

Міністерство освіти і науки України Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського" Фізико-технічний інститут

# Криптографія

Лабораторна №1

Виконали: Студенти групи ФБ-82 **Козачок Вячеслав Ілля Кузнєцов** Перевірив:

## 1 Порядок виконання роботи

- 1. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
- 2. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку Н1 та Н2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення Н1 та Н2 на довільно обраному тексті російською мовою достатньюї довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення Н1 та Н2 на тому ж тексті, в якому вилучено всі пробіли.
- 3. За допомогою програми CoolPinkProgram оцінити значення H(10), H(20), H(30)
- 4. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

## 2 Методичні вказівки

Звичайні текстові файли містять багато символів окрім власне літер; для обчислення значень ентропій вони повинні пройти попередню фільтрацію: всі символи, окрім текстових, повинні вилучатись або замінюватись на пробіли; прописні літери - замінюватись на відповідні стрічні; послідовність пробілів (або інших розділових знаків, наприклад, символів кінця рядку) повинна трактуватись як один пробіл або вилучатись, якщо пробіл не входить до алфавіту.

При підрахунку частот біграм треба розглядати як пари букв, що перетинаються, так і пари букв, що не перетинаються (тобто рухатися вздовж тексту з кроком 2). Одержані результати не повинні суттєво відрізнятись, однак в першому випадку використовується більше статистики, а тому чисельні дані більш точні. Таблицю частот символів потрібно подавати відсортованою за спаданням частот. Таблицю частот біграм зручно подавати у вигляді квадратної матриці, індексованої першою та другою літерами біграм.

Програма CoolPinkProgram використовує текст, що лежить у допоміжному файлі text. Цей текст написаний російською мовою без знаків пунктуації та великих літер; буква «ё» замінена буквою «е», а «ъ» — буквою «ь». Пробіл також вважається буквою. Таким чином, кількість букв алфавіту m=32. При підрахунку H(10), H(20), H(30) виконати не менш ніж 50 експериментів.

## 3 Частота

### 3.1 Монограмы

#### 3.1.1 Монограми без пробіла

count percentage o: 162385, 11.50640% e: 123629, 8.76020% a: 117081, 8.29622% 98121, 6.95273% 93860, 6.65080% 84627, 5.99656% T: 75105, 5.32185% 70907, 5.02438% л: 66549, 4.71558% в: 56282, 3.98807% p: 48456, 3.43353% ĸ: 41627, 2.94964% д: 40517, 2.87098% M: 38126, 2.70156% у: 34091, 2.41565% п: 30439, 2.15687% я: 27849, 1.97335% ь: 26210, 1.85721% ы: г: 25689, 1.82029% 24717, 1.75142% б: 23851, 1.69005% ч: 23117, 1.63804% 16018, 1.13502% ж: 14860, 1.05296% й: 12067, 0.85505% ш: 10981, 0.77810% x: 8811, 0.62434% ю: 5018, 0.35557% э: щ: 4054, 0.28726% 3992, 0.28287% ц: 1779, 0.12606% ф: 412, 0.02919% ъ: 31, 0.00220% ë:

Enthropy: 4.44178288163354 Redundancy 0.8654005187383775

---

#### 3.1.2 Монограми з пробілом

count percentage : 281670, 16.63804% o: 162385, 9.59196% e: 123629, 7.30267% a: 117081, 6.91589% 98121, 5.79593% 93860, 5.54424% 84627, 4.99885% 75105, 4.43640% c: 70907, 4.18842% л: 66549, 3.93100% в: 56282, 3.32454% p: 48456, 2.86226% ĸ: 41627, 2.45888% д: 40517, 2.39331% 38126, 2.25207% у: 34091, 2.01373% п: 30439, 1.79801% я: 27849, 1.64502% 26210, 1.54821% 25689, 1.51743% г: 24717, 1.46001% б: 23851, 1.40886% ч: 23117, 1.36550% з: 16018, 0.94617% ж: й: 14860, 0.87777% 12067, 0.71279% ш: 10981, 0.64864% x:8811, 0.52046% ю: э: 5018, 0.29641% 4054, 0.23947% щ: 3992, 0.23580% ц: 1779, 0.10508% ф: 412, 0.02434% ъ: ë: 31, 0.00183%

Enthropy: 4.3521146215748905, Redundancy: 0.8719966287772091

--

#### 3.2 Биграмы

#### 3.2.1 Біграми з пробілами та з перетинами

--- Bigrams with spaces intersected ---. а б в г д е ж з и й к л м н о п р с т у ф х ц ч ш щ ъ м ь 0 5329 11281 24995 5815 12319 8813 4276 6949 18705 4 15266 6114 9355 27479 22801 23723 6520 27214 13249 6991 839 2587 438 10574 1181 172 0 1 1 52 3727  $a \quad 31521 \qquad 22 \quad 993 \quad 5119 \quad 987 \quad 3643 \quad 1723 \quad 1881 \quad 6997 \quad 214 \quad 885 \quad 7006 \quad 16129 \quad 4038 \quad 7774 \qquad 11 \quad 996 \quad 3934 \quad 5898 \quad 7011 \quad 86997 \quad 1881 \quad 18$ 429 1416 202 1312 1067 340 0 9 4 1301 2 893 7589 14 416 187 6943 1399 2 143 9 54 3344 0 400 1344 96 502 3306 0 1838 134 13 1631 0 91 10 10 9147 9176 27 37 53 522 7168 0 252 1302 242 2122 11765 344 1780 5806 384 1103 39 131 1781 13 7 4011 315 r 1153 1402 1 12 4 1921 599 8 1 1258 0 199 2516 1 289 14057 0 1143 49 11 984 2 1 0 63 9 0 л 1615 7082 37 1274 7 64 7549 23 2 3425 0 349 1332 116 2549 6188 108 2475 566 311 3001 0 102 293 45 170 1 97 887 1288 e 31684 87 2154 4193 6130 4091 3617 1322 1952 162 4193 2937 9643 6711 12246 323 2254 8757 6450 7643 152 18 1313 269 1506 1552 1019 0 0 0 w 463 2322 73 0 9 1285 6803 19 0 2447 0 150 6 4 1660 74 0 1 159 0 405 67 0 0 1 0 0 Ω 2 3 1698 8939 267 1349 865 1284 397 175 98 506 0 158 308 463 3255 599 35 340 87 22 434 0 0 7 92 24 637 266 11 0 w 30604 158 644 4455 697 2583 4047 523 2995 1030 2238 2655 7411 4668 6909 154 233 897 3591 6398 17 23 2190 1366 3148 655 248 0 0 1 417 2905 0 8 0 0 340 2 0 1 0 0 108 144 88 461 3 0 4 832 424 0 70 118 87 1 ₩ 12162 7 0 0 0 ¥ 7840 12951 0 281 0 8 839 23 7 4794 2 1039 13140 1 2315 1552 692 2360 12 0 18 0 29 549 0 0 0 0 1557 5912 0 1684 2905 т 12120 12771 29 6 195 23 8789 630 5 8843 0 474 425 4 495 9707 71 0 2023 41 1986 1 194 6 1 m 11762 4153 21 6 48 3 4966 1 5 4413 0 88 153 70 2121 5263 95 121 183 90 4576 2 0 16 31 8 0 1550 116 0 5 644 H 9183 18415 22 14 170 1062 16580 7 23 12993 0 422 0 3 4878 16319 1 142 1773 705 4069 27 2 580 425 2 264 0 4874 2060 o 39760 13 5768 12304 7740 7046 4158 3177 1707 1198 5141 2592 9211 8104 14006 310 1901 9131 9863 10949 90 272 572 150 2680 2068 247 0 0 0 73 1320 834  $\pi$  71 1927 12 0 0 0 3515 0 0 1360 0 97 765 0 106 13697 40 10400 14 81 858 2 0 24 52 22 5 0 420 59 1105 11537 379 507 599 566 8820 502 26 7389 0 943 83 219 1116 11580 99 27 345 767 3785 18 160 47 113 351 51 0 2050 1437 4860 2985 87 3044 17 527 7352 43 31 2217 0 8476 4242 1370 1587 4040 2806 203 1014 17646 862 13 225 97 715 102 3 38 550 4691 3 208 5024 7393 8319 21 4469 11 195 8392 8 3 6411 0 800 430 35 1803 24151 81 4067 1432 138 1942 35 17 101 458 1 21 58 2471 10409 y 11098 62 873 1853 1624 2535 428 2858 492 5 199 1258 1885 1951 363 6 954 926 1677 1760 3 59 492 6 1451 1160 328 0 0 25 1604 191 64 149 0 0 0 0 204 0 0 521 0 0 41 11 3 163 0 418 18 31 42 12 0 0 0 0 0 0 143 81 164 3708 4641 1367 0 159 2 2 59 0 0 188 1 113 70 16 250 0 15 440 784 1 63 0 0 1122 0 0 321 0 185 1 2 0 457 1 0 1 2 349 0 0 12 0 0 0 0 251 71 8 814 82 1285 3720 0 0 5646 0 0 2161 0 519 0 100 8 7356 1519 0 0 0 0 279 0 106 1511 0 140 0 0 3775 0 0 2539 0 593 903 8 613 465 10 2 0 106 473 2 0 4 0 4 Ω 13 485 0 0 0 2191 0 0 1114 0 0 0 0 53 2 0 9 0 0 148 0 0 0 0 0 0 0 0 0 202 0 0 0 0 0 0 0 0 0 0 0 0 Λ 0 0 12 9 1995 271 4303 1939 548 7124 0 774 1630 149 223 1590 41 103 0 189 463 1279 1188 0 1382 1 292 704 ъ 17616 4 151 48 83 43 812 2 281 681 0 1976 0 360 1969 14 6 0 1370 123 0 4 2 170 101 679 25 646 4 3 0 0 10 0 5 66 71 3 40 0 36 12 7 4748 0 7 396 0 11 6 2 0 6 24 27 46 0 14 67 281 435 0 3 17 18 68 42 554 0 0 41 431 464 889 127 477 526 64 200 180 1358 538 1111 0 92 77 1409 2287 1 0 314 58 203 79 336 0 0 0 0 166 225 я 18786 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 --- Enthropy: 3.945625089661013, Redundancy 0.8804356033436057 ---

#### 3.2.2 Біграми без пробілів та з перетинами

--- Bigrams without spaces intersected ---749 2148 8015 1632 5022 3013 2155 7654 2296 885 8989 16823 4949 10972 2863 3494 4565 8856 8424 840 506 1641 238 2305 1189 352 5 159 13 57 3370 9 7 1363 0 406 1349 101 519 3492 10 1839 153 18 1700 1 92 11 13 15 416 187 6943 259 481 472 1078 7542 91 1035 7926 0 1013 1482 542 2759 12414 1218 2103 6871 1003 1246 24 142 66 385 1853 7 4011 315 414 9291 14 59 141 20 2007 638 21 37 1327 0 236 2530 28 400 14130 163 1174 174 41 1019 5 0 76 12 0 7114 105 1387 42 115 7596 43 57 3547 0 458 1364 157 2757 6282 213 2519 769 393 3033 105 294 108 172 1 97 887 1288 12 671 3671 7071 6770 5625 4323 1726 3098 2268 4193 4265 10414 8546 14811 2639 5184 9571 9522 9089 995 105 1734 302 2407 1671 1038 0 518 92 31 11 1311 6824 21 14 2478 0 170 12 16 1719 108 47 6 192 43 413 0 1 72 1 0 14 1 303 1483 916 1389 444 194 131 568 0 309 343 569 3376 680 197 396 275 128 491 24 637 266 14 126 512 1910 7567 1330 3918 4913 847 3834 2762 2240 4274 8033 5587 9811 2781 3274 1623 6567 7888 886 128 2458 1431 4075 805 277 405 423 3300 0 0 426 1008 311 1020 177 263 253 1121 0 832 407 502 1355 797 1061 449 2039 836 247 66 117 105 480 160 14 0 106 596 0 459 707 287 2022 13774 485 2463 2245 1028 2515 22 128 к 13076 555 810 133 331 1008 349 143 5237 252 44 12 188 23 308 1194 415 362 9563 688 203 9545 0 1167 899 232 1394 11300 842 213 3329 568 2250 34 60 21 925 431 1004 372 532 5151 202 273 5352 0 787 413 385 3115 6227 1113 385 1344 535 4878 38 135 29 743 101 14 0 1550 116 118 0 726 138 201 5971 16831 1040 310 2749 1045 4421 39 135 596 688 37 266 0 4874 2060 501 910 337 1401 16724 97 330 13519 422 7788 15978 8453 8642 5410 4008 2615 3297 5141 4276 9987 9308 17921 4121 5039 9995 13453 13166 1077 377 918 192 4651 2198 269 0 0 781 1325 1577 15 6 0 2 3517 3 2 1369 0 103 768 2 111 13704 41 10401 24 84 858 2 5 560 2 0 24 53 22 5 p 11563 421 591 626 607 8841 507 53 7517 0 1037 98 248 1222 11679 192 48 472 800 3814 26 162 48 133 354 53 0 2050 1437 0 8852 4351 1548 2230 4309 3134 328 1377 18060 1062 26 258 105 783 118 6 c 3087 240 3384 149 778 7452 321 123 2365 38 551 4691 0 1166 592 303 2539 24605 626 4238 2137 498 2098 74 110 746 21 58 2471 10410 206 8462 421 5141 115 513 8692 141 152 6904 74 28 313 1190 2767 1802 3033 646 3011 719 1058 200 1992 2092 2325 1726 754 1794 1128 2643 2283 187 87 584 17 2170 1188 336 0 0 125 1605 351 0 205 0 526 41 13 9 164 3 419 24 34 43 12 0 0 159 5 14 1 0 0 3 0 3 0 1444 189 595 131 971 159 77 109 583 0 262 322 249 522 4096 452 246 534 158 389 17 32 15 87 35 0 212 2 18 109 12 17 1132 4 11 351 15 40 498 46 42 19 363 14 9 2 0 251 0 12 794 10 Ω 1 8 52 2241 0 569 89 25 941 173 160 127 145 7428 1574 0 13 - 1 24 283 3748 38 178 19 51 5661 ш 1513 9 147 3 1 3775 1 2 2547 0 599 905 17 618 470 23 6 7 114 474 2 5 4 4 ш 485 0 2 0 2191 0 0 1114 0 0 2 1 54 3 3 9 2 0 148 0 0 Ω 0 0 0 0 202 1 0 0 Ο 0 0 0 0 - 1 0 0 0 0 0 Ω 0 0 0 0 107 1125 2272 296 560 1763 128 362 609 1995 625 4427 2179 1435 435 853 580 1987 1482 149 18 1481 11 448 722 5 238 598 1689 327 710 1776 164 720 2013 0 3010 332 887 3603 1379 1372 375 3156 937 403 50 157 191 951 741 31 0 0 430 646 963 4 3 0 0 10 0 5 66 71 3 42 0 36 12 7 4748 0 3 0 0 145 1180 383 135 634 156 188 135 494 0 360 168 212 454 465 434 274 795 666 174 47 55 37 339 94 562 0 0 88 484 710 2092 841 1637 771 750 989 1561 201 1232 1745 1079 3672 1206 1541 470 3211 3111 378 58 479 92 795 140 346 0 0 0 262 169 417 1 0 1 0 1 2 0 1 0 1 5 1 0 4 0 1 0 2 0 --- Enthropy: 4.123476319796595, Redundancy 0.8750461721273759 ---

#### 3.2.4 Біграми без пробілів без та без перетинів

--- Bigrams without spaces not-intersected --й к п м н о а 6 в г д е ж з и й к л м н о п р с т у ф х ц ч ш щ ъ м ь э ю я 409 1099 4110 828 2488 1511 1086 3757 1149 421 4445 8445 2484 5450 1419 1754 2272 4474 4248 427 254 817 114 1138 560 183 0 0 0 303 692 2318 724 1 75 5 29 1648 4 3 705 0 199 676 47 260 1766 5 878 69 9 879 1 44 8 6 6 180 99 3431 6 66 5 412 B 4516 125 231 230 521 3660 50 520 4017 0 509 722 289 1389 6165 616 1040 3458 506 631 14 69 35 178 931 6 2 2022 175 207 681 26 83 11 1013 324 8 18 680 0 131 1254 17 217 7129 83 611 90 22 518 3 2 0 33 6 0 ٥ 0 0 q д 3498 44 690 22 55 3838 27 23 1755 0 223 666 66 1403 3247 104 1296 371 190 1502 5 48 144 45 86 1 48 467 661 223 1870 3525 3345 2853 2142 837 1549 1175 2133 2101 5203 4247 7357 1297 2618 4817 4792 4438 494 51 872 147 1223 816 515 0 0 0 281 326 475 46 14 6 640 3385 12 5 1239 0 86 5 7 874 54 21 1 76 20 200 0 0 0 25 0 0 56 291 0 142 191 285 1690 311 96 183 137 56 237 4543 165 730 475 723 205 86 59 11 0 8 347 134 944 3840 670 1916 2474 399 1931 1375 1122 2111 3991 2705 4928 1363 1620 804 3271 3930 437 66 1247 721 2030 407 134 0 0 0 193 203 1598 251 0 400 195 246 691 398 495 223 1035 410 122 31 57 46 239 67 8 207 507 148 503 89 125 117 559 52 1 6582 279 400 67 165 517 179 70 2656 0 216 364 147 1022 6904 258 1228 1143 515 1280 6 71 12 140 22 6444 157 613 219 169 4879 362 108 4766 0 567 456 123 659 5635 420 105 1642 280 1134 19 28 7 458 31 4 0 779 3027 113 844 1526 2182 197 507 187 270 2617 106 146 2723 0 400 186 189 1589 3182 561 194 676 261 2425 18 70 14 352 48 0 750 58 62 9506 261 437 152 714 8352 52 177 6747 0 379 64 101 3000 8290 533 157 1393 513 2209 18 65 293 351 19 129 0 2499 1046 37 136 1/5/ 217 3991 8012 4128 4327 2793 1993 1331 1705 2628 2107 4863 4596 8929 2022 2501 5008 6679 6538 530 191 452 101 2320 1138 139 0 0 384 666 947 4 0 0 1752 2 1 690 0 46 390 1 49 6932 21 5205 14 45 415 2 0 11 21 14 4 0 209 2 5750 202 270 304 309 4453 251 30 3775 0 512 49 116 609 5914 88 25 212 392 1878 10 86 22 62 174 21 0 1014 726 131 1682 65 376 3727 165 60 1172 0 4389 2155 791 1091 2124 1560 157 688 9026 511 15 117 54 380 54 6 22 275 2413 39 103 2608 226 2645 4194 64 250 4375 86 77 3460 0 597 296 157 1305 12304 327 2121 1037 247 1037 33 34 49 375 12 13 28 1235 5218 107  $601 \quad 1408 \quad 923 \quad 1472 \quad 326 \quad 1555 \quad 352 \quad 516 \quad 100 \quad 1009 \quad 1039 \quad 1173 \quad 865 \quad 396 \quad 851 \quad 584 \quad 1366 \quad 1171 \quad 96$ 38 276 12 1066 614 170 0 60 754 0 108 0 245 0 2 24 4 84 2 219 11 21 22 0 0 47 732 89 259 65 137 73 35 57 304 0 136 166 122 256 2049 219 114 266 71 195 16 33 20 0 18 548 5 181 0 119 14 249 22 24 198 3 0 0 122 65 3 0 20 92 10 25 2854 6 28 1136 0 297 52 10 477 87 96 62 69 3751 787 0 3 12 136 0 0 1890 1 1896 0 1250 0 298 468 318 238 16 3 2 53 227 2 754 79 0 425 2 2 0 1128 0 563 ٥ 0 2 1 27 2 2 0 73 Ω 0 Ο Ω 0 100 0 0 0 0 0 0 Ω 0 1 0 ٥ 0 0 Ω 0 0 0 0 ٥ 563 1116 147 305 823 49 173 308 1017 314 2235 1102 729 210 412 278 971 737 64 8 770 208 368 2 ٥ 296 848 160 343 848 94 333 998 0 1506 167 425 1759 695 674 180 1595 454 188 28 85 89 464 354 13 111 0 0 3 0 5 29 36 3 20 0 19 6 3 2399 0 2 0 0 0 Ο 2 0 0 1 74 560 197 71 335 85 97 70 261 0 183 82 102 232 241 234 139 408 339 84 24 33 18 181 41 272 0 0 45 28 225 345 1018 402 794 377 403 502 776 111 599 889 542 1824 626 779 234 1584 1521 183 29 233 52 389 71 176 0 0 0 130 94 204 0 0 5 1 0 0 0 1 2 0 1 0 0 0 2 0 0 3 0 1 0 2 0 0 0 0 --- Enthropy: 4.122368938570627, Redundancy 0.8750797291342234 ---