст

многогранную личность достоевского можно рассматривать с четырех сторон как писателя как н  
евротика как мыслителя этика как грешника как жеразобраться в этой невольной мушкетерской на  
сложности наименее спорен как писатель место его в одном ряду с Шекспиром братья Карамазо  
вывеличайший роман из всех когда-либо написанных легенда о великом инквизиторе одно из вы  
сочайших достижений мировой литературы переоценить которое невозможно к сожалению пер  
ед проблемой писательского творчества психоанализ должен сложить оружие достоевский скор  
ее всего уязвим как моралист представляя его человеком высоко нравственным на том основании  
что только тот достигает высшего нравственного совершенства кто прошел через глубочайшие б  
ездны греховности мы игнорируем одно изображение ведь нравственным является человек реаг  
ирующий уже на внутреннее испытание и искушение при этом ему не поддаваясь к тому же по пере  
менному грешит и рассказывает себе высокие нравственные цели того легко упрекнуть в то  
м что он слишком удобен для себя строит свою жизнь он не исполняет основного принципа нравст  
венности необходимости отречения во время как нравственный образ жизни в практических ин  
тересах всего человечества этим он напоминает варваров эпохи переселения народов варваров у  
бывавших из атемкаявших ся в этом так что пока я не установилось техническим примером расчи  
щавшим путь новым убийствам так же поступали вангрозный этас делка совестью характерна  
я русская черта достаточно бесславы и конечный итог нравственной борьбы достоевского после  
исступленной борьбы во имя примирения притязаний первичных позывов индивида требование  
ми человеческого общества он вынужден регрессирует к подчинению мирскому и духовному  
у авторитету поклонению царю и христианскому богу у русскому мелкодушному национализму  
у которому менее значительные умы пришли к гораздо меньшим усилиям чем он в этом слабое мест  
обольшой личности достоевский упустил возможность стать учителем и освободителем челове  
чества и присоединился к тюремщикам культу рабудущего немногим будет ему обязана в этом по  
всей вероятности проявился его невроз из-за которого он был осужден на такую неудачу помощи  
постижения и силе любви к людям было открыт другой апостольский путь служения нам предс  
тавляется отталкивающим рассматривание достоевского как качества грешника или преступника  
но это отталкивание не должно основываться на обывательской оценке преступника выявлять по  
длинную мотивацию преступления не долго для преступника существенны две черты безгранич  
ное себялюбие и сильная деструктивная склонность общим для обеих черт предпосылкой для и  
х проявлений является безлюбивость нехватка эмоционально оценочного отношения к челове  
ку тут сразу вспоминаешь противоположное этому у достоевского его большую потребность в л  
юбви и его огромную способность любить проявившуюся в его сверхдоброте и позволявшую ему  
любить и помогать там где он имел бы право ненавидеть и мстить например по отношению к его пе  
рвой жене и ее любовнику но тогда возникает вопрос откуда приходится обаяние и причисления дост  
оевского к преступникам ответ из выбора его сюжетов это преимущественно насильники и убий  
цы эгоцентрические характеры что свидетельствует о существовании таких склонностей в его в  
нутреннем мире а так же из некоторых фактов его жизни страсти его казартными играми может быть  
сексуального растрепанности незрелой девочки и исповедь это противоречие разрешается следующим  
образом сильная деструктивная устремленность достоевского которая могла бы сделать его пре

ступником была его жизнь направлена главным образом на самого себя, внутрь, в место того, что бы изнутри и таким образом выразилась в мазохизме и чувстве вины. В сетаки его личности не было ни садистических черт, выявляющихся в его раздражительности, мучительстве, нетерпимости, даже по отношению к любимым людям, а также его манере обращения с читателем, так в мелочах, ни садиства, ни в важном, садиство по отношению к самому себе, следовательно, мазохист, это мягчайший и добродушный и всегда готовый помочь человек в сложной личности Достоевского мы выделяли три фактора: один количественный и два качественных: его чрезвычайное повышение аффективности, его устремленность к перверзии, и которая должна была привести его к садомазохизму или сделать преступником, и его неподдающееся анализу творческое дарование, такое сочетание вполне могло бы существовать без невроза, ведь бывают жестокие процентные мазохисты без наличия невроза, в соотношении сил притязаний и первичных позывов и противоборствующих им торможений, присоединяя сюда возможность сублимирования Достоевского, все это можно было бы отнести к разряду импульсивных характеров, но положение вещей затемняется наличием невроза, не обязательно, но как бы сказано, приданных обстоятельствах, но все же возникающего тем самым, чемнасыщеннее осложнение, подлежащее ссоры, человеческое, преодоление невроза, это только знак того, что такой синтез не удался, что оно при этой попытке, оплатилось своим единством, в чем же в строгом смысле, проявляется невроз Достоевский, называл себя сам, и другие также считали его эпилептиком, на том основании, что он был подвержен тяжелой припадкам, сопровождавшимся потерей сознания, судорогами и последующим упадочным настроением, весьмавероятно, что эта так называемая эпилепсия, была лишь симптомом его невроза, который в таком случае, следует определить как истероэпилепсию, то есть как тяжелую истерию, утверждать это, с полной уверенностью, нельзя, по двум причинам: во-первых, потому, что даты и анамнезических припадков, так называемой эпилепсии Достоевского, недостаточны и ненадежны, а во-вторых, потому, что, по мнению исследователей, связанных с эпилептиками, и припадками болезненных состояний, остается неясным, а

## Висновки

У результаті виконання лабораторної роботи ми здобули навички частотного аналізу на прикладі розкриття моноалфавітної підстановки, а також опанували засоби роботи в модулярній арифметиці.

Під час порівняння методів розпізнавання змістовного тексту російською мовою, ми помітили, що другий підхід працює краще, ніж перший. Перший метод використовує надлишковість природної мови, але різниця між надлишковістю істинного тексту і найкращим хибним результатом не дуже велика. Проте інший метод, який перевіряє частоти рідкісних біграм, працює набагато краще, адже рідкісна біграма може зустрітися тільки в тому випадку, коли між двома словами було видалено пробіл.