

СИМЕТРИЧНА КРИПТОГРАФІЯ

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Виконали студенти групи ФІ-94
Костюк Кирило і Панасюк Єгор
Варіант-4

Мета роботи

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Хід роботи(типо код з сpp)

```
#include <iostream>
#include <fstream>
#include <string>
#include <vector>
#include <cctype>
#include <algorithm>
#include <cwctype>
#include <map>
#include <cmath>
#include <sstream>
#include <locale>
#include <codecvt>

using namespace std;
const map<wchar_t, wchar_t> lowercase =
{
    {L'A',L'a'},
    {L'B',L'b'},
    {L'В',L'в'},
    {L'Г',L'г'},
    {L'Д',L'д'},
    {L'E',L'e'},
    {L'Ё',L'e'},
    {L'Ж',L'ж'},
    {L'З',L'з'},
    {L'И',L'и'},
```

{L'Й',L'й'},
{L'К',L'к'},
{L'Л',L'л'},
{L'М',L'м'},
{L'Н',L'н'},
{L'О',L'о'},
{L'П',L'п'},
{L'Р',L'р'},
{L'С',L'с'},
{L'Т',L'т'},
{L'У',L'у'},
{L'Ф',L'ф'},
{L'Х',L'х'},
{L'Ц',L'ц'},
{L'Ч',L'ч'},
{L'Ш',L'ш'},
{L'Щ',L'щ'},
{L'Ъ',L'ъ'},
{L'Ы',L'ы'},
{L'Ь',L'ь'},
{L'Э',L'э'},
{L'Ю',L'ю'},
{L'Я',L'я'},
{L'а',L'а'},
{L'б',L'б'},
{L'в',L'в'},
{L'г',L'г'},
{L'д',L'д'},
{L'е',L'е'},
{L'ё',L'е'},
{L'ж',L'ж'},
{L'з',L'з'},
{L'и',L'и'},
{L'й',L'й'},
{L'к',L'к'},
{L'л',L'л'},
{L'м',L'м'},
{L'н',L'н'},
{L'о',L'о'},
{L'п',L'п'},
{L'р',L'р'},
{L'с',L'с'},
{L'т',L'т'},

```

        {L'y', L'y'},
        {L'Ѡ', L'Ѡ'},
        {L'х', L'х'},
        {L'ц', L'ц'},
        {L'ч', L'ч'},
        {L'ш', L'ш'},
        {L'щ', L'щ'},
        {L'ъ', L'ъ'},
        {L'ы', L'ы'},
        {L'ь', L'ь'},
        {L'э', L'э'},
        {L'ю', L'ю'},
        {L'я', L'я'},
        {L' ', L' '}
    };

wstring delete_rubish(const wstring& input) {
    wstring result;
    for (const auto& c : input) {
        if (lowercase.find(c) != lowercase.end()) {
            result.push_back(c);
        }
    }
    return result;
}

wstring lower_case(wstring input) {
    input = delete_rubish(input);
    for_each(input.begin(), input.end(), [&](wchar_t& c) {c =
lowercase.at(c); });
    return input;
}

wstring parse(const wstring& input) {

    wstring low_input = lower_case(input);
    wstringstream ss;
    ss << low_input;
    wstring buf, result;
    while (!ss.eof())
    {
        ss >> buf;
        auto c = ss.peek();

```

```

        do {
            ss.ignore(1);
            c = ss.peek();
        } while (c == L' ');
        result += buf + L' ';
    }
    result.erase(result.end() - 1);
    return result;
}

void output_frequency(const map<wchar_t, size_t>& letter_counter) {
    for (const auto& p : letter_counter) {
        wcout << L"letter: " << p.first << L" consized: " << p.second <<
endl;
    }
}

auto letter_frequency(wstring input) {
    map <wchar_t, size_t> let_counter;
    for (const auto& c : input) {
        ++let_counter[c];
    }

    output_frequency(let_counter);
    return let_counter;
}

void output_bigram(const map<wstring, size_t>& bigram_counter) {
    for (const auto& p : bigram_counter) {
        wcout << L"Bigram: " << p.first << L" consized " << p.second <<
endl;
    }
}

auto bigram_count(wstring input) {
    map <wstring, size_t> bigram_counter;
    for (int i = 0; i < input.length() - 1; i++) {

        ++bigram_counter[input.substr(i, 2)];
    }
    output_bigram(bigram_counter);
    return bigram_counter;
}

double letter_probability( double text_size, double letter_count )

```

```

{
    return letter_count / text_size;
}

double entropy_term( double probability )
{
    return probability * log2( probability );
}

double letter_entrop( const wstring& input )
{
    auto parsed_input = parse( input );
    auto text_size = parsed_input.size();
    auto letters_count = letter_frequency( parsed_input );
    double result = 0;

    for ( const auto& pair_let_count : letters_count )
    {
        auto probability = letter_probability( text_size ,
pair_let_count.second );
        result += entropy_term( probability );
    }

    result *= -1;

    return result;
}

double bigram_sum( const map<wstring, size_t>& bigrams )
{
    size_t dick = 0;
    for ( const auto& bigram : bigrams )
    {
        dick += bigram.second;
    }

    return dick;
}

double bigram_entrop( const wstring& input )
{
    auto parsed_input = parse( input );
    auto bigrams_count = bigram_count( parsed_input );

```

```

    auto all_bigrams = bigram_sum( bigrams_count );
    double result = 0;

    for ( const auto& pair_bigram_count : bigrams_count )
    {
        auto probability = letter_probability( all_bigrams,
pair_bigram_count.second );
        result += entropy_term( probability );
    }

    result *= -1;

    return result;
}

int main()
{
    setlocale(LC_ALL, "");
#ifdef _WIN32
    system("chcp 1251");
#endif

    std::wstring_convert<std::codecvt_utf8<wchar_t>> converter;
    wstring input;
    string line;
    std::ifstream fin( "../dead_souls.txt" );
    while (getline(fin, line, '\n') && !line.empty())
    {
        input += converter.from_bytes(line);
    }
    fin.close( );

    input = parse(input);
    wcout << input << endl;

    wcout << letter_entrop( input ) << std::endl;
    auto entropia_bigrams = bigram_entrop( input );
    wcout << entropia_bigrams << std::endl;
    wcout << "Result: " << entropia_bigrams / 2 << std::endl;
}

```

Output

letter: consized: 114174

letter: a consized: 46443
letter: б consized: 11610
letter: в consized: 28300
letter: г consized: 10596
letter: д consized: 18914
letter: е consized: 51768
letter: ж consized: 6739
letter: з consized: 10215
letter: и consized: 41357
letter: й consized: 6375
letter: к consized: 25022
letter: л consized: 27035
letter: м consized: 18004
letter: н consized: 37392
letter: о consized: 68219
letter: п consized: 16848
letter: р consized: 25008
letter: с consized: 30878
letter: т consized: 37309
letter: у consized: 18332
letter: ф consized: 726
letter: х consized: 5900
letter: ц consized: 1916
letter: ч consized: 11493
letter: ш consized: 5789
letter: щ consized: 1792
letter: ы consized: 11431
letter: ь consized: 12874
letter: э consized: 1361
letter: ю consized: 3631
letter: я consized: 11507

H1 = 4.38921

Bigram: a consized 1210
Bigram: б consized 5339
Bigram: в consized 10949
Bigram: г consized 2850
Bigram: д consized 5839

Bigram: е consized 2347
Bigram: ж consized 1531
Bigram: з consized 3052
Bigram: и consized 7509
Bigram: к consized 6862
Bigram: л consized 1645
Bigram: м consized 3628
Bigram: н consized 11162
Bigram: о consized 6458
Bigram: п consized 11928
Bigram: р consized 2689
Bigram: с consized 10658
Bigram: т consized 5905
Bigram: у consized 2967
Bigram: ф consized 373
Bigram: х consized 1283
Bigram: ц consized 205
Bigram: ч consized 4965
Bigram: ш consized 464
Bigram: щ consized 102
Bigram: ь consized 2
Bigram: э consized 1277
Bigram: ю consized 40
Bigram: я consized 935
Bigram: а consized 10001
Bigram: аа consized 21
Bigram: аб consized 467
Bigram: ав consized 2250
Bigram: аг consized 470
Bigram: ад consized 1056
Bigram: ае consized 894
Bigram: аж consized 985
Bigram: аз consized 2644
Bigram: аи consized 75
Bigram: ай consized 484
Bigram: ак consized 5081
Bigram: ал consized 4990

Bigram: ам consized 2520
Bigram: ан consized 2651
Bigram: ао consized 14
Bigram: ап consized 618
Bigram: ар consized 1462
Bigram: ас consized 2256
Bigram: ат consized 3444
Bigram: ау consized 75
Bigram: аф consized 84
Bigram: ах consized 717
Bigram: ац consized 35
Bigram: ач consized 611
Bigram: аш consized 601
Bigram: ащ consized 171
Bigram: аэ consized 8
Bigram: аю consized 560
Bigram: ая consized 1197
Bigram: б consized 175
Bigram: ба consized 875
Bigram: бв consized 33
Bigram: бг consized 5
Bigram: бд consized 32
Bigram: бе consized 1647
Bigram: бж consized 10
Bigram: бз consized 3
Bigram: би consized 427
Bigram: бк consized 94
Bigram: бл consized 560
Bigram: бм consized 21
Bigram: бн consized 223
Bigram: бо consized 1448
Bigram: бр consized 1086
Bigram: бс consized 110
Bigram: бт consized 14
Bigram: бу consized 1098
Bigram: бх consized 39
Bigram: бц consized 8

Bigram: бч consized 7
Bigram: бш consized 12
Bigram: бщ consized 86
Bigram: бы consized 3125
Bigram: бь consized 107
Bigram: бю consized 6
Bigram: бя consized 359
Bigram: в consized 4722
Bigram: ва consized 3906
Bigram: вб consized 10
Bigram: вв consized 65
Bigram: вг consized 6
Bigram: вд consized 243
Bigram: ве consized 3823
Bigram: вж consized 1
Bigram: вз consized 314
Bigram: ви consized 2164
Bigram: вк consized 214
Bigram: вл consized 396
Bigram: вm consized 106
Bigram: вn consized 848
Bigram: во consized 4706
Bigram: вп consized 178
Bigram: вр consized 449
Bigram: вс consized 2241
Bigram: вт consized 123
Bigram: ву consized 453
Bigram: vx consized 18
Bigram: вц consized 31
Bigram: вч consized 43
Bigram: вш consized 982
Bigram: вщ consized 5
Bigram: вы consized 1898
Bigram: вь consized 154
Bigram: вэ consized 8
Bigram: вю consized 1
Bigram: вя consized 192

Bigram: г consized 500
Bigram: га consized 568
Bigram: гб consized 4
Bigram: гв consized 11
Bigram: гд consized 688
Bigram: ге consized 286
Bigram: ги consized 480
Bigram: гк consized 70
Bigram: гл consized 817
Bigram: гм consized 7
Bigram: гн consized 177
Bigram: го consized 5850
Bigram: гр consized 530
Bigram: гс consized 6
Bigram: гт consized 3
Bigram: гу consized 578
Bigram: гч consized 18
Bigram: гш consized 3
Bigram: д consized 737
Bigram: да consized 3634
Bigram: дб consized 29
Bigram: дв consized 615
Bigram: дг consized 3
Bigram: дд consized 14
Bigram: де consized 3517
Bigram: дж consized 4
Bigram: дз consized 6
Bigram: ди consized 1404
Bigram: дк consized 185
Bigram: дл consized 407
Bigram: дм consized 75
Bigram: дн consized 1263
Bigram: до consized 2287
Bigram: дп consized 85
Bigram: др consized 1098
Bigram: дс consized 393
Bigram: дт consized 115

Bigram: ду consized 1423
Bigram: дх consized 40
Bigram: дц consized 180
Bigram: дч consized 25
Bigram: дш consized 44
Bigram: ды consized 335
Bigram: дь consized 729
Bigram: дю consized 18
Bigram: дя consized 249
Bigram: е consized 13263
Bigram: еа consized 34
Bigram: еб consized 1067
Bigram: ев consized 1370
Bigram: ег consized 2090
Bigram: ед consized 2184
Bigram: ее consized 743
Bigram: еж consized 533
Bigram: ез consized 815
Bigram: еи consized 93
Bigram: ей consized 1519
Bigram: ек consized 1053
Bigram: ел consized 4403
Bigram: ем consized 2870
Bigram: ен consized 4685
Bigram: ео consized 182
Bigram: еп consized 717
Bigram: ер consized 3876
Bigram: ес consized 3324
Bigram: ет consized 3807
Bigram: еу consized 63
Bigram: еф consized 5
Bigram: ех consized 528
Bigram: ец consized 387
Bigram: еч consized 721
Bigram: еш consized 580
Bigram: ещ consized 597
Bigram: еэ consized 5

Bigram: ею consized 148
Bigram: ея consized 106
Bigram: ж consized 488
Bigram: жа consized 633
Bigram: жб consized 58
Bigram: жд consized 498
Bigram: же consized 2903
Bigram: жж consized 7
Bigram: жи consized 1038
Bigram: жк consized 102
Bigram: жл consized 11
Bigram: жм consized 14
Bigram: жн consized 697
Bigram: жо consized 70
Bigram: жр consized 1
Bigram: жс consized 17
Bigram: жу consized 148
Bigram: жч consized 19
Bigram: жъ consized 35
Bigram: з consized 858
Bigram: за consized 3732
Bigram: зб consized 151
Bigram: зв consized 702
Bigram: зг consized 181
Bigram: зд consized 694
Bigram: зе consized 299
Bigram: зж consized 106
Bigram: зз consized 39
Bigram: зи consized 213
Bigram: зк consized 77
Bigram: зл consized 118
Bigram: зм consized 110
Bigram: зн consized 1233
Bigram: зо consized 422
Bigram: зп consized 17
Bigram: зр consized 87
Bigram: зс consized 31

Bigram: зт consized 15
Bigram: зу consized 214
Bigram: зх consized 1
Bigram: зц consized 6
Bigram: зч consized 29
Bigram: зш consized 7
Bigram: зы consized 259
Bigram: зь consized 176
Bigram: зэ consized 1
Bigram: зю consized 4
Bigram: зя consized 433
Bigram: и consized 13093
Bigram: иа consized 33
Bigram: иб consized 578
Bigram: ив consized 1589
Bigram: иг consized 369
Bigram: ид consized 895
Bigram: ие consized 1428
Bigram: иж consized 196
Bigram: из consized 1481
Bigram: ии consized 270
Bigram: ий consized 826
Bigram: ик consized 2427
Bigram: ил consized 3169
Bigram: им consized 1899
Bigram: ин consized 2326
Bigram: ио consized 141
Bigram: ип consized 183
Bigram: ир consized 398
Bigram: ис consized 1667
Bigram: ит consized 2853
Bigram: иу consized 8
Bigram: иф consized 110
Bigram: их consized 1289
Bigram: иц consized 610
Bigram: ич consized 1953
Bigram: иш consized 353

Bigram: ищ consized 146

Bigram: иэ consized 5

Bigram: ию consized 188

Bigram: ия consized 874

Bigram: й consized 4817

Bigram: йа consized 6

Bigram: йб consized 30

Bigram: йв consized 5

Bigram: йг consized 3

Bigram: йд consized 106

Bigram: йе consized 5

Bigram: йз consized 1

Bigram: йи consized 4

Bigram: йк consized 212

Bigram: йл consized 26

Bigram: йм consized 71

Bigram: йн consized 288

Bigram: йо consized 6

Bigram: йп consized 11

Bigram: йр consized 1

Bigram: йс consized 356

Bigram: йт consized 235

Bigram: йх consized 1

Bigram: йц consized 19

Bigram: йч consized 36

Bigram: йш consized 134

Bigram: йэ consized 1

Bigram: йя consized 1

Bigram: к consized 3815

Bigram: ка consized 6340

Bigram: кб consized 2

Bigram: кв consized 122

Bigram: кг consized 5

Bigram: кд consized 19

Bigram: ке consized 497

Bigram: кж consized 39

Bigram: кз consized 4

Bigram: ки consized 2452
Bigram: кк consized 8
Bigram: кл consized 385
Bigram: км consized 2
Bigram: кн consized 452
Bigram: ко consized 7628
Bigram: кп consized 8
Bigram: кр consized 1028
Bigram: кс consized 67
Bigram: кт consized 413
Bigram: ку consized 1717
Bigram: кх consized 1
Bigram: кц consized 7
Bigram: кч consized 5
Bigram: кш consized 4
Bigram: кэ consized 2
Bigram: л consized 4604
Bigram: ла consized 3120
Bigram: лб consized 23
Bigram: лв consized 21
Bigram: лг consized 123
Bigram: лд consized 36
Bigram: ле consized 2532
Bigram: лж consized 203
Bigram: лз consized 4
Bigram: ли consized 4458
Bigram: лк consized 304
Bigram: лл consized 69
Bigram: лм consized 8
Bigram: лн consized 146
Bigram: ло consized 4427
Bigram: лп consized 24
Bigram: лр consized 1
Bigram: лс consized 945
Bigram: лт consized 54
Bigram: лу consized 979
Bigram: лф consized 11

Bigram: лц consized 2
Bigram: лч consized 57
Bigram: лщ consized 7
Bigram: лы consized 402
Bigram: ль consized 2869
Bigram: лэ consized 4
Bigram: лю consized 645
Bigram: ля consized 957
Bigram: м consized 5457
Bigram: ма consized 1662
Bigram: мб consized 21
Bigram: мв consized 7
Bigram: мг consized 19
Bigram: мд consized 12
Bigram: ме consized 2693
Bigram: мж consized 1
Bigram: мз consized 5
Bigram: ми consized 2177
Bigram: мк consized 77
Bigram: мл consized 138
Bigram: мм consized 23
Bigram: мн consized 961
Bigram: мо consized 1990
Bigram: мп consized 41
Bigram: мр consized 21
Bigram: мс consized 84
Bigram: мт consized 47
Bigram: му consized 1571
Bigram: мф consized 4
Bigram: мх consized 3
Bigram: мц consized 13
Bigram: мч consized 24
Bigram: мш consized 3
Bigram: мщ consized 7
Bigram: мы consized 579
Bigram: мь consized 42
Bigram: мэ consized 3

Bigram: мя consized 319
Bigram: н consized 2442
Bigram: на consized 6345
Bigram: нб consized 13
Bigram: нв consized 12
Bigram: нг consized 19
Bigram: нд consized 169
Bigram: не consized 6793
Bigram: нж consized 50
Bigram: нз consized 16
Bigram: ни consized 5246
Bigram: нк consized 276
Bigram: нл consized 2
Bigram: нм consized 2
Bigram: нн consized 1904
Bigram: но consized 6908
Bigram: нп consized 2
Bigram: нр consized 44
Bigram: нс consized 286
Bigram: нт consized 253
Bigram: ну consized 1831
Bigram: нф consized 11
Bigram: нх consized 1
Bigram: нц consized 162
Bigram: нч consized 128
Bigram: нш consized 4
Bigram: нщ consized 33
Bigram: ны consized 2314
Bigram: нь consized 1227
Bigram: нэ consized 1
Bigram: ню consized 119
Bigram: ня consized 779
Bigram: о consized 15248
Bigram: оа consized 17
Bigram: об consized 2781
Bigram: ов consized 6242
Bigram: ог consized 3175

Bigram: од consized 3619
Bigram: ое consized 1474
Bigram: ож consized 1358
Bigram: оз consized 1134
Bigram: ои consized 511
Bigram: ой consized 2248
Bigram: ок consized 1438
Bigram: ол consized 3826
Bigram: ом consized 3642
Bigram: он consized 3579
Bigram: оо consized 150
Bigram: оп consized 915
Bigram: ор consized 4229
Bigram: ос consized 4308
Bigram: от consized 4615
Bigram: оу consized 72
Bigram: оъ consized 66
Bigram: ох consized 359
Bigram: оц consized 41
Bigram: оч consized 1373
Bigram: ош consized 758
Bigram: ощ consized 110
Bigram: оэ consized 27
Bigram: ою consized 526
Bigram: оя consized 378
Bigram: п consized 31
Bigram: па consized 837
Bigram: пв consized 1
Bigram: пд consized 1
Bigram: пе consized 1526
Bigram: пи consized 712
Bigram: пк consized 108
Bigram: пл consized 550
Bigram: пн consized 71
Bigram: по consized 6970
Bigram: пп consized 16
Bigram: пр consized 5035

Bigram: пс consized 7
Bigram: пт consized 40
Bigram: пу consized 453
Bigram: пх consized 1
Bigram: пц consized 36
Bigram: пч consized 40
Bigram: пш consized 11
Bigram: пщ consized 15
Bigram: пы consized 141
Bigram: пь consized 44
Bigram: пя consized 202
Bigram: р consized 510
Bigram: ра consized 4637
Bigram: рб consized 79
Bigram: рв consized 180
Bigram: рг consized 85
Bigram: рд consized 219
Bigram: ре consized 4084
Bigram: рж consized 163
Bigram: рз consized 27
Bigram: ри consized 3456
Bigram: рк consized 203
Bigram: рл consized 70
Bigram: рм consized 117
Bigram: рн consized 514
Bigram: ро consized 5108
Bigram: рп consized 78
Bigram: рр consized 9
Bigram: рс consized 195
Bigram: рт consized 562
Bigram: ру consized 2000
Bigram: рф consized 29
Bigram: рх consized 124
Bigram: рц consized 15
Bigram: рч consized 45
Bigram: рш consized 288
Bigram: рщ consized 15

Bigram: ры consized 950
Bigram: рь consized 407
Bigram: рю consized 128
Bigram: ря consized 711
Bigram: с consized 1846
Bigram: са consized 1231
Bigram: сб consized 39
Bigram: св consized 886
Bigram: сг consized 13
Bigram: сд consized 218
Bigram: се consized 2866
Bigram: сж consized 4
Bigram: сэ consized 11
Bigram: си consized 976
Bigram: ск consized 2754
Bigram: сл consized 1981
Bigram: см consized 488
Bigram: сн consized 554
Bigram: со consized 1786
Bigram: сп consized 986
Bigram: ср consized 128
Bigram: сс consized 480
Bigram: ст consized 7494
Bigram: су consized 580
Bigram: сф consized 2
Bigram: сх consized 167
Bigram: сц consized 17
Bigram: сч consized 191
Bigram: сш consized 54
Bigram: сы consized 200
Bigram: съ consized 1938
Bigram: сю consized 135
Bigram: ся consized 2853
Bigram: т consized 3992
Bigram: та consized 4120
Bigram: тб consized 13
Bigram: тв consized 1673

Bigram: тг consized 9
Bigram: тд consized 111
Bigram: те consized 4002
Bigram: тз consized 19
Bigram: ти consized 2247
Bigram: тк consized 385
Bigram: тл consized 138
Bigram: тм consized 37
Bigram: тн consized 998
Bigram: то consized 9549
Bigram: тп consized 94
Bigram: тр consized 1738
Bigram: тс consized 805
Bigram: тт consized 69
Bigram: ту consized 1092
Bigram: тф consized 7
Bigram: тх consized 10
Bigram: тц consized 34
Bigram: тч consized 110
Bigram: тш consized 9
Bigram: ты consized 1078
Bigram: ть consized 4551
Bigram: тю consized 70
Bigram: тя consized 349
Bigram: у consized 4614
Bigram: уа consized 23
Bigram: уб consized 521
Bigram: ув consized 492
Bigram: уг consized 940
Bigram: уд consized 1473
Bigram: уе consized 191
Bigram: уж consized 1413
Bigram: уз consized 264
Bigram: уи consized 8
Bigram: уй consized 93
Bigram: ук consized 595
Bigram: ул consized 694

Bigram: ум consized 913
Bigram: ун consized 165
Bigram: уо consized 4
Bigram: уп consized 693
Bigram: ур consized 484
Bigram: ус consized 877
Bigram: ут consized 1194
Bigram: уу consized 1
Bigram: уф consized 10
Bigram: ух consized 360
Bigram: уц consized 2
Bigram: уч consized 707
Bigram: уш consized 651
Bigram: ущ consized 163
Bigram: уэ consized 5
Bigram: ую consized 746
Bigram: уя consized 36
Bigram: ф consized 7
Bigram: фа consized 210
Bigram: фе consized 134
Bigram: фи consized 136
Bigram: фл consized 9
Bigram: фм consized 2
Bigram: фн consized 1
Bigram: фо consized 59
Bigram: фр consized 109
Bigram: фс consized 6
Bigram: фт consized 10
Bigram: фу consized 19
Bigram: фч consized 2
Bigram: фы consized 9
Bigram: фь consized 9
Bigram: фю consized 1
Bigram: фя consized 3
Bigram: х consized 2779
Bigram: ха consized 500
Bigram: хв consized 156

Bigram: хг consized 5
Bigram: хд consized 5
Bigram: хе consized 71
Bigram: хз consized 2
Bigram: хи consized 118
Bigram: хк consized 3
Bigram: хл consized 213
Bigram: хм consized 17
Bigram: хн consized 142
Bigram: хо consized 1596
Bigram: хп consized 5
Bigram: хр consized 80
Bigram: хс consized 44
Bigram: хт consized 24
Bigram: ху consized 115
Bigram: хх consized 2
Bigram: хц consized 1
Bigram: хч consized 8
Bigram: хш consized 8
Bigram: хъ consized 1
Bigram: хэ consized 3
Bigram: хя consized 2
Bigram: ц consized 371
Bigram: ца consized 465
Bigram: цв consized 44
Bigram: цд consized 1
Bigram: це consized 507
Bigram: ци consized 86
Bigram: цк consized 31
Bigram: цл consized 2
Bigram: цн consized 1
Bigram: цо consized 133
Bigram: цу consized 134
Bigram: цы consized 141
Bigram: ч consized 377
Bigram: ча consized 1361
Bigram: чв consized 4

Bigram: чг consized 1
Bigram: чд consized 4
Bigram: че consized 2463
Bigram: чз consized 1
Bigram: чи consized 2801
Bigram: чк consized 304
Bigram: чл consized 10
Bigram: чм consized 6
Bigram: чн consized 567
Bigram: чо consized 62
Bigram: чп consized 3
Bigram: чр consized 31
Bigram: чс consized 1
Bigram: чт consized 2728
Bigram: чу consized 487
Bigram: чх consized 2
Bigram: чч consized 3
Bigram: чш consized 115
Bigram: чъ consized 161
Bigram: чя consized 1
Bigram: ш consized 115
Bigram: ша consized 642
Bigram: шв consized 17
Bigram: шд consized 1
Bigram: ше consized 1473
Bigram: ши consized 1609
Bigram: шк consized 568
Bigram: шл consized 285
Bigram: шм consized 3
Bigram: шн consized 193
Bigram: шо consized 228
Bigram: шп consized 23
Bigram: шр consized 1
Bigram: шт consized 47
Bigram: шу consized 178
Bigram: шц consized 4
Bigram: шч consized 1

Bigram: шъ consized 401
Bigram: щ consized 9
Bigram: ща consized 239
Bigram: ще consized 921
Bigram: щи consized 521
Bigram: щн consized 11
Bigram: що consized 1
Bigram: щр consized 7
Bigram: щу consized 62
Bigram: щъ consized 21
Bigram: ы consized 3183
Bigram: ыа consized 3
Bigram: ыб consized 99
Bigram: ыв consized 612
Bigram: ыг consized 104
Bigram: ыд consized 80
Bigram: ые consized 897
Bigram: ыж consized 13
Bigram: ыз consized 28
Bigram: ыи consized 17
Bigram: ый consized 1049
Bigram: ык consized 196
Bigram: ыл consized 1718
Bigram: ым consized 855
Bigram: ын consized 130
Bigram: ыо consized 8
Bigram: ып consized 139
Bigram: ыр consized 251
Bigram: ыс consized 457
Bigram: ыт consized 451
Bigram: ыу consized 7
Bigram: ых consized 716
Bigram: ыц consized 3
Bigram: ыч consized 86
Bigram: ыш consized 309
Bigram: ыщ consized 18
Bigram: ыэ consized 1

Bigram: ыя consized 1
Bigram: ь consized 7505
Bigram: ыа consized 27
Bigram: ыб consized 65
Bigram: ыв consized 24
Bigram: ыг consized 60
Bigram: ыд consized 59
Bigram: ые consized 693
Bigram: ыж consized 4
Bigram: ыз consized 145
Bigram: ыи consized 37
Bigram: ык consized 988
Bigram: ыл consized 3
Bigram: ым consized 223
Bigram: ын consized 843
Bigram: ыо consized 16
Bigram: ып consized 16
Bigram: ыс consized 853
Bigram: ыт consized 148
Bigram: ыу consized 2
Bigram: ыф consized 3
Bigram: ых consized 5
Bigram: ыц consized 79
Bigram: ыч consized 52
Bigram: ыш consized 281
Bigram: ыщ consized 4
Bigram: ыэ consized 2
Bigram: ыю consized 219
Bigram: ыя consized 518
Bigram: э consized 5
Bigram: эг consized 1
Bigram: эд consized 53
Bigram: эй consized 18
Bigram: эк consized 76
Bigram: эл consized 6
Bigram: эм consized 12
Bigram: эп consized 11

Bigram: эс consized 4
Bigram: эт consized 1141
Bigram: эф consized 2
Bigram: эх consized 32
Bigram: ю consized 2083
Bigram: юа consized 2
Bigram: юб consized 200
Bigram: юв consized 9
Bigram: юг consized 9
Bigram: юд consized 232
Bigram: юе consized 2
Bigram: юж consized 19
Bigram: юз consized 6
Bigram: юи consized 4
Bigram: юк consized 19
Bigram: юл consized 7
Bigram: юм consized 20
Bigram: юн consized 66
Bigram: юо consized 3
Bigram: юп consized 9
Bigram: юр consized 70
Bigram: юс consized 104
Bigram: ют consized 298
Bigram: юф consized 7
Bigram: юх consized 34
Bigram: юц consized 4
Bigram: юч consized 79
Bigram: юш consized 105
Bigram: ющ consized 209
Bigram: юэ consized 5
Bigram: юю consized 22
Bigram: юя consized 4
Bigram: я consized 6527
Bigram: яа consized 10
Bigram: яб consized 21
Bigram: яв consized 209
Bigram: яг consized 51

Bigram: яд consized 353
Bigram: яе consized 57
Bigram: яж consized 99
Bigram: яз consized 194
Bigram: яи consized 108
Bigram: яй consized 138
Bigram: як consized 312
Bigram: ял consized 387
Bigram: ям consized 313
Bigram: ян consized 541
Bigram: яо consized 9
Bigram: яп consized 48
Bigram: яр consized 65
Bigram: яс consized 295
Bigram: ят consized 1207
Bigram: яу consized 5
Bigram: яф consized 2
Bigram: ях consized 167
Bigram: яц consized 15
Bigram: яч consized 155
Bigram: яш consized 9
Bigram: ящ consized 104
Bigram: яэ consized 3
Bigram: юю consized 55
Bigram: яя consized 48
7.94273
Result: H2 = 3.97137

Як видно з виводу програми H1 та H2 на “Мертвих душах” Гоголя видається:

H1 = 4.38921

H2 = 3.97137

З програмки:

Лабораторная работа №1

Произвольная часть текста:
ими_эле_не_более_чем_в_цвете_их_волос_я_знаю_что_по_мнению_некоторых_людей_

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: e

Символ по счету: 1

Номер эксперимента: 52

Неравенство для энтропии:
 $1.86328911854908 < H < 2.58769131187048$

Двоичная таблица угаданных символов:

00100000000000000000000000000000	▲
00000000000001000000000000000000	
00010000000000000000000000000000	
10000000000000000000000000000000	
00010000000000000000000000000000	▼

Вероятности:

q[1] = 0.5
q[2] = 0.1346153
q[3] = 0.0769230
q[4] = 0.0769230
q[5] = 0
q[6] = 0.0192307
q[7] = 0.0192307
q[8] = 0
q[9] = 0
q[10] = 0
q[11] = 0.019230
q[12] = 0
q[13] = 0
q[14] = 0.019230
q[15] = 0
q[16] = 0.019230
q[17] = 0
q[18] = 0
q[19] = 0.038461
q[20] = 0
q[21] = 0.019230
q[22] = 0.019230
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0.038461
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

Поле ввода символов:
e

Продолжить Другой

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

$$1.86 < H(10) < 2.58$$

Лабораторная работа №1

Произвольная часть текста:
ей_семьи_или_к_тем_кто_живет_вокруг_или_вообще_ко_всем_людям_однако_они_все

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: к

Символ по счету: 1

Номер эксперимента: 52

Неравенство для энтропии:
 $1.51357636042729 < H < 2.29214029359501$

Двоичная таблица угаданных символов:

10000000000000000000000000000000	▲
00010000000000000000000000000000	
00000000000000010000000000000000	
10000000000000000000000000000000	
10000000000000000000000000000000	▼

Вероятности:

q[1] = 0.5961538
q[2] = 0.1346153
q[3] = 0.0384615
q[4] = 0.0192307
q[5] = 0.0192307
q[6] = 0.0192307
q[7] = 0
q[8] = 0
q[9] = 0
q[10] = 0.019230
q[11] = 0.019230
q[12] = 0.019230
q[13] = 0
q[14] = 0
q[15] = 0.019230
q[16] = 0.038461
q[17] = 0.019230
q[18] = 0
q[19] = 0
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0.019230
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0
q[31] = 0.019230
q[32] = 0

Поле ввода символов:
к

Продолжить Другой

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

$$1.51 < H(10) < 2.29$$

