Міністерство освіти і науки України Національний технічний університет України “Київський політехнічний інститут ім. Ігоря Сікорського” Фізико-технічний інститут

**КРИПТОГРАФІЯ КОМП’ЮТЕРНИЙ ПРАКТИКУМ №1 Експериментальна оцінка ентропії на символ джерела відкритого тексту**

Виконали: Студент групи ФБ-05 Даниленко Данило,  
Студентка ФБ-05 Мірошніченко Ілона

**Київ – 2022**

**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.

**Порядок виконання роботи**

0. Уважно прочитати методичні вказівки до виконання комп’ютерного практикуму.

1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку 1 H та 2 H за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення 1 H та 2 H на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення 1 H та 2 H на тому ж тексті, в якому вилучено всі пробіли.

2. За допомогою програми CoolPinkProgram оцінити значення (10) H, (20) H , (30) H . 3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

**Хід роботи**

Перед виконанням роботи були розглянуті теоретичні відомості в методичних вказівках. В якості експериментального тексту була взята книга «Good Omens» в перекладі на російську мову. Оригінал тексту можна знайти у файлі badtxt.txt. Відредагований текст з пробілами міститься у файлі spaces.txt, а без пробілів у nospaces.txt. В ході виконання роботи було прийнято рішення пробіли замінити на «\_» для кращого сприйняття. Усі таблиці також наведені у файлах з відповідними назвами. Єдина відмінність - таблиці біграм були виконанні у двох варіантах(таблиці у вигляді матриці і звичайні)

Monograms with space

h1 with space - 4.4769572317595205

r1 with space - 0.1200019790888095

о - 0.102851

а - 0.088136

е - 0.078324

и - 0.075033

н - 0.065523

т - 0.062623

с - 0.050766

р - 0.048078

л - 0.042801

в - 0.042185

к - 0.036785

у - 0.035314

м - 0.029008

д - 0.028709

п - 0.026668

я - 0.021505

ы - 0.018928

ь - 0.018739

б - 0.018396

з - 0.017728

г - 0.015919

ч - 0.014189

х - 0.011443

й - 0.010844

ж - 0.010705

ш - 0.00962

ю - 0.005917

щ - 0.004627

ц - 0.004129

э - 0.003246

ф - 0.00091

ъ - 0.00018

ё - 0.00017

Monograms without space

h1 with space - 4.4769572317595205

r1 without space - 0.11248977413377115

о - 0.102851

а - 0.088136

е - 0.078324

и - 0.075033

н - 0.065523

т - 0.062623

с - 0.050766

р - 0.048078

л - 0.042801

в - 0.042185

к - 0.036785

у - 0.035314

м - 0.029008

д - 0.028709

п - 0.026668

я - 0.021505

ы - 0.018928

ь - 0.018739

б - 0.018396

з - 0.017728

г - 0.015919

ч - 0.014189

х - 0.011443

й - 0.010844

ж - 0.010705

ш - 0.00962

ю - 0.005917

щ - 0.004627

ц - 0.004129

э - 0.003246

ф - 0.00091

ъ - 0.00018

ё - 0.00017

Bigrams with space

h2 with space - 4.112855148278171

r2 with space - 0.19157047891727508

то - 0.016059

но - 0.015962

на - 0.013675

ст - 0.013129

ра - 0.012474

по - 0.011972

не - 0.011686

ка - 0.010931

ни - 0.010508

ро - 0.009712

ко - 0.009658

ал - 0.009501

ть - 0.009323

от - 0.009215

та - 0.009139

ен - 0.009005

ре - 0.008767

ли - 0.008652

ак - 0.008572

ет - 0.008374

ос - 0.00826

ва - 0.008142

ер - 0.007925

ат - 0.007908

пр - 0.007869

во - 0.007767

ла - 0.007706

ов - 0.007621

он - 0.007581

да - 0.007293

го - 0.006995

ин - 0.006739

об - 0.006622

ри - 0.006515

ил - 0.006464

ор - 0.006464

ом - 0.006445

те - 0.006421

де - 0.006334

ан - 0.006322

ар - 0.006249

ел - 0.006196

ит - 0.006135

ло - 0.006084

ся - 0.00598

за - 0.005977

ле - 0.0059

ол - 0.005777

од - 0.005481

ру - 0.005364

ти - 0.005358

ме - 0.00535

ки - 0.005229

ма - 0.00517

ас - 0.005096

ши - 0.005084

мо - 0.004831

ог - 0.004815

се - 0.004811

ес - 0.004767

ам - 0.004733

ны - 0.004706

ик - 0.004607

ем - 0.004605

ну - 0.004604

ль - 0.004604

аз - 0.004474

ку - 0.004457

же - 0.004414

тр - 0.004414

ве - 0.004409

ой - 0.004333

до - 0.004309

би - 0.004299

ск - 0.003962

кр - 0.003909

ав - 0.003902

ая - 0.003882

ис - 0.003878

сь - 0.003878

чт - 0.003839

ми - 0.00379

со - 0.003679

уд - 0.00362

ив - 0.003577

пе - 0.00357

вы - 0.003524

сл - 0.003479

ча - 0.003449

ад - 0.003444

нн - 0.003412

ут - 0.00338

бы - 0.003375

им - 0.003344

са - 0.003305

чи - 0.003302

ед - 0.0033

вс - 0.003279

ры - 0.003166

ок - 0.003131

ви - 0.003111

из - 0.00305

аь - 0.003032

ег - 0.002947

че - 0.002836

жи - 0.002813

ож - 0.002722

бе - 0.002653

сп - 0.002624

ул - 0.002583

аг - 0.002575

тв - 0.002566

ей - 0.002566

ук - 0.00253

бо - 0.002523

ир - 0.002476

ше - 0.002412

эт - 0.002386

ие - 0.002366

га - 0.002363

ыв - 0.002342

ще - 0.002283

ду - 0.002262

их - 0.002219

ев - 0.002217

иь - 0.002208

хо - 0.002179

цу - 0.002174

ус - 0.00217

ня - 0.002165

оч - 0.002128

ьн - 0.002092

си - 0.002076

вн - 0.002075

тс - 0.002072

бу - 0.002072

оз - 0.002062

ид - 0.002044

ге - 0.002037

ят - 0.002023

ун - 0.002015

ыл - 0.002003

ди - 0.001993

па - 0.001947

ох - 0.001929

ай - 0.001917

хи - 0.0019

ля - 0.001886

ты - 0.001881

ия - 0.001843

ря - 0.00183

ту - 0.001825

еч - 0.00182

еь - 0.001786

аю - 0.001774

оп - 0.001773

ае - 0.001765

лс - 0.001764

ба - 0.001704

гл - 0.001688

ое - 0.001682

жа - 0.001674

аж - 0.001665

ый - 0.001654

ех - 0.001638

уж - 0.00163

мн - 0.001615

дж - 0.001605

ум - 0.001592

еб - 0.001579

зн - 0.001571

уч - 0.001571

му - 0.001563

еп - 0.001544

ез - 0.00152

тн - 0.001519

ее - 0.001506

ьс - 0.001484

дн - 0.001473

ыа - 0.001468

зе - 0.001442

ур - 0.001441

мы - 0.001438

ач - 0.001433

вз - 0.001433

щи - 0.001425

оь - 0.001415

ых - 0.001414

жн - 0.001406

ич - 0.001395

бр - 0.001388

ые - 0.001388

ую - 0.001373

нв - 0.00136

ек - 0.001336

ьо - 0.001326

нй - 0.001317

вр - 0.001317

см - 0.001311

ах - 0.001304

хн - 0.001296

рь - 0.001294

зу - 0.00128

пл - 0.00128

лу - 0.001273

чн - 0.001247

зд - 0.001246

гр - 0.001218

оы - 0.001212

лы - 0.001204

км - 0.001201

ий - 0.00119

зо - 0.001187

ац - 0.001186

ющ - 0.001178

аб - 0.001174

св - 0.00116

ещ - 0.001148

йн - 0.001138

пд - 0.001132

ым - 0.001132

кл - 0.001129

вл - 0.001127

ха - 0.001122

эо - 0.001117

ап - 0.001095

ыс - 0.001087

вш - 0.001084

тк - 0.001078

ша - 0.001049

зр - 0.001045

сн - 0.001032

бл - 0.001027

еш - 0.001026

дв - 0.001024

рн - 0.001014

ды - 0.001013

яд - 0.001006

лю - 0.000967

уг - 0.000966

ои - 0.000963

зв - 0.000962

су - 0.000958

рт - 0.000955

пи - 0.000953

пу - 0.000953

ср - 0.000949

др - 0.000949

яс - 0.00094

аы - 0.000935

гд - 0.000916

уш - 0.000909

ощ - 0.000902

ье - 0.000892

це - 0.000889

нд - 0.000884

аш - 0.000884

ыт - 0.000881

ыо - 0.000877

ву - 0.000864

уп - 0.000857

сс - 0.000853

ии - 0.000851

ги - 0.000851

иб - 0.000847

уб - 0.000838

кт - 0.000828

оя - 0.000824

ьк - 0.000817

ял - 0.000815

рй - 0.000814

чу - 0.000804

зм - 0.000793

дл - 0.000778

нт - 0.000774

еж - 0.000769

ща - 0.000768

жу - 0.000765

лн - 0.000764

ош - 0.000763

эй - 0.00076

жд - 0.000758

ют - 0.000733

пы - 0.000729

гн - 0.000716

иг - 0.000715

яь - 0.000712

ци - 0.000709

рв - 0.000704

гу - 0.0007

нс - 0.0007

ке - 0.000686

ув - 0.000685

нь - 0.000681

тя - 0.000676

хв - 0.00067

ух - 0.000666

хд - 0.000663

ям - 0.000663

тх - 0.000659

кн - 0.00065

еы - 0.000648

рж - 0.000646

мл - 0.000646

бт - 0.000639

ян - 0.000638

жт - 0.000635

кз - 0.000627

пэ - 0.000627

рм - 0.000624

иы - 0.000621

мж - 0.000618

тб - 0.000613

вй - 0.000609

чс - 0.000591

ец - 0.000591

иш - 0.000588

кы - 0.000586

йс - 0.000586

ию - 0.000586

ью - 0.00057

ущ - 0.000562

иц - 0.000556

ащ - 0.000555

щу - 0.000538

юи - 0.000523

шь - 0.000519

ыр - 0.000519

зи - 0.000514

кй - 0.000513

шк - 0.000505

ио - 0.000497

ыш - 0.000491

пя - 0.000489

вп - 0.000489

зл - 0.000487

ьш - 0.00048

тг - 0.000467

бя - 0.000467

тш - 0.000466

ху - 0.000465

ья - 0.000461

яа - 0.000458

нч - 0.000458

зг - 0.000458

ца - 0.000453

вя - 0.000453

дя - 0.000452

шл - 0.00045

мя - 0.000448

зы - 0.000445

пм - 0.000439

оо - 0.000437

сд - 0.000433

бн - 0.000431

йд - 0.000423

йк - 0.00042

вю - 0.000419

сы - 0.000418

ьр - 0.000415

яу - 0.000409

йя - 0.000405

зп - 0.000394

кд - 0.000394

йп - 0.000391

цн - 0.000391

тй - 0.000388

дй - 0.000381

гя - 0.00038

пй - 0.000376

хр - 0.000371

вк - 0.000371

нк - 0.00037

сх - 0.00037

дс - 0.00037

яя - 0.00037

иж - 0.000362

лк - 0.000359

ыб - 0.000358

ою - 0.000358

лж - 0.000349

йт - 0.000345

дь - 0.000345

йе - 0.000344

йч - 0.000344

вт - 0.000344

уь - 0.000338

ьз - 0.000338

дз - 0.000336

яв - 0.000335

чк - 0.000331

уз - 0.000325

жв - 0.000324

ык - 0.000323

шн - 0.000322

сч - 0.000319

яз - 0.000319

уя - 0.000312

бщ - 0.00031

ящ - 0.000307

рх - 0.000306

бз - 0.000303

хе - 0.000303

щй - 0.000298

уе - 0.000293

пв - 0.000292

ьы - 0.000289

мю - 0.000288

шу - 0.000288

кп - 0.000285

уы - 0.000284

нц - 0.000284

ып - 0.000279

лх - 0.000272

ип - 0.000258

мщ - 0.000254

ыь - 0.000254

дт - 0.000251

аа - 0.00025

юб - 0.00025

йм - 0.00025

рк - 0.00025

шо - 0.000247

оц - 0.000247

тд - 0.000242

фу - 0.000242

сй - 0.000241

вд - 0.000241

зб - 0.000241

юа - 0.00024

нз - 0.00024

ьа - 0.000238

ях - 0.000238

яе - 0.000236

юн - 0.000234

кю - 0.000233

фи - 0.00023

ыч - 0.000229

кч - 0.000228

тл - 0.000226

ьм - 0.000225

сю - 0.000223

ыг - 0.000217

фо - 0.000216

йш - 0.000215

лй - 0.000213

чь - 0.000213

вм - 0.000213

юч - 0.000212

яч - 0.000211

нг - 0.000211

ищ - 0.000211

щт - 0.000208

жо - 0.000208

лц - 0.000206

рг - 0.000204

шй - 0.000203

ыд - 0.000203

мб - 0.000203

кя - 0.000202

рд - 0.000202

яр - 0.0002

шс - 0.000199

пг - 0.000199

щь - 0.000199

юс - 0.000199

зя - 0.000199

яю - 0.000199

мх - 0.000198

пн - 0.000198

сб - 0.000198

лр - 0.000197

дц - 0.000197

мш - 0.000195

юд - 0.000194

уи - 0.000194

хл - 0.000193

цо - 0.000193

гб - 0.000191

юу - 0.000191

бй - 0.000191

нб - 0.00019

вь - 0.000189

мй - 0.000187

яж - 0.000187

юо - 0.000186

иа - 0.000181

йр - 0.00018

бс - 0.000177

ню - 0.000176

бь - 0.000174

чш - 0.000172

ын - 0.000172

юз - 0.000171

оэ - 0.000169

дк - 0.000169

кв - 0.000168

ео - 0.000167

тп - 0.000167

як - 0.000165

чо - 0.000161

уо - 0.000161

зь - 0.00016

мс - 0.00016

ао - 0.000159

рп - 0.000159

щя - 0.000158

пб - 0.000155

пх - 0.000155

щн - 0.000155

зх - 0.000154

оу - 0.000154

цл - 0.000152

щл - 0.000151

юр - 0.000151

аи - 0.000151

жы - 0.000148

шз - 0.000147

хм - 0.000147

ея - 0.000147

бж - 0.000146

яг - 0.000146

хя - 0.000144

гк - 0.000144

тз - 0.000143

лл - 0.000141

шт - 0.000141

шя - 0.000139

кс - 0.000139

аэ - 0.000138

пт - 0.000138

уу - 0.000138

еэ - 0.000137

эи - 0.000137

щм - 0.000137

цч - 0.000137

рш - 0.000137

ыя - 0.000135

еу - 0.000135

лг - 0.000133

тч - 0.00013

уа - 0.00013

щх - 0.000129

тю - 0.000129

еа - 0.000126

мк - 0.000125

зк - 0.000125

щщ - 0.000124

еи - 0.000124

чг - 0.000122

дп - 0.000122

вв - 0.000122

уй - 0.000122

ыы - 0.000121

ыж - 0.00012

пз - 0.000118

юг - 0.000118

яы - 0.000117

юя - 0.000117

ыз - 0.000117

бг - 0.000116

шх - 0.000115

вэ - 0.000115

бю - 0.000115

цы - 0.000115

юь - 0.000113

ьц - 0.000113

ею - 0.000113

ъя - 0.000113

шг - 0.000111

хс - 0.000111

нр - 0.000108

фе - 0.000108

дд - 0.000107

рс - 0.000107

хч - 0.000104

тт - 0.000103

юц - 0.000103

мч - 0.000103

щю - 0.0001

эа - 0.0001

зю - 0.0001

лч - 0.0001

фа - 9.9e-05

нм - 9.9e-05

цю - 9.8e-05

зщ - 9.8e-05

фк - 9.8e-05

ъе - 9.6e-05

рю - 9.6e-05

жк - 9.4e-05

эн - 9.2e-05

бк - 9.1e-05

нх - 9e-05

ьл - 8.9e-05

сж - 8.9e-05

дг - 8.9e-05

мр - 8.7e-05

рч - 8.7e-05

тм - 8.6e-05

сф - 8.6e-05

пк - 8.6e-05

шы - 8.5e-05

ьх - 8.5e-05

гч - 8.5e-05

уц - 8.5e-05

шю - 8.3e-05

рл - 8.3e-05

эл - 8.3e-05

щг - 8.2e-05

эр - 7.9e-05

эк - 7.9e-05

мг - 7.9e-05

яэ - 7.8e-05

оъ - 7.5e-05

ау - 7.5e-05

жз - 7.4e-05

хж - 7.4e-05

дб - 7.4e-05

чы - 7.3e-05

дщ - 7.3e-05

хй - 7.2e-05

жя - 7.2e-05

чб - 7e-05

гй - 6.9e-05

йж - 6.9e-05

бч - 6.6e-05

сш - 6.6e-05

чя - 6.5e-05

чх - 6.5e-05

гю - 6.5e-05

цп - 6.5e-05

мэ - 6.5e-05

эв - 6.5e-05

пп - 6.5e-05

ця - 6.4e-05

ьт - 6.4e-05

мь - 6.4e-05

дх - 6.4e-05

цм - 6.1e-05

мм - 6.1e-05

чй - 6e-05

йх - 6e-05

сц - 6e-05

зз - 5.9e-05

йо - 5.9e-05

мп - 5.7e-05

жж - 5.6e-05

мц - 5.5e-05

зй - 5.3e-05

пж - 5.3e-05

чз - 5.3e-05

мт - 5.3e-05

вч - 5.3e-05

лт - 5.3e-05

жю - 5.2e-05

йц - 5.2e-05

иф - 5.2e-05

хг - 5.1e-05

цр - 4.9e-05

эе - 4.8e-05

хы - 4.8e-05

жб - 4.8e-05

хь - 4.8e-05

фл - 4.8e-05

гы - 4.7e-05

йы - 4.7e-05

кь - 4.7e-05

цв - 4.7e-05

кж - 4.6e-05

пь - 4.6e-05

эу - 4.4e-05

хю - 4.4e-05

нф - 4.4e-05

дч - 4.4e-05

фг - 4.3e-05

щс - 4.3e-05

бд - 4.3e-05

фф - 4.3e-05

эд - 4.3e-05

кэ - 4.3e-05

бх - 4.2e-05

ыи - 4.2e-05

йи - 4.2e-05

ёт - 4.2e-05

щп - 4e-05

дш - 4e-05

кк - 4e-05

эм - 4e-05

шв - 4e-05

фн - 3.9e-05

кц - 3.9e-05

вб - 3.8e-05

эф - 3.8e-05

пц - 3.6e-05

юл - 3.5e-05

вщ - 3.5e-05

рр - 3.5e-05

вг - 3.5e-05

гщ - 3.4e-05

йг - 3.4e-05

ьб - 3.4e-05

бм - 3.4e-05

оф - 3.4e-05

лэ - 3.3e-05

чщ - 3.3e-05

жй - 3.3e-05

лз - 3.3e-05

ьи - 3.3e-05

лп - 3.3e-05

юж - 3.3e-05

вх - 3.3e-05

зж - 3.3e-05

бб - 3.1e-05

шд - 3.1e-05

яф - 3.1e-05

лб - 3.1e-05

жь - 3.1e-05

йф - 3e-05

пщ - 3e-05

эх - 3e-05

юю - 3e-05

яй - 3e-05

бв - 3e-05

нш - 3e-05

сг - 3e-05

бш - 3e-05

ьж - 2.9e-05

гп - 2.9e-05

зъ - 2.9e-05

еф - 2.7e-05

жш - 2.7e-05

мё - 2.7e-05

эп - 2.7e-05

пс - 2.7e-05

пч - 2.7e-05

рц - 2.7e-05

ёа - 2.6e-05

яц - 2.6e-05

тц - 2.6e-05

нэ - 2.6e-05

аф - 2.6e-05

ъд - 2.5e-05

чю - 2.5e-05

цс - 2.5e-05

ыэ - 2.5e-05

рз - 2.5e-05

шм - 2.5e-05

шр - 2.5e-05

юэ - 2.3e-05

чё - 2.3e-05

що - 2.3e-05

лф - 2.3e-05

шб - 2.3e-05

ьд - 2.3e-05

зт - 2.3e-05

йа - 2.3e-05

гг - 2.2e-05

жщ - 2.2e-05

ъс - 2.2e-05

кх - 2.2e-05

рб - 2.2e-05

эс - 2.2e-05

яп - 2.2e-05

фы - 2.2e-05

фр - 2.2e-05

ьф - 2.1e-05

гз - 2.1e-05

хщ - 2.1e-05

щк - 2.1e-05

лм - 2.1e-05

вц - 2.1e-05

иу - 2.1e-05

ъм - 2e-05

цй - 2e-05

рэ - 2e-05

ёс - 2e-05

юк - 2e-05

рщ - 2e-05

ёи - 1.8e-05

йэ - 1.8e-05

йё - 1.8e-05

нж - 1.7e-05

чл - 1.7e-05

сз - 1.7e-05

ьщ - 1.6e-05

бп - 1.6e-05

ыц - 1.6e-05

её - 1.6e-05

юш - 1.6e-05

нщ - 1.6e-05

яш - 1.6e-05

кш - 1.6e-05

дю - 1.6e-05

щы - 1.4e-05

хт - 1.4e-05

хп - 1.4e-05

жл - 1.4e-05

дм - 1.4e-05

жг - 1.4e-05

тэ - 1.4e-05

гт - 1.4e-05

цд - 1.3e-05

ъв - 1.3e-05

оа - 1.3e-05

мд - 1.3e-05

гв - 1.3e-05

съ - 1.3e-05

ыщ - 1.3e-05

ьч - 1.3e-05

ёо - 1.2e-05

фс - 1.2e-05

юы - 1.2e-05

дэ - 1.2e-05

вф - 1.2e-05

хш - 1.2e-05

ёб - 1.2e-05

шп - 1.2e-05

жр - 1.2e-05

чж - 1e-05

гц - 1e-05

щв - 1e-05

цх - 1e-05

жп - 1e-05

ём - 1e-05

гс - 1e-05

ьв - 1e-05

чр - 1e-05

тщ - 1e-05

хф - 9e-06

йй - 9e-06

эз - 9e-06

фц - 9e-06

аъ - 9e-06

чд - 9e-06

хк - 9e-06

вж - 9e-06

фз - 8e-06

ът - 8e-06

юф - 8e-06

лё - 8e-06

йщ - 8e-06

кф - 8e-06

ыу - 8e-06

дф - 8e-06

уф - 8e-06

фм - 8e-06

иъ - 7e-06

цг - 7e-06

хц - 7e-06

кё - 7e-06

хэ - 7e-06

шц - 7e-06

юх - 7e-06

цк - 7e-06

лв - 7e-06

чм - 7e-06

ьг - 7e-06

лд - 7e-06

мф - 7e-06

фй - 5e-06

ъз - 5e-06

вё - 5e-06

цф - 5e-06

жх - 5e-06

ър - 5e-06

йв - 5e-06

чв - 5e-06

хз - 5e-06

юв - 5e-06

сщ - 5e-06

тф - 5e-06

зц - 5e-06

рф - 5e-06

жч - 5e-06

йз - 5e-06

ёш - 4e-06

кщ - 4e-06

цш - 4e-06

ёу - 4e-06

зф - 4e-06

цз - 4e-06

цб - 4e-06

шч - 4e-06

ыф - 4e-06

фч - 4e-06

ьп - 4e-06

юм - 4e-06

зш - 4e-06

жё - 4e-06

щр - 4e-06

сэ - 4e-06

йл - 4e-06

шш - 4e-06

жм - 4e-06

цт - 4e-06

уэ - 4e-06

иэ - 4e-06

хх - 4e-06

яо - 4e-06

гм - 4e-06

яб - 4e-06

гх - 4e-06

эг - 4e-06

жс - 4e-06

гэ - 4e-06

ъш - 3e-06

эш - 3e-06

шщ - 3e-06

йю - 3e-06

ьь - 3e-06

ёк - 3e-06

ён - 3e-06

йб - 3e-06

фш - 3e-06

нп - 3e-06

ьэ - 3e-06

фт - 3e-06

ёж - 3e-06

зэ - 3e-06

кб - 3e-06

рё - 3e-06

мв - 3e-06

гш - 3e-06

лщ - 3e-06

дъ - 3e-06

ээ - 3e-06

зч - 3e-06

яи - 3e-06

юе - 3e-06

пш - 3e-06

тж - 3e-06

ъх - 1e-06

ъл - 1e-06

щш - 1e-06

бф - 1e-06

ёэ - 1e-06

аё - 1e-06

бц - 1e-06

въ - 1e-06

юп - 1e-06

лш - 1e-06

ьу - 1e-06

кг - 1e-06

хб - 1e-06

пю - 1e-06

ёз - 1e-06

мз - 1e-06

фя - 1e-06

зс - 1e-06

пф - 1e-06

оё - 1e-06

бэ - 1e-06

нл - 1e-06

йу - 1e-06

Bigrams without space

h2 without space - 4.112855148278171

r2 without space - 0.16571603296123505

то - 0.016059

но - 0.015962

на - 0.013675

ст - 0.013129

ра - 0.012474

по - 0.011972

не - 0.011686

ка - 0.010931

ни - 0.010508

ро - 0.009712

ко - 0.009658

ал - 0.009501

ть - 0.009323

от - 0.009215

та - 0.009139

ен - 0.009005

ре - 0.008767

ли - 0.008652

ак - 0.008572

ет - 0.008374

ос - 0.00826

ва - 0.008142

ер - 0.007925

ат - 0.007908

пр - 0.007869

во - 0.007767

ла - 0.007706

ов - 0.007621

он - 0.007581

да - 0.007293

го - 0.006995

ин - 0.006739

об - 0.006622

ри - 0.006515

ил - 0.006464

ор - 0.006464

ом - 0.006445

те - 0.006421

де - 0.006334

ан - 0.006322

ар - 0.006249

ел - 0.006196

ит - 0.006135

ло - 0.006084

ся - 0.00598

за - 0.005977

ле - 0.0059

ол - 0.005777

од - 0.005481

ру - 0.005364

ти - 0.005358

ме - 0.00535

ки - 0.005229

ма - 0.00517

ас - 0.005096

ши - 0.005084

мо - 0.004831

ог - 0.004815

се - 0.004811

ес - 0.004767

ам - 0.004733

ны - 0.004706

ик - 0.004607

ем - 0.004605

ну - 0.004604

ль - 0.004604

аз - 0.004474

ку - 0.004457

же - 0.004414

тр - 0.004414

ве - 0.004409

ой - 0.004333

до - 0.004309

би - 0.004299

ск - 0.003962

кр - 0.003909

ав - 0.003902

ая - 0.003882

ис - 0.003878

сь - 0.003878

чт - 0.003839

ми - 0.00379

со - 0.003679

уд - 0.00362

ив - 0.003577

пе - 0.00357

вы - 0.003524

сл - 0.003479

ча - 0.003449

ад - 0.003444

нн - 0.003412

ут - 0.00338

бы - 0.003375

им - 0.003344

са - 0.003305

чи - 0.003302

ед - 0.0033

вс - 0.003279

ры - 0.003166

ок - 0.003131

ви - 0.003111

из - 0.00305

аь - 0.003032

ег - 0.002947

че - 0.002836

жи - 0.002813

ож - 0.002722

бе - 0.002653

сп - 0.002624

ул - 0.002583

аг - 0.002575

тв - 0.002566

ей - 0.002566

ук - 0.00253

бо - 0.002523

ир - 0.002476

ше - 0.002412

эт - 0.002386

ие - 0.002366

га - 0.002363

ыв - 0.002342

ще - 0.002283

ду - 0.002262

их - 0.002219

ев - 0.002217

иь - 0.002208

хо - 0.002179

цу - 0.002174

ус - 0.00217

ня - 0.002165

оч - 0.002128

ьн - 0.002092

си - 0.002076

вн - 0.002075

тс - 0.002072

бу - 0.002072

оз - 0.002062

ид - 0.002044

ге - 0.002037

ят - 0.002023

ун - 0.002015

ыл - 0.002003

ди - 0.001993

па - 0.001947

ох - 0.001929

ай - 0.001917

хи - 0.0019

ля - 0.001886

ты - 0.001881

ия - 0.001843

ря - 0.00183

ту - 0.001825

еч - 0.00182

еь - 0.001786

аю - 0.001774

оп - 0.001773

ае - 0.001765

лс - 0.001764

ба - 0.001704

гл - 0.001688

ое - 0.001682

жа - 0.001674

аж - 0.001665

ый - 0.001654

ех - 0.001638

уж - 0.00163

мн - 0.001615

дж - 0.001605

ум - 0.001592

еб - 0.001579

зн - 0.001571

уч - 0.001571

му - 0.001563

еп - 0.001544

ез - 0.00152

тн - 0.001519

ее - 0.001506

ьс - 0.001484

дн - 0.001473

ыа - 0.001468

зе - 0.001442

ур - 0.001441

мы - 0.001438

ач - 0.001433

вз - 0.001433

щи - 0.001425

оь - 0.001415

ых - 0.001414

жн - 0.001406

ич - 0.001395

бр - 0.001388

ые - 0.001388

ую - 0.001373

нв - 0.00136

ек - 0.001336

ьо - 0.001326

нй - 0.001317

вр - 0.001317

см - 0.001311

ах - 0.001304

хн - 0.001296

рь - 0.001294

зу - 0.00128

пл - 0.00128

лу - 0.001273

чн - 0.001247

зд - 0.001246

гр - 0.001218

оы - 0.001212

лы - 0.001204

км - 0.001201

ий - 0.00119

зо - 0.001187

ац - 0.001186

ющ - 0.001178

аб - 0.001174

св - 0.00116

ещ - 0.001148

йн - 0.001138

пд - 0.001132

ым - 0.001132

кл - 0.001129

вл - 0.001127

ха - 0.001122

эо - 0.001117

ап - 0.001095

ыс - 0.001087

вш - 0.001084

тк - 0.001078

ша - 0.001049

зр - 0.001045

сн - 0.001032

бл - 0.001027

еш - 0.001026

дв - 0.001024

рн - 0.001014

ды - 0.001013

яд - 0.001006

лю - 0.000967

уг - 0.000966

ои - 0.000963

зв - 0.000962

су - 0.000958

рт - 0.000955

пи - 0.000953

пу - 0.000953

ср - 0.000949

др - 0.000949

яс - 0.00094

аы - 0.000935

гд - 0.000916

уш - 0.000909

ощ - 0.000902

ье - 0.000892

це - 0.000889

нд - 0.000884

аш - 0.000884

ыт - 0.000881

ыо - 0.000877

ву - 0.000864

уп - 0.000857

сс - 0.000853

ии - 0.000851

ги - 0.000851

иб - 0.000847

уб - 0.000838

кт - 0.000828

оя - 0.000824

ьк - 0.000817

ял - 0.000815

рй - 0.000814

чу - 0.000804

зм - 0.000793

дл - 0.000778

нт - 0.000774

еж - 0.000769

ща - 0.000768

жу - 0.000765

лн - 0.000764

ош - 0.000763

эй - 0.00076

жд - 0.000758

ют - 0.000733

пы - 0.000729

гн - 0.000716

иг - 0.000715

яь - 0.000712

ци - 0.000709

рв - 0.000704

гу - 0.0007

нс - 0.0007

ке - 0.000686

ув - 0.000685

нь - 0.000681

тя - 0.000676

хв - 0.00067

ух - 0.000666

хд - 0.000663

ям - 0.000663

тх - 0.000659

кн - 0.00065

еы - 0.000648

рж - 0.000646

мл - 0.000646

бт - 0.000639

ян - 0.000638

жт - 0.000635

кз - 0.000627

пэ - 0.000627

рм - 0.000624

иы - 0.000621

мж - 0.000618

тб - 0.000613

вй - 0.000609

чс - 0.000591

ец - 0.000591

иш - 0.000588

кы - 0.000586

йс - 0.000586

ию - 0.000586

ью - 0.00057

ущ - 0.000562

иц - 0.000556

ащ - 0.000555

щу - 0.000538

юи - 0.000523

шь - 0.000519

ыр - 0.000519

зи - 0.000514

кй - 0.000513

шк - 0.000505

ио - 0.000497

ыш - 0.000491

пя - 0.000489

вп - 0.000489

зл - 0.000487

ьш - 0.00048

тг - 0.000467

бя - 0.000467

тш - 0.000466

ху - 0.000465

ья - 0.000461

яа - 0.000458

нч - 0.000458

зг - 0.000458

ца - 0.000453

вя - 0.000453

дя - 0.000452

шл - 0.00045

мя - 0.000448

зы - 0.000445

пм - 0.000439

оо - 0.000437

сд - 0.000433

бн - 0.000431

йд - 0.000423

йк - 0.00042

вю - 0.000419

сы - 0.000418

ьр - 0.000415

яу - 0.000409

йя - 0.000405

зп - 0.000394

кд - 0.000394

йп - 0.000391

цн - 0.000391

тй - 0.000388

дй - 0.000381

гя - 0.00038

пй - 0.000376

хр - 0.000371

вк - 0.000371

нк - 0.00037

сх - 0.00037

дс - 0.00037

яя - 0.00037

иж - 0.000362

лк - 0.000359

ыб - 0.000358

ою - 0.000358

лж - 0.000349

йт - 0.000345

дь - 0.000345

йе - 0.000344

йч - 0.000344

вт - 0.000344

уь - 0.000338

ьз - 0.000338

дз - 0.000336

яв - 0.000335

чк - 0.000331

уз - 0.000325

жв - 0.000324

ык - 0.000323

шн - 0.000322

сч - 0.000319

яз - 0.000319

уя - 0.000312

бщ - 0.00031

ящ - 0.000307

рх - 0.000306

бз - 0.000303

хе - 0.000303

щй - 0.000298

уе - 0.000293

пв - 0.000292

ьы - 0.000289

мю - 0.000288

шу - 0.000288

кп - 0.000285

уы - 0.000284

нц - 0.000284

ып - 0.000279

лх - 0.000272

ип - 0.000258

мщ - 0.000254

ыь - 0.000254

дт - 0.000251

аа - 0.00025

юб - 0.00025

йм - 0.00025

рк - 0.00025

шо - 0.000247

оц - 0.000247

тд - 0.000242

фу - 0.000242

сй - 0.000241

вд - 0.000241

зб - 0.000241

юа - 0.00024

нз - 0.00024

ьа - 0.000238

ях - 0.000238

яе - 0.000236

юн - 0.000234

кю - 0.000233

фи - 0.00023

ыч - 0.000229

кч - 0.000228

тл - 0.000226

ьм - 0.000225

сю - 0.000223

ыг - 0.000217

фо - 0.000216

йш - 0.000215

лй - 0.000213

чь - 0.000213

вм - 0.000213

юч - 0.000212

яч - 0.000211

нг - 0.000211

ищ - 0.000211

щт - 0.000208

жо - 0.000208

лц - 0.000206

рг - 0.000204

шй - 0.000203

ыд - 0.000203

мб - 0.000203

кя - 0.000202

рд - 0.000202

яр - 0.0002

шс - 0.000199

пг - 0.000199

щь - 0.000199

юс - 0.000199

зя - 0.000199

яю - 0.000199

мх - 0.000198

пн - 0.000198

сб - 0.000198

лр - 0.000197

дц - 0.000197

мш - 0.000195

юд - 0.000194

уи - 0.000194

хл - 0.000193

цо - 0.000193

гб - 0.000191

юу - 0.000191

бй - 0.000191

нб - 0.00019

вь - 0.000189

мй - 0.000187

яж - 0.000187

юо - 0.000186

иа - 0.000181

йр - 0.00018

бс - 0.000177

ню - 0.000176

бь - 0.000174

чш - 0.000172

ын - 0.000172

юз - 0.000171

оэ - 0.000169

дк - 0.000169

кв - 0.000168

ео - 0.000167

тп - 0.000167

як - 0.000165

чо - 0.000161

уо - 0.000161

зь - 0.00016

мс - 0.00016

ао - 0.000159

рп - 0.000159

щя - 0.000158

пб - 0.000155

пх - 0.000155

щн - 0.000155

зх - 0.000154

оу - 0.000154

цл - 0.000152

щл - 0.000151

юр - 0.000151

аи - 0.000151

жы - 0.000148

шз - 0.000147

хм - 0.000147

ея - 0.000147

бж - 0.000146

яг - 0.000146

хя - 0.000144

гк - 0.000144

тз - 0.000143

лл - 0.000141

шт - 0.000141

шя - 0.000139

кс - 0.000139

аэ - 0.000138

пт - 0.000138

уу - 0.000138

еэ - 0.000137

эи - 0.000137

щм - 0.000137

цч - 0.000137

рш - 0.000137

ыя - 0.000135

еу - 0.000135

лг - 0.000133

тч - 0.00013

уа - 0.00013

щх - 0.000129

тю - 0.000129

еа - 0.000126

мк - 0.000125

зк - 0.000125

щщ - 0.000124

еи - 0.000124

чг - 0.000122

дп - 0.000122

вв - 0.000122

уй - 0.000122

ыы - 0.000121

ыж - 0.00012

пз - 0.000118

юг - 0.000118

яы - 0.000117

юя - 0.000117

ыз - 0.000117

бг - 0.000116

шх - 0.000115

вэ - 0.000115

бю - 0.000115

цы - 0.000115

юь - 0.000113

ьц - 0.000113

ею - 0.000113

ъя - 0.000113

шг - 0.000111

хс - 0.000111

нр - 0.000108

фе - 0.000108

дд - 0.000107

рс - 0.000107

хч - 0.000104

тт - 0.000103

юц - 0.000103

мч - 0.000103

щю - 0.0001

эа - 0.0001

зю - 0.0001

лч - 0.0001

фа - 9.9e-05

нм - 9.9e-05

цю - 9.8e-05

зщ - 9.8e-05

фк - 9.8e-05

ъе - 9.6e-05

рю - 9.6e-05

жк - 9.4e-05

эн - 9.2e-05

бк - 9.1e-05

нх - 9e-05

ьл - 8.9e-05

сж - 8.9e-05

дг - 8.9e-05

мр - 8.7e-05

рч - 8.7e-05

тм - 8.6e-05

сф - 8.6e-05

пк - 8.6e-05

шы - 8.5e-05

ьх - 8.5e-05

гч - 8.5e-05

уц - 8.5e-05

шю - 8.3e-05

рл - 8.3e-05

эл - 8.3e-05

щг - 8.2e-05

эр - 7.9e-05

эк - 7.9e-05

мг - 7.9e-05

яэ - 7.8e-05

оъ - 7.5e-05

ау - 7.5e-05

жз - 7.4e-05

хж - 7.4e-05

дб - 7.4e-05

чы - 7.3e-05

дщ - 7.3e-05

хй - 7.2e-05

жя - 7.2e-05

чб - 7e-05

гй - 6.9e-05

йж - 6.9e-05

бч - 6.6e-05

сш - 6.6e-05

чя - 6.5e-05

чх - 6.5e-05

гю - 6.5e-05

цп - 6.5e-05

мэ - 6.5e-05

эв - 6.5e-05

пп - 6.5e-05

ця - 6.4e-05

ьт - 6.4e-05

мь - 6.4e-05

дх - 6.4e-05

цм - 6.1e-05

мм - 6.1e-05

чй - 6e-05

йх - 6e-05

сц - 6e-05

зз - 5.9e-05

йо - 5.9e-05

мп - 5.7e-05

жж - 5.6e-05

мц - 5.5e-05

зй - 5.3e-05

пж - 5.3e-05

чз - 5.3e-05

мт - 5.3e-05

вч - 5.3e-05

лт - 5.3e-05

жю - 5.2e-05

йц - 5.2e-05

иф - 5.2e-05

хг - 5.1e-05

цр - 4.9e-05

эе - 4.8e-05

хы - 4.8e-05

жб - 4.8e-05

хь - 4.8e-05

фл - 4.8e-05

гы - 4.7e-05

йы - 4.7e-05

кь - 4.7e-05

цв - 4.7e-05

кж - 4.6e-05

пь - 4.6e-05

эу - 4.4e-05

хю - 4.4e-05

нф - 4.4e-05

дч - 4.4e-05

фг - 4.3e-05

щс - 4.3e-05

бд - 4.3e-05

фф - 4.3e-05

эд - 4.3e-05

кэ - 4.3e-05

бх - 4.2e-05

ыи - 4.2e-05

йи - 4.2e-05

ёт - 4.2e-05

щп - 4e-05

дш - 4e-05

кк - 4e-05

эм - 4e-05

шв - 4e-05

фн - 3.9e-05

кц - 3.9e-05

вб - 3.8e-05

эф - 3.8e-05

пц - 3.6e-05

юл - 3.5e-05

вщ - 3.5e-05

рр - 3.5e-05

вг - 3.5e-05

гщ - 3.4e-05

йг - 3.4e-05

ьб - 3.4e-05

бм - 3.4e-05

оф - 3.4e-05

лэ - 3.3e-05

чщ - 3.3e-05

жй - 3.3e-05

лз - 3.3e-05

ьи - 3.3e-05

лп - 3.3e-05

юж - 3.3e-05

вх - 3.3e-05

зж - 3.3e-05

бб - 3.1e-05

шд - 3.1e-05

яф - 3.1e-05

лб - 3.1e-05

жь - 3.1e-05

йф - 3e-05

пщ - 3e-05

эх - 3e-05

юю - 3e-05

яй - 3e-05

бв - 3e-05

нш - 3e-05

сг - 3e-05

бш - 3e-05

ьж - 2.9e-05

гп - 2.9e-05

зъ - 2.9e-05

еф - 2.7e-05

жш - 2.7e-05

мё - 2.7e-05

эп - 2.7e-05

пс - 2.7e-05

пч - 2.7e-05

рц - 2.7e-05

ёа - 2.6e-05

яц - 2.6e-05

тц - 2.6e-05

нэ - 2.6e-05

аф - 2.6e-05

ъд - 2.5e-05

чю - 2.5e-05

цс - 2.5e-05

ыэ - 2.5e-05

рз - 2.5e-05

шм - 2.5e-05

шр - 2.5e-05

юэ - 2.3e-05

чё - 2.3e-05

що - 2.3e-05

лф - 2.3e-05

шб - 2.3e-05

ьд - 2.3e-05

зт - 2.3e-05

йа - 2.3e-05

гг - 2.2e-05

жщ - 2.2e-05

ъс - 2.2e-05

кх - 2.2e-05

рб - 2.2e-05

эс - 2.2e-05

яп - 2.2e-05

фы - 2.2e-05

фр - 2.2e-05

ьф - 2.1e-05

гз - 2.1e-05

хщ - 2.1e-05

щк - 2.1e-05

лм - 2.1e-05

вц - 2.1e-05

иу - 2.1e-05

ъм - 2e-05

цй - 2e-05

рэ - 2e-05

ёс - 2e-05

юк - 2e-05

рщ - 2e-05

ёи - 1.8e-05

йэ - 1.8e-05

йё - 1.8e-05

нж - 1.7e-05

чл - 1.7e-05

сз - 1.7e-05

ьщ - 1.6e-05

бп - 1.6e-05

ыц - 1.6e-05

её - 1.6e-05

юш - 1.6e-05

нщ - 1.6e-05

яш - 1.6e-05

кш - 1.6e-05

дю - 1.6e-05

щы - 1.4e-05

хт - 1.4e-05

хп - 1.4e-05

жл - 1.4e-05

дм - 1.4e-05

жг - 1.4e-05

тэ - 1.4e-05

гт - 1.4e-05

цд - 1.3e-05

ъв - 1.3e-05

оа - 1.3e-05

мд - 1.3e-05

гв - 1.3e-05

съ - 1.3e-05

ыщ - 1.3e-05

ьч - 1.3e-05

ёо - 1.2e-05

фс - 1.2e-05

юы - 1.2e-05

дэ - 1.2e-05

вф - 1.2e-05

хш - 1.2e-05

ёб - 1.2e-05

шп - 1.2e-05

жр - 1.2e-05

чж - 1e-05

гц - 1e-05

щв - 1e-05

цх - 1e-05

жп - 1e-05

ём - 1e-05

гс - 1e-05

ьв - 1e-05

чр - 1e-05

тщ - 1e-05

хф - 9e-06

йй - 9e-06

эз - 9e-06

фц - 9e-06

аъ - 9e-06

чд - 9e-06

хк - 9e-06

вж - 9e-06

фз - 8e-06

ът - 8e-06

юф - 8e-06

лё - 8e-06

йщ - 8e-06

кф - 8e-06

ыу - 8e-06

дф - 8e-06

уф - 8e-06

фм - 8e-06

иъ - 7e-06

цг - 7e-06

хц - 7e-06

кё - 7e-06

хэ - 7e-06

шц - 7e-06

юх - 7e-06

цк - 7e-06

лв - 7e-06

чм - 7e-06

ьг - 7e-06

лд - 7e-06

мф - 7e-06

фй - 5e-06

ъз - 5e-06

вё - 5e-06

цф - 5e-06

жх - 5e-06

ър - 5e-06

йв - 5e-06

чв - 5e-06

хз - 5e-06

юв - 5e-06

сщ - 5e-06

тф - 5e-06

зц - 5e-06

рф - 5e-06

жч - 5e-06

йз - 5e-06

ёш - 4e-06

кщ - 4e-06

цш - 4e-06

ёу - 4e-06

зф - 4e-06

цз - 4e-06

цб - 4e-06

шч - 4e-06

ыф - 4e-06

фч - 4e-06

ьп - 4e-06

юм - 4e-06

зш - 4e-06

жё - 4e-06

щр - 4e-06

сэ - 4e-06

йл - 4e-06

шш - 4e-06

жм - 4e-06

цт - 4e-06

уэ - 4e-06

иэ - 4e-06

хх - 4e-06

яо - 4e-06

гм - 4e-06

яб - 4e-06

гх - 4e-06

эг - 4e-06

жс - 4e-06

гэ - 4e-06

ъш - 3e-06

эш - 3e-06

шщ - 3e-06

йю - 3e-06

ьь - 3e-06

ёк - 3e-06

ён - 3e-06

йб - 3e-06

фш - 3e-06

нп - 3e-06

ьэ - 3e-06

фт - 3e-06

ёж - 3e-06

зэ - 3e-06

кб - 3e-06

рё - 3e-06

мв - 3e-06

гш - 3e-06

лщ - 3e-06

дъ - 3e-06

ээ - 3e-06

зч - 3e-06

яи - 3e-06

юе - 3e-06

пш - 3e-06

тж - 3e-06

ъх - 1e-06

ъл - 1e-06

щш - 1e-06

бф - 1e-06

ёэ - 1e-06

аё - 1e-06

бц - 1e-06

въ - 1e-06

юп - 1e-06

лш - 1e-06

ьу - 1e-06

кг - 1e-06

хб - 1e-06

пю - 1e-06

ёз - 1e-06

мз - 1e-06

фя - 1e-06

зс - 1e-06

пф - 1e-06

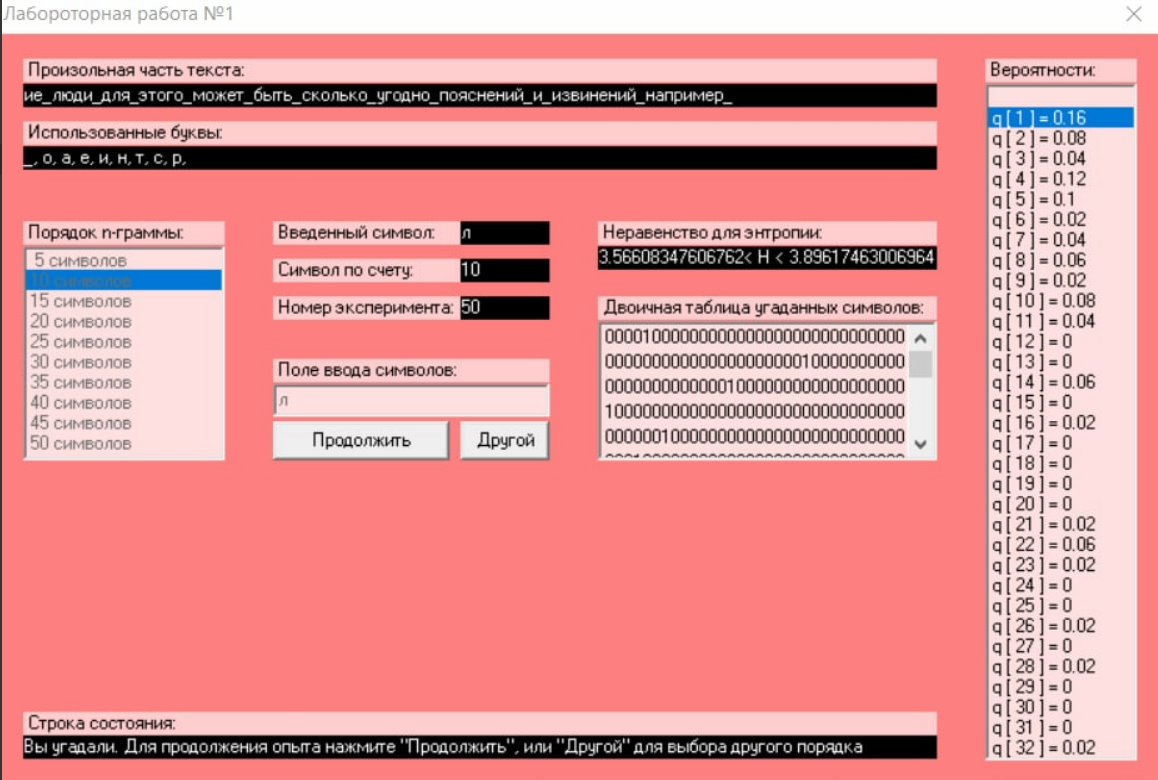
оё - 1e-06

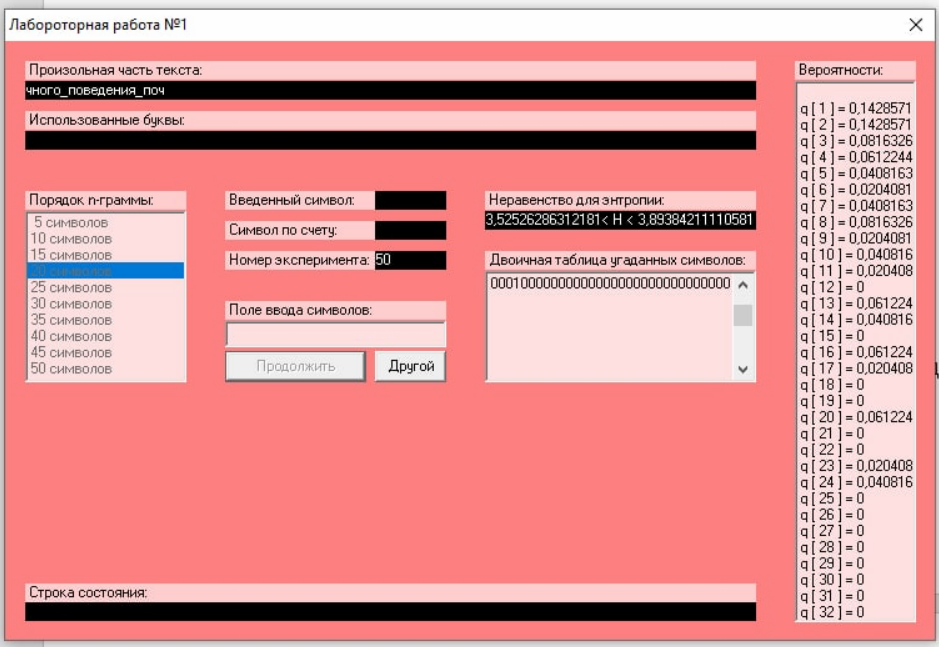
бэ - 1e-06

нл - 1e-06

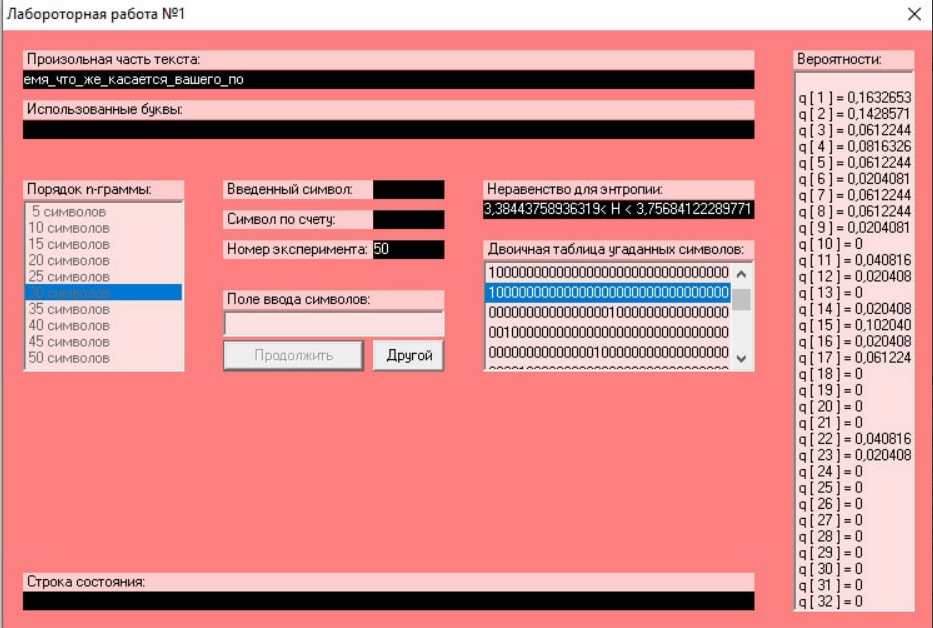
йу - 1e-06

**Результати експериментів у СoolPinkProgram**

n=10:

n=20:

n=30:



**Висновки**

Під час виконання даної лабораторної роботи, ми здобули навички з аналізу тексту. Використовуючи regex нам вдалося в досить простій формі отримати працююче рішення. Крім того, нами було засвоєно визначення ентропїї та надлишковості, які ми інтегрували в наш розв’язок. В результаті ми перевірили статистику тексту за допомогою CoolPinkProgram. На нашу думку, дані навички стануть у нагоді у подальшому розвитку.