

Temat 7

Najlepsze i najcięższe – *algorytmy sortowania*

Streszczenie

Komputery są często używane porządkowania różnych danych, na przykład nazwisk (w porządku alfabetycznym), terminów spotkań lub e-maili (według daty), czy po prostu porządkowania wartości liczbowych. Dzięki istnieniu uporządkowania danych można skrócić czas wyszukiwania ich, kiedy są potrzebne. Uzyskujemy też błyskawiczny dostęp do wartości najmniejszej i największej.

Użycie nieefektywnej metody może znacznie wydłużyć czas sortowania dużej ilości danych, nawet jeśli używamy do tego komputera. Na szczęście znanych jest kilka szybkich metod sortowania. W czasie tych zajęć dzieci będą mogły odkryć różne metody sortowania i zorientować się, jak użycie sprytnej metody pozwala skrócić wykonanie zadania.

Wiek

- ✓ 8 lat i więcej

Materiały

Każda grupa dzieci otrzyma:

- ✓ zbiory 8 pojemników tej samej wielkości ale różnej wagi (np. pojemników po kliszy fotograficznej wypełnionych piaskiem)
- ✓ wagę szalkową
- ✓ kartę pracy: Sortowanie na wadze (s. 66)
- ✓ kartę pracy: Dziel i zwyciężaj (s. 67)

Najlżejsze i najcięższe

Dyskusja

Komputerów często używa się do porządkowania różnych zbiorów danych. W jakich sytuacjach życia codziennego ważne jest uporządkowanie jakiegoś zbioru? Co by się stało, gdyby elementy tego zbioru nie były uporządkowane?

Komputery zazwyczaj porównują tylko dwie wartości w tym samym czasie. Zadanie opisane na następnej stronie umożliwi dzieciom zrozumienie takiego ograniczenia.

Zadanie

1. Należy podzielić dzieci na grupy.
2. Każda z grup otrzymuje kartę pracy (s. 66), odważniki i wagę.
3. Dzieci wykonują zadanie i dyskutują o jego rezultatach.

Karta pracy: Sortowanie na wadze

Cel: Odkryć najlepszą metodę sortowania grupy odważników o nieznannej wadze.

Środki: Piasek lub woda, osiem identycznych pojemników, wagi szalkowe.

Kolejne kroki:

1. Napełnij każdy z pojemników różną ilością piasku lub wody. Możesz też użyć odpowiedniej liczby monet. Zamknij dobrze każdy z pojemników.
2. Ustaw odważniki (wypełnione pojemniki) w przypadkowej kolejności.
3. Znajdź najlżejszy odważnik. W jaki sposób można to najłatwiej zrobić?

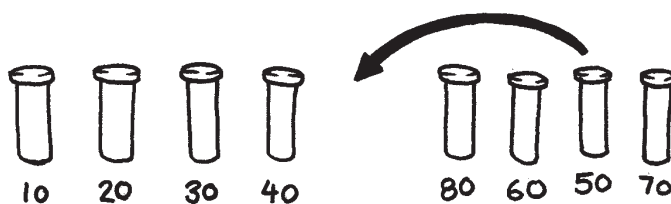
Uwaga: Możesz używać tylko i wyłącznie wagi szalkowej. W danym momencie możesz porównywać wagę tylko dwóch odważników.

4. Wybierz trzy dowolne odważniki i uporządkuj je od najlżejszego do najcięższego, używając tylko i wyłącznie wagi szalkowej. W jaki sposób to zrobić? Ilu co najmniej porównań (wagęń) trzeba wykonać? Dlaczego?
5. Uporządkuj wszystkie odważniki od najlżejszego do najcięższego.

Sprawdź czy uzyskałeś właściwe uporządkowanie, porównując na wadze każdą parę stojących obok siebie odważników.

Sortowanie przez wybór

Jedną z metod sortowania, która może być używana przez komputer jest *sortowanie przez wybór*. Działa ona w następujący sposób: Znajdujemy najpierw najlżejszy odważnik i odkładamy go. Następnie znajdujemy najlżejszy z tych, które zostały i kładziemy obok wcześniej znalezionej. Powtarzamy tą procedurę tak długo, aż wszystkie odważniki nie zostaną odłożone na bok.



Oblicz liczbę porównań, które musisz wykonać.

Zadanie dodatkowe: Podaj arytmetyczne uzasadnienie dla liczby ww. porównań. Jak wyglądałoby to w przypadku dziewięciu odważników? 20?

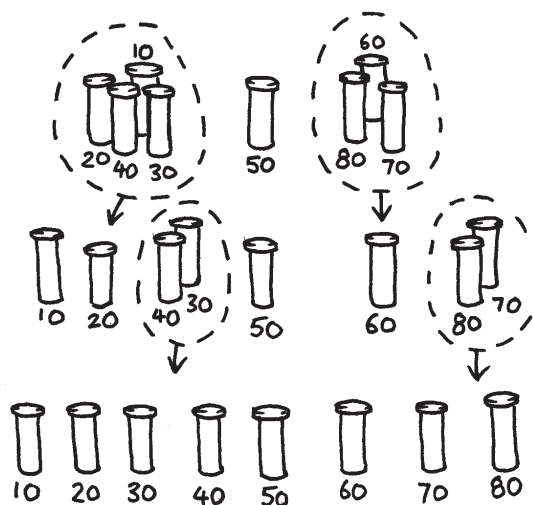
Karta pracy: Dziel i zwyciężaj

Quicksort

Quicksort (dosł. szybkie sortowanie) jest metodą sortowania, która jest o wiele szybsza od metody sortowania przez wybór, szczególnie w przypadku dużych zbiorów. Jest jedną z najlepszych znanych dziś metod sortowania.

Działa w następujący sposób:

- Wybierz dowolny z odważników i połącz go na jednej z szalek wagi.
- Porównaj jego wagę z wagą każdego z pozostałych odważników. Te, które były lżejsze połącz po lewej stronie, wybrany na początku – na środku, a cięższe – po prawej stronie. (Może się tak zdarzyć, że z jednej strony będzie o wiele więcej odważników niż z drugiej.)
- Wybierz jedną z grup i powtórz dla niej całą wyżej opisaną procedurę. Zrób to samo z drugą grupą odważników. Pamiętaj o tym, który pozostał w środku.
- Powtórz ww. procedurę dla każdej z powstałych grup odważników aż nie powstaną grupy jednoelementowe. W tym momencie przedmioty będą już uporządkowane od najlżejszego do najcięższego.



Ilu porównań wymagał ten cały proces sortowania?

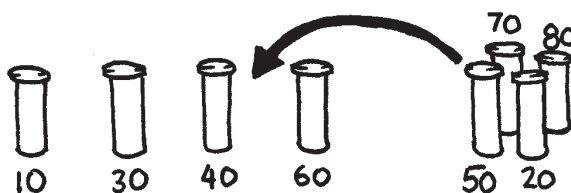
W tym momencie powinno okazać się, że ta metoda jest bardziej efektywna (chyba że jako pierwszy wybrany został najlżejszy lub najcięższy z odważników). W najlepszym przypadku, jeśli na początku wybrany był jeden ze „środkowych” odważników wystarczy mogło 14 porównań (w przypadku sortowania przez wybór było ich zawsze 28). Wniosek: metoda *quicksort* nie będzie nigdy gorsza od metody sortowania przez wybór a czasami może być o wiele lepsza!

Zadanie dodatkowe: Ile porównań byłoby wykonywanych w przypadku, gdyby w każdym kroku metody *quicksort* wybierany byłby najlżejszy z odważników?

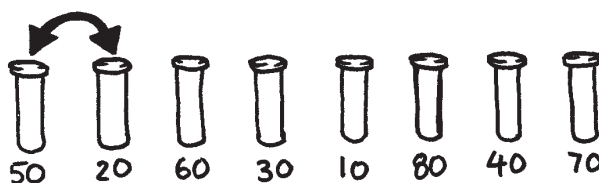
Wersja rozszerzona

Istnieje wiele innych metod sortowania. Możesz spróbować posortować odważniki używając jednej z nich.

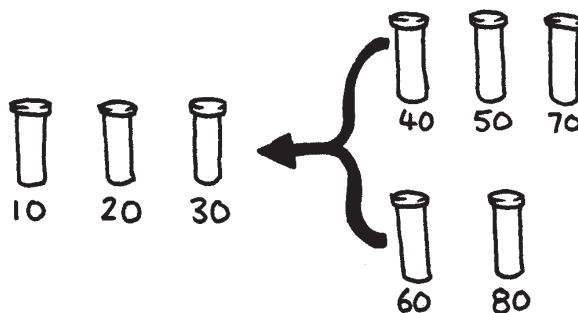
Sortowanie przez wstawianie polega na dodaniu kolejnych przedmiotów z grupy nieposortowanych we właściwe miejsce grupy przedmiotów już posortowanych (jak na rysunku). Ta procedura związana jest ze stopniowym powiększeniem grupy tych posortowanych aż do uzyskania oczekiwanego uporządkowania. Metoda ta jest często stosowana podczas gry w karty (do ułożenia kart trzymanyh w ręce przez grającego).



Sortowanie bąbelkowe (pęcherzykowe) polega na wielokrotnym powtórzeniu dla całej grupy przedmiotów następującej czynności: jeśli kolejne dwa przedmioty znajdują się w niewłaściwym porządku to zamieniamy je miejscami. Procedura nie jest więcej powtarzana, gdy podczas kolejnego przeglądania grupy przedmiotów żadne dwa z nich nie są już zamieniane miejscami. Ta prosta metoda jest jednak nieefektywna.



Sortowanie przez scalanie jest kolejną metodą, która działa w oparciu o zasadę „dziel i rządź”. Najpierw grupę przedmiotów należy podzielić w sposób losowy na dwie grupy o równej liczbie elementów (lub prawie równej, jeśli przedmiotów jest nieparzysta liczba). Każda z tych grup jest następnie sortowana i w końcu łączy się je (scala) w jedną. Takie scalanie dwóch posortowanych grup jest proste – z przedmiotów znajdujących się w danej chwili na początku każdej z dwóch uporządkowanych grup wybiera się ten o mniejszej wadze. W poniższym przykładzie: 40- i 60-gramowe odważniki znajdują się z przodu w posortowanych grupach – wybieramy odważnik 40-gramowy. W jaki sposób sortujemy te mniejsze grupy? Po prostu używamy również dla nich metody sortowania przez scalanie! Postępujemy tak wielokrotnie, aż kolejna grupa będzie składać się z pojedynczego przedmiotu.



O co w tym wszystkim chodzi?

Potrzebne nam informacje o wiele łatwiej jest odnaleźć w zbiorach, w których są uporządkowane. Książki telefoniczne, słowniki i indeksy (skorowidze) w książkach składają się z informacji uporządkowanych. Ich praktyczne użycie byłoby znacznie trudniejsze, gdyby tego uporządkowania nie było. Kiedy zestaw liczb (np. zestawienie wydatków) jest posortowany, łatwo dostrzec wartości najmniejsze i największe, podobnie jak wartości powtarzające się (duplety).

Komputery podczas działania dużą część czasu muszą poświęcać na sortowanie różnych elementów. Dlatego informatycy ciągle poszukują szybkich i efektywnych metod sortowania. W pewnych szczególnych sytuacjach swoje zastosowanie znajdują wolniejsze z ww. metod (np. sortowanie przez wstawianie czy sortowanie przez wybór). Zazwyczaj jednak w praktyce używa się dziś szybszych metod takich jak *quicksort*.

Quicksort działa w oparciu o zasadę zwaną rekurencją (rekursją). Wiąże się to z tym, że zasadę sortowania związaną z podziałem zbioru elementów stosuje się ponownie dla każdego z wyznaczonych podzbiorów itd. W tym przypadku mamy do czynienia z metodą zwaną dziel i zwyciężaj. Zbiór jest wielokrotnie dzielony na podzbiory tak długo aż staje się zbiorem jednoelementowym. Uporządkować taki zbiór jest rzeczą trywialną! Chociaż metoda ta może wydawać się zawiła, w praktyce działa znacząco szybciej niż inne metody.

Rozwiązania i wskazówki

4. Najlepsza metoda znalezienia najlżejszego odważnika polega na porównywaniu kolejnych odważników z najlżejszym znanym w danym momencie.
5. Porównaj odważniki za pomocą wagi szalkowej. Można to zrobić za pomocą trzech ważeń (czasami wystarczą dwa – trzeba zauważyć, że operacja porównywania ma własność przechodniości, tzn. jeśli A jest lżejszy niż B a B jest lżejszy niż C, to A jest lżejszy niż C).

Dodatkowe:

Poniżej znajduje się szkic rozumowania dotyczącego liczby porównań używanych podczas sortowania przez wybór.

Do znalezienia mniejszego z dwóch obiektów potrzeba jednego porównania. W przypadku trzech, liczba porównań jest równa dwa, w przypadku czterech – trzy itd. Dla ośmiu obiektów metoda sortowania przez wybór wymaga siedmiu porównań dla znalezienia najmniejszego elementu, sześciu – do znalezienia drugiego najmniejszego, pięciu – do znalezienia kolejnego itd. To w sumie daje nam:

$$7 + 6 + 5 + 4 + 3 + 2 + 1 = 28 \text{ porównań.}$$

W przypadku n obiektów liczba wymaganych porównań to $1 + 2 + 3 + 4 + \dots + n - 1$.

Dodanie tych liczb będzie łatwiejsze, jeśli przegrupujemy składniki tej sumy.

Dla przykładu, sumę $1 + 2 + 3 + \dots + 20$, możemy zapisać tak

$$\begin{aligned} &(1 + 20) + (2 + 19) + (3 + 18) + (4 + 17) + (5 + 16) + \\ &(6 + 15) + (7 + 14) + (8 + 13) + (9 + 12) + (10 + 11) \\ &= 21 \times 10 \\ &= 210 \end{aligned}$$

Ogólnie, suma $1 + 2 + 3 + 4 \dots + n - 1 = n(n - 1)/2$.