Київський національний університет імені Тараса Шевченка

Факультет комп’ютерних наук та кібернетики

Кафедра інтелектуальних програмних систем

Алгоритми та складність

Лабораторна робота №2

«Багатофазне сортування злиттям»

Варіант № 1

Виконав студент 2-го курсу

Групи ІПС21

Вербицький Артем Віталійович

Київ – 2023

**Завдання:**

Реалізуйте багатофазне сортування злиттям.

**Теорія**

Алгоритм сортування — це алгоритм, що розв'язує задачу сортування, тобто здійснює впорядкування лінійного списку (масиву) елементів.

Багатофазне сортування прийнято називати «зовнішнім » сортуванням послідовних файлів, розташованих у зовнішній пам'яті і дуже великих, щоб можна було повністю перемістити їх в основну пам'ять і застосувати один з розглянутих у попередньому розділі методів внутрішнього сортування. Найчастіше зовнішнє сортування застосовується в системах керування базами даних при виконанні запитів, і від ефективності вживаних методів істотно залежить продуктивність СУБД. Швидкість виконання зовнішнього сортування залежить від розміру буферу (або буферів) основної пам'яті, яка може бути використана для цих цілей.

Система управління базами даних (СУБД, СКБД англ. Database Management System, DBMS) — набір взаємопов'язаних даних (база даних) і програм для доступу до цих даних. Надає можливості створення, збереження, оновлення та пошуку інформації в базах даних з контролем доступу до даних.

Принцип «розділяй та владарюй»

* *Поділ* задачі на декілька підзадач.
* *Підкорення* – рекурсивне розв'язання цих підзадач. Коли об'єм підзадачі досить малий, вона вирішується безпосередньо.
* *Комбінування* розв'язку початкової задачі з розв'язків допоміжних задач.

Сортування злиттям (англ. merge sort) — алгоритм сортування, в основі якого лежить принцип «Розділяй та володарюй».

* *Поділ:* послідовність, що сортується і складається з n елементів, розбивається на дві менші по n/2 елементів кожна. При цьому рекурсія закінчується, коли послідовність містить єдиний (і очевидно відсортований) елемент.
* *Підкорення:* сортування обох допоміжних підпослідовностей шляхом злиття.
* *Комбінування:* злиття двох відсортованих підпослідовностей для отримання кінцевого результату.
* Основна операція – об'єднання двох відсортованих послідовностей в ході комбінування (останній етап).
* Вводиться допоміжна процедура *Merge*(A,p,q,r), де А – масив; р, q, r – індекси елементів масиву (p≤q<r). Вважається, що елементи підмасивів A[p..q] та A[(q+1)..r] впорядковані. Процедура зливає їх в один відсортований, елементи якого замінюють поточні елементи підмасиву А[p..r].
* Додаткова ідея: в кінець допоміжних масивів додається сигнальний елемент, який слугуватиме ознакою вичерпання підмасиву, його значення перевищує значення будь-якого іншого елемента. Тоді злиття виконається рівно за (r–p+1) крок.
* Алгоритм має складність Θ(n log n).
* Сортування допускає природне розпаралелення.
* Не має "поганих" вхідних даних, але й відсутній виграш у швидкості роботи для майже відсортованих даних порівняно з середнім випадком.
* Єдиний недолік – використання додаткової пам’яті (пропорційно до кількості вхідних елементів). Можлива реалізація без задіювання додаткової пам’яті, однак вона додає великий множник до часу роботи.

**Алгоритм**

Оцінимо спочатку число записів, які можна зберігати в оперативній пам'яті одночасно. Оголосимо масив, довжина S якого дорівнює цій величині. На першому кроці ми прочитаємо S записів і відсортуємо їх за допомогою внутрішнього сортування. Цей набір вже відсортованих записів перепишемо в файл A. Потім прочитаємо ще S записів, відсортуємо їх і перепишемо в файл B. Цей процес триває, причому відсортовані блоки записів пишуться поперемінно то в файл A, то в файл B.

Після того, як вхідний файл повністю розбитий на відсортовані відрізки, ми починаємо їх зливати. Кожен з файлів A і B містить деяку послідовність відсортованих відрізків, однак, як і в випадку сортування злиттям, ми нічого не можемо сказати про порядок записів в двох різних відрізках.

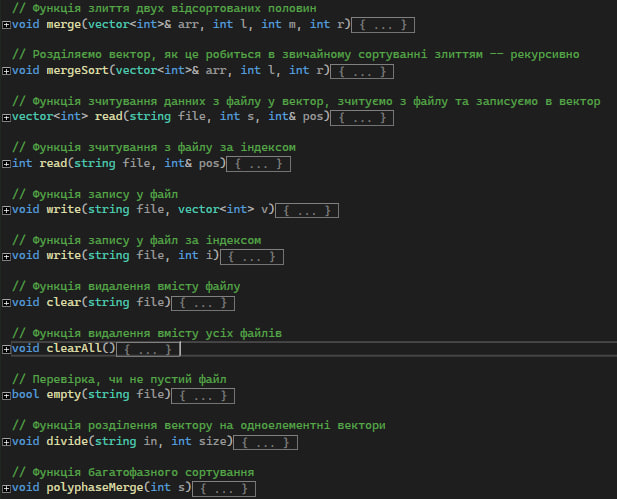
Процес злиття буде аналогічним до звичайного сортування злиттям, проте тепер замість того, щоб переписувати записи в новий масив, вони записуються в новий файл. Тому ми починаємо з читання половинок перших відрізків з файлів A і B. Читаємо ми лише по половині відрізків, оскільки ми вже з'ясували, що в пам'яті може перебувати одночасно лише S записів, а нам потрібні записи з обох файлів. Будемо тепер зливати ці половинки відрізків в один відрізок файлу C. Після того, як одна з половинок закінчиться, ми прочитаємо другу половинку з того ж файлу. Коли обробка одного з відрізків буде завершена, кінець другого відрізка буде переписаний у файл C. Після того, як злиття перших двох відрізків з файлів A і B буде завершено, наступні два відрізки зливаються в файл D. Цей процес злиття відрізків триває з поперемінним записом злитих відрізків в файли C і D. По завершенні ми отримуємо два файли, розбитих на відсортовані відрізки довжини 2S. Потім процес повторюється, причому відрізки зчитуються з файлів C і D, а злиті відрізки довжини 4S записуються в файли A і B. Врешті-решт відрізки зіллються в один відсортований список в одному з файлів.

**Складність алгоритму**

Час роботи має бути рівним , де n — число сортованих елементів, а N — кількість списків в одній операції злиття. І саме є середньою складністю алгоритму.

**Мова реалізації алгоритму** С++

**Модулі програми:**



**Інтерфейс користувача**

Вхідні дані (набір чисел розділених пробілами) вводяться у текстовий файл DataFile, вихідні дані виводяться також у DataFile.

**Тестові приклади**

*Вхідні дані:*

64642 62791 4807 10708 51627 92437 21907 16243 83949 98050 54007 48363 72300 90194 68182 39169 21427 92965 30681 86527 11170 46621 79185 53395 51183 69227 67077 9633 25239 93880 79632 78229 87472 52377 39560 59475 82810 90828 41544 82409 74717 49439 52187 63290 99186 55477 94814 81147 51816 57273 16119 37289 30002 10 5436 12929 92315 43721 73950 46931 15444 10205 79237 88437 10892 9754 94298 79655 91570 80359 45485 23091 11144 83348 74758 98114 46368 63751 13047 22220 52269 39104 14533 69467 14374 90945 36595 17871 50404 94834 23483 34213 88233 8193 6407 53004 98532 35504 40799 65750 88419 51612 85592 83044 61420 77570 85286 71244 59349 38077 91718 93631 62167 49449 26683 86529 9737 74226 20936 39406 72782 80994 25154 68664 43555 37519 79537 85727 31586 86785 52735 21961 70823 67496 63573 17298 6534 67608 16721 42997 15099 65655 82594 58499 83751 69725 53263 46162 56436 6488 86405 72784 94092 90465 29119 73253 97807 69779 46519 16410 38506 19611 12461 98252 79552 41106 38445 72555 70244 41496 49072 44274 33699 71350 62011 58104 33380 77527 6377 14945 75303 64798 43284 49758 83901 79144 15347 96130 50423 1799 71550 79745 46793 54582 33455 54267 86541 31183 70610 39882 53005 77967 70977 38207 64601 74313 95571 12735 17663 45558 87778 6450 98632 75781 69165 7400 16199 13470 24523 4323 88548 42939 55039 1411 75248 29221 47802 40941 79299 13441 29478 21411 33920 16156 85313 53716 43783 76401 15968 18009 16191 73089 26623 87607 8846 12175 4772 58380 2811 51915 978 60568 31686 69971 75342 3951 34616 2951 4337 39311 31759 70615 33188 76527 85617 51382 61193 28654 59408 11006 96571 51613 48603 58069 82527 29519 95897 54711 38986 48753 13493 37757 79608 48556 74303 80495 88533 24686 21947 25591 6759 58919 24451 66001 89266 8429 25423 39877 6662 88079 45311 28342 89404 22662 20235 60279 12045 66171 47297 20415 82358 1807 48277 19225 89346 45344 10261 10912 22183 71227 50914 19771 35583 22721 14861 63624 39612 86692 47373 62851 27171 51010 14982 69322 68305 77919 17252 66880 34919 24984 72607 64976 81308 90977 39658 76012 52478 77077 20612 35145 55459 62168 78240 62338 40117 84180 80774 734 92959 39752 39151 92564 30200 81138 20351 68890 67725 57243 42323 98361 65575 47590 55492 85277 21846 66623 94249 14644 59293 45344 4515 79247 91733 53429 2043 78579 76323 44825 67061 62421 37853 30137 19657 80319 51869 70221 45282 66678 56971 78029 93794 47444 97487 11476 69690 34160 25868 67725 41308 67120 50754 37123 42528 91497 56593 11833 57165 89119 91444 62362 27339 64443 30230 81903 47001 32426 70474 267 68630 99700 21021 24049 53373 38579 68182 72789 76176 90564 79317 40165 75949 90300 99852 88588 58789 92657 70052 37769 58590 34508 39270 1589 19988 10975 57745 1871 86818 56715 90749 21285 68823 53586 65656 77813 40884 81165 21514 64237 22375 66410 36765 10993 51160 77160 18532 706 28008 19236 46198 36916 46228 40245 49277 58299 45393 88740 75037 7501 85210 20723 28122 15818 33209 52167 2183 34470 27699 18543 53436 11556 59822 15133 37117 27038 9000 90324 74162 90427 20311 71813 69086 40118 87406 40052 11392 8789 42369 7068 8942 597 64786 7910 65335 54612 19200 75438 73984 29001 36361 84964 88307 772 24524 26896 97895 53889 60629 39228 99291 82491 67229 77985 33373 59472 38672 88397 12583 40520 90514 60685 24610 45770 27958 24156 28032 55448 31669 33973 73789 32545 48744 66772 10765 84904 32035 31087 72352 22821 9219 79981 93471 79456 35877 20836 96435 84992 26468 17808 61123 19620 78247 90088 59195 25382 38161 57994 41214 42804 76905 20307 23866 10974 70915 99828 69025 97518 66922 4786 97519 23001 91479 32457 69090 65743 6078 46683 2615 98711 46268 95551 24820 85163 34275 22868 67332 45299 1596 24769 28218 93440 5046 52572 20108 34303 92313 76096 29098 76645 79803 3272 8542 60932 65860 26551 68589 16279 54368 2280 20122 98298 46931 44316 37325 68399 59272 10714 46099 38894 99509 64836 69093 54858 81531 7811 79516 60128 90892 32798 36601 95955 80313 49622 48617 39744 72031 76334 6262 16740 33455 20017 21806 81394 15703 26986 90736 72741 33770 75969 63712 76465 85761 33545 8377 64429 2243 37195 33290 71800 49226 94753 24488 3179 80848 7075 96951 46068 60079 76929 11579 73077 4606 44235 22580 69232 9712 10401 24027 62706 77409 37512 53691 45790 69786 34215 99053 1784 11581 5815 40635 57399 92382 22048 47630 7690 9293 47181 38380 43945 75641 70776 15033 93620 47596 14727 83599 51280 46703 13939 54183 73980 59865 59359 96275 73945 25562 59146 72878 76208 81097 68332 38624 38962 36621 63204 66015 70774 41153 96290 1572 26906 79735 53016 644 54613 33784 19969 21279 91752 33289 33803 31231 7044 69434 28562 24871 49900 97785 53788 68830 89853 14152 83216 93207 3735 89158 8734 22357 37533 37750 30576 37723 95054 33341 47046 15390 47665 35174 81870 74127 19102 84508 6842 79836 64104 40826 14892 61567 76126 53657 18164 93046 55966 27334 84043 29742 84722 62647 2423 48187 95641 85087 65300 16586 15377 43015 35022 62530 39361 80968 29358 17941 9577 68360 61776 79691 66627 63692 42956 1488 46402 38501 40936 22809 65835 32614 51664 49076 45405 66516 63724 91673 72474 6192 96280 18971 11858 5079 35237 90355 75699 97400 94152 18115 33643 67862 17620 20276 24594 54903 72369 26310 48148 13571 16723 54839 27170 58756 28860 147 59463 12568 71217 28987 69151 92803 85520 95893 88862 40541 95729 26714 21323 53098 81131 50950 44484 46370 91174 21568 61153 21745 84627 54621 24050 9626 16149 1227 17077 90950 45176 6056 50947 29604 87865 21958 5237 43371 33952 35428 51876 8138 95150 52741 39281 77937 21078 55901 95412 71712 69060 85245 84669 48071 24650 60393 709 4571 51787 11765 24865 20300 57773 38118 25139 17013 56034 63506 73721 90180 42138 52255 85243 90829 52680 23321 7676 94901 12757 56758 651 52451 25525 57040 18739 91740 52856 8912 11972 19525 37745 90486 41176 78489 4324 43888 35222 79037 19851 10146 77193 51512 51235 45468 24089 65801 45988 54398 78880 46184 3183 82698 51527 58379 18131 5352 91853 49183 95925 45590 87926 39185 87790 63643 59045 5202 20545 87710 36748 39989 39637 77469 26857 9166 24349 88082 1419 41099 23108 66265 2424 78609 26281 4043 66234 96305 56307 66671 48797 52902 21613 96158 61937 18454 96821 67926 55127 50316 88549 58100

*Вихідні дані:*

10 147 267 597 644 651 706 709 734 772 978 1227 1411 1419 1488 1572 1589 1596 1784 1799 1807 1871 2043 2183 2243 2280 2423 2424 2615 2811 2951 3179 3183 3272 3735 3951 4043 4323 4324 4337 4515 4571 4606 4772 4786 4807 5046 5079 5202 5237 5352 5436 5815 6056 6078 6192 6262 6377 6407 6450 6488 6534 6662 6759 6842 7044 7068 7075 7400 7501 7676 7690 7811 7910 8138 8193 8377 8429 8542 8734 8789 8846 8912 8942 9000 9166 9219 9293 9577 9626 9633 9712 9737 9754 10146 10205 10261 10401 10708 10714 10765 10892 10912 10974 10975 10993 11006 11144 11170 11392 11476 11556 11579 11581 11765 11833 11858 11972 12045 12175 12461 12568 12583 12735 12757 12929 13047 13441 13470 13493 13571 13939 14152 14374 14533 14644 14727 14861 14892 14945 14982 15033 15099 15133 15347 15377 15390 15444 15703 15818 15968 16119 16149 16156 16191 16199 16243 16279 16410 16586 16721 16723 16740 17013 17077 17252 17298 17620 17663 17808 17871 17941 18009 18115 18131 18164 18454 18532 18543 18739 18971 19102 19200 19225 19236 19525 19611 19620 19657 19771 19851 19969 19988 20017 20108 20122 20235 20276 20300 20307 20311 20351 20415 20545 20612 20723 20836 20936 21021 21078 21279 21285 21323 21411 21427 21514 21568 21613 21745 21806 21846 21907 21947 21958 21961 22048 22183 22220 22357 22375 22580 22662 22721 22809 22821 22868 23001 23091 23108 23321 23483 23866 24027 24049 24050 24089 24156 24349 24451 24488 24523 24524 24594 24610 24650 24686 24769 24820 24865 24871 24984 25139 25154 25239 25382 25423 25525 25562 25591 25868 26281 26310 26468 26551 26623 26683 26714 26857 26896 26906 26986 27038 27170 27171 27334 27339 27699 27958 28008 28032 28122 28218 28342 28562 28654 28860 28987 29001 29098 29119 29221 29358 29478 29519 29604 29742 30002 30137 30200 30230 30576 30681 31087 31183 31231 31586 31669 31686 31759 32035 32426 32457 32545 32614 32798 33188 33209 33289 33290 33341 33373 33380 33455 33455 33545 33643 33699 33770 33784 33803 33920 33952 33973 34160 34213 34215 34275 34303 34470 34508 34616 34919 35022 35145 35174 35222 35237 35428 35504 35583 35877 36361 36595 36601 36621 36748 36765 36916 37117 37123 37195 37289 37325 37512 37519 37533 37723 37745 37750 37757 37769 37853 38077 38118 38161 38207 38380 38445 38501 38506 38579 38624 38672 38894 38962 38986 39104 39151 39169 39185 39228 39270 39281 39311 39361 39406 39560 39612 39637 39658 39744 39752 39877 39882 39989 40052 40117 40118 40165 40245 40520 40541 40635 40799 40826 40884 40936 40941 41099 41106 41153 41176 41214 41308 41496 41544 42138 42323 42369 42528 42804 42939 42956 42997 43015 43284 43371 43555 43721 43783 43888 43945 44235 44274 44316 44484 44825 45176 45282 45299 45311 45344 45344 45393 45405 45468 45485 45558 45590 45770 45790 45988 46068 46099 46162 46184 46198 46228 46268 46368 46370 46402 46519 46621 46683 46703 46793 46931 46931 47001 47046 47181 47297 47373 47444 47590 47596 47630 47665 47802 48071 48148 48187 48277 48363 48556 48603 48617 48744 48753 48797 49072 49076 49183 49226 49277 49439 49449 49622 49758 49900 50316 50404 50423 50754 50914 50947 50950 51010 51160 51183 51235 51280 51382 51512 51527 51612 51613 51627 51664 51787 51816 51869 51876 51915 52167 52187 52255 52269 52377 52451 52478 52572 52680 52735 52741 52856 52902 53004 53005 53016 53098 53263 53373 53395 53429 53436 53586 53657 53691 53716 53788 53889 54007 54183 54267 54368 54398 54582 54612 54613 54621 54711 54839 54858 54903 55039 55127 55448 55459 55477 55492 55901 55966 56034 56307 56436 56593 56715 56758 56971 57040 57165 57243 57273 57399 57745 57773 57994 58069 58100 58100 58104 58299 58379 58380 58499 58590 58756 58789 58919 59045 59146 59195 59272 59293 59349 59359 59408 59463 59472 59475 59822 59865 60079 60128 60279 60393 60568 60629 60685 60932 61123 61153 61193 61420 61567 61776 61937 62011 62167 62168 62338 62362 62421 62530 62647 62706 62791 62851 63204 63290 63506 63573 63624 63643 63692 63712 63724 63751 64104 64237 64429 64443 64601 64642 64786 64798 64836 64976 65300 65335 65575 65655 65656 65743 65750 65801 65835 65860 66001 66015 66171 66234 66265 66410 66516 66623 66627 66671 66678 66772 66880 66922 67061 67077 67120 67229 67332 67496 67608 67725 67725 67862 67926 68182 68182 68305 68332 68360 68399 68589 68630 68664 68823 68830 68890 69025 69060 69086 69090 69093 69151 69165 69227 69232 69322 69434 69467 69690 69725 69779 69786 69971 70052 70221 70244 70474 70610 70615 70774 70776 70823 70915 70977 71217 71227 71244 71350 71550 71712 71800 71813 72031 72300 72352 72369 72474 72555 72607 72741 72782 72784 72789 72878 73077 73089 73253 73721 73789 73945 73950 73980 73984 74127 74162 74226 74303 74313 74717 74758 75037 75248 75303 75342 75438 75641 75699 75781 75949 75969 76012 76096 76126 76176 76208 76323 76334 76401 76465 76527 76645 76905 76929 77077 77160 77193 77409 77469 77527 77570 77813 77919 77937 77967 77985 78029 78229 78240 78247 78489 78579 78609 78880 79037 79144 79185 79237 79247 79299 79317 79456 79516 79537 79552 79608 79632 79655 79691 79735 79745 79803 79836 79981 80313 80319 80359 80495 80774 80848 80968 80994 81097 81131 81138 81147 81165 81308 81394 81531 81870 81903 82358 82409 82491 82527 82594 82698 82810 83044 83216 83348 83599 83751 83901 83949 84043 84180 84508 84627 84669 84722 84904 84964 84992 85087 85163 85210 85243 85245 85277 85286 85313 85520 85592 85617 85727 85761 86405 86527 86529 86541 86692 86785 86818 87406 87472 87607 87710 87778 87790 87865 87926 88079 88082 88233 88307 88397 88419 88437 88533 88548 88549 88588 88740 88862 89119 89158 89266 89346 89404 89853 90088 90180 90194 90300 90324 90355 90427 90465 90486 90514 90564 90736 90749 90828 90829 90892 90945 90950 90977 91174 91444 91479 91497 91570 91673 91718 91733 91740 91752 91853 92313 92315 92382 92437 92564 92657 92803 92959 92965 93046 93207 93440 93471 93620 93631 93794 93880 94092 94152 94249 94298 94753 94814 94834 94901 95054 95150 95412 95551 95571 95641 95729 95893 95897 95925 95955 96130 96158 96275 96280 96290 96305 96435 96571 96821 96951 97400 97487 97518 97519 97785 97807 97895 98050 98114 98252 98298 98361 98532 98632 98711 99053 99186 99291 99509 99700 99828 99852

**Висновки**

Багатофазне сортування злиттям використовується коли ми маємо досить великий об’єм даних, які потрібно відсортувати, проте нам потрібно, щоб витрати в часі були не досить великими.

**Використані літературні джерела**

* Лекція №3
* Кормен, Ч. Лейзерсон, Р. Ривест, К. Штайн. Алгоритмы: построение и анализ. 3-е издание. – М.: ИД "Вильямс", 2013. (с. 53-61).
* <https://ozlib.com/868204/informatika/mnogofaznaya_sortirovka>
* <http://um.co.ua/4/4-16/4-163692.html>
* <https://uk.wikipedia.org/wiki/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0_%D1%83%D0%BF%D1%80%D0%B0%D0%B2%D0%BB%D1%96%D0%BD%D0%BD%D1%8F_%D0%B1%D0%B0%D0%B7%D0%B0%D0%BC%D0%B8_%D0%B4%D0%B0%D0%BD%D0%B8%D1%85>
* <https://uk.wikipedia.org/wiki/%D0%90%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC_%D1%81%D0%BE%D1%80%D1%82%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F>