Prowadząca zajęcia:  
dr hab. inż. Małgorzata Sterna, prof. nadzw.

**ALGORYTMY I STRUKTURY DANYCH**

**Ćwiczenie 5**

**Programowanie dynamiczne**

Stanisław Jasiewicz  
nr 116753  
Wojciech Regulski  
nr 132312  
  
Informatyka(WI) I1

# Cel

Zapoznanie się z metodami dokładnymi oraz heurystycznymi wyznaczania rozwiązania problemu plecakowego oraz porównanie ich pod kątem złożoności czasowej oraz dokładności.

# Pomiary i wykresy

## Czas obliczeń t w funkcji liczby paczek n dla b=50% ∑s(ai)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **n** | **GH4** | **PD** | **BF1** | **BF2** |
| 10 | 1,02E-06 | 0,000194 | 5,98E-05 | 6,70E-06 |
| 11 | 1,22E-06 | 0,000223 | 0,000134 | 1,26E-05 |
| 12 | 1,14E-06 | 0,000278 | 0,000241 | 3,08E-05 |
| 13 | 1,62E-06 | 0,000334 | 0,000473 | 5,05E-05 |
| 14 | 1,62E-06 | 0,000343 | 0,001034 | 9,16E-05 |
| 15 | 2,18E-06 | 0,000642 | 0,003345 | 0,000299 |
| 16 | 2,42E-06 | 0,000568 | 0,004463 | 0,000356 |
| 17 | 2,59E-06 | 0,000686 | 0,00955 | 0,00072 |
| 18 | 2,54E-06 | 0,000718 | 0,020359 | 0,001557 |
| 19 | 2,50E-06 | 0,000828 | 0,043381 | 0,003161 |
| 20 | 2,90E-06 | 0,00103 | 0,091128 | 0,006802 |
| 21 | 2,82E-06 | 0,000981 | 0,182956 | 0,011813 |
| 22 | 3,01E-06 | 0,001272 | 0,381086 | 0,028149 |
| 23 | 3,06E-06 | 0,001227 | 0,757785 | 0,047459 |
| 24 | 3,14E-06 | 0,001249 | 1,477723 | 0,083394 |
| 25 | 3,42E-06 | 0,001269 | 2,996016 | 0,170293 |

## Czas obliczeń t w funkcji liczby paczek n dla b=25% ∑s(ai) oraz b=75%% ∑s(ai)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n** | **b** | **GH4** | **PD** | **BF1** | **BF2** |
| 10 | 25% | 1,02E-06 | 5.20E-05 | 4.96E-05 | 1.70E-06 |
| 11 | 1,66E-06 | 8.26E-05 | 0.000111 | 2,93E-06 |
| 12 | 1,30E-06 | 0.0001 | 0.00023 | 5.66E-06 |
| 13 | 1,78E-06 | 0.000259 | 0.00059 | 1.07E-05 |
| 14 | 1,91E-06 | 0.000241 | 0.001221 | 2.18E-05 |
| 15 | 1,98E-06 | 0.00025 | 0.002403 | 3.12E-05 |
| 16 | 2,30E-06 | 0.000328 | 0.004849 | 5.89E-05 |
| 17 | 2,58E-06 | 0.000317 | 0.010432 | 8.97E-05 |
| 18 | 2,38E-06 | 0.000344 | 0.019698 | 0.000145 |
| 19 | 2,82E-06 | 0.000419 | 0.041893 | 0.000284 |
| 20 | 2,78E-06 | 0.000438 | 0.083627 | 0.000463 |
| 21 | 2,98E-06 | 0.000588 | 0.178165 | 0.000851 |
| 22 | 3,10E-06 | 0.0005 | 0.366335 | 0.001306 |
| 23 | 3,21E-06 | 0.000543 | 0.751562 | 0.002494 |
| 24 | 3,46E-06 | 0.000636 | 1.496658 | 0.005561 |
| 25 | 3,38E-06 | 0.000651 | 3.071512 | 0.008151 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n** | **b** | **GH4** | **PD** | **BF1** | **BF2** |
| 10 | 75% | 1,58E-06 | 0.000449 | 6.77E-05 | 1.33E-05 |
| 11 | 1,34E-06 | 0.000459 | 0.000117 | 1.88E-05 |
| 12 | 1,54E-06 | 0.000607 | 0.000250 | 3.59E-05 |
| 13 | 1,30E-06 | 0.000610 | 0.000554 | 7.13E-05 |
| 14 | 1,70E-06 | 0.000758 | 0.001067 | 0.000155 |
| 15 | 1,78E-06 | 0.000753 | 0.002309 | 0.000276 |
| 16 | 2,10E-06 | 0.000834 | 0.004783 | 0.000603 |
| 17 | 2,38E-06 | 0.000959 | 0.009680 | 0.001139 |
| 18 | 2,90E-06 | 0.001260 | 0.022541 | 0.002712 |
| 19 | 2,74E-06 | 0.001311 | 0.043611 | 0.004848 |
| 20 | 2,70E-06 | 0.001299 | 0.084842 | 0.009350 |
| 21 | 3,10E-06 | 0.001625 | 0.182456 | 0.020269 |
| 22 | 3,54E-06 | 0.001669 | 0.368278 | 0.038788 |
| 23 | 3,34E-06 | 0.001521 | 0.743414 | 0.077418 |
| 24 | 3,52E-06 | 0.001846 | 1.513369 | 0.151340 |
| 25 | 3,78E-06 | 0.002099 | 3.079206 | 0.300001 |

## Błędy metod heurystycznych dla różnych wielkości plecaka

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n** | **b** | **GH1[%]** | **GH2[%]** | **GH3[%]** | **GH4[%]** |
| **100** | 75% | 28.76439 | 7.465838 | 1.310306 | 0.050197 |
| **200** | 21.23481 | 9.076982 | 1.238762 | 0 |
| **300** | 23.78241 | 8.219692 | 1.113411 | 0.004149 |
| **400** | 17.9412 | 10.7495 | 1.059614 | 0.010869 |
| **500** | 22.53358 | 10.07867 | 1.284672 | 0.005932 |
| **600** | 21.03525 | 9.03839 | 1.211943 | 0.006907 |
| **700** | 18.76446 | 9.307653 | 1.979697 | 0.004267 |
| **800** | 18.56022 | 8.674371 | 1.387884 | 0.008492 |
| **900** | 22.33318 | 9.739407 | 1.682494 | 0 |
| **1000** | 21.99263 | 8.885094 | 2.018746 | 0.009717 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n** | **b** | **GH1[%]** | **GH2[%]** | **GH3[%]** | **GH4[%]** |
| **100** | 50% | 47.98053 | 17.79928 | 4.058701 | 0.019234 |
| **200** | 35.22508 | 13.09878 | 5.196774 | 0.080333 |
| **300** | 36.60582 | 12.96039 | 6.654174 | 0.037393 |
| **400** | 38.50858 | 13.88088 | 4.988326 | 0.020384 |
| **500** | 39.22889 | 12.9186 | 5.935709 | 0.031945 |
| **600** | 40.83878 | 12.25851 | 7.783585 | 0.008498 |
| **700** | 39.99411 | 11.95873 | 7.48661 | 0.019781 |
| **800** | 37.83451 | 13.7958 | 7.68242 | 0.012686 |
| **900** | 40.3 | 13.18288 | 8.096158 | 0.012457 |
| **1000** | 39.30664 | 12.4677 | 6.380442 | 0.013248 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **n** | **b** | **GH1[%]** | **GH2[%]** | **GH3[%]** | **GH4[%]** |
| **100** | 25% | 52.91549 | 12.77143 | 13.22978 | 0.063925 |
| **200** | 54.24052 | 10.40117 | 19.85525 | 0.019951 |
| **300** | 60.1734 | 13.12807 | 19.52201 | 0 |
| **400** | 57.14923 | 9.505067 | 25.18285 | 0.018178 |
| **500** | 58.25943 | 16.57334 | 22.09522 | 0.003594 |
| **600** | 53.27514 | 14.51869 | 22.76037 | 0.009578 |
| **700** | 55.20266 | 14.58614 | 20.77172 | 0.014439 |
| **800** | 57.39895 | 13.92106 | 22.61656 | 0.021029 |
| **900** | 54.08863 | 15.10512 | 22.81534 | 0.025253 |
| **1000** | 58.34154 | 14.1097 | 22.84961 | 0.008492 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Średni błąd** | **GH1[%]** | **GH2[%]** | **GH3[%]** | **GH4[%]** |
| **b=25%** | 56.1045 | 13.46198 | 21.16987 | 0.018444 |
| **b=50%** | 39.5823 | 13.43216 | 6.42629 | 0.025596 |
| **b=75%** | 21.6942 | 9.12356 | 1.428753 | 0.010053 |
| **Średnia** | 39.127 | 12.0059 | 9.674971 | 0.018031 |

# Wnioski

Wersja decyzyjna problemu plecakowego należy do klasy problemów NP-zupełnych. Oznacza to, że jest ona nierozwiązywalna w czasie wielomianowym. Wersja optymalizacyjna jest problemem NP-trudnym. Jest to problem rozwiązywalny algorytmem psuedowielomianowym, stąd wniosek, że nie jest on silnie NP-zupełny, lecz NP-zupełny w zwykłym sensie.

Warunkiem możliwości zastosowania programowania dynamicznego jest możliwośc podziału na zależne od siebie podproblemy oraz istnienie optymalnej podstruktury. Oznacza to, że dla każdego optymalnego wyboru n elementów wybór n-1 elementów również jest optymalny.

W ćwiczeniu badane były metody dokładne (programowanie dynamiczne, brute force) oraz heurystyczne (wybór losowy, wybór po minimalnej wielkości elementów, wybór po maksymalnej wartości elementów, wybór po maksymalnym stosunku wartości do wielkości elementów). Programowanie dynamiczne cechuje się złożonością pseudowielomianową O(n\*b). Złożoność nie jest wielomianowa, ponieważ uwarunkowana jest również wielkością plecaka, a nie tylko liczbą paczek. Algorytmy siłowe mają wykładniczą złożoność O(2n). Jeden z wykorzystanych algorytmów siłowych eliminował niedopuszczalne rozwiązania, dzięki czemu działa szybciej od niezmodyfikowanej wersji, lecz jego złożoność pozostaje ta sama. Wybór losowy pozwala znaleźć rozwiązanie w czasie liniowym O(n), lecz charakteryzuje się bardzo dużym błędem. Pozostałe metody heurystyczne cechują się złożonością wynoszącą O(n\*log2n) ze względu na wewnętrzne zastosowanie algorytmu szybkiego sortowania.

Wybór algorytmu powinien być uwarunkowany potrzebami. Jeżeli istotne jest znalezienie optymalnego rozwiązania, należy użyć metody dokładnej - najlepiej programowania dynamicznego, ze względu na znacząco mniejszą złożoność obliczeniową w porównaniu do algorytmów siłowych. Jeżeli użytkownikowi bardziej zależy na szybkim otrzymaniu rozwiązania, najlepiej zastosować heurystykę. Najmniejszym błędem cechuje się wybór po największym stosunku wartości do wielkości. Najszybszą metodą jest pierwsza heurystyka polegająca na wyborze losowym, lecz powinna być stosowana jedynie, gdy nie gra roli maksymalizacja wartości, a jedynie odszukanie dowolnej możliwej kombinacji przedmiotów, które zmieszczą się w plecaku.

Czas wykonywania algorytmów wzrasta wraz ze wzrostem liczby paczek oraz w większości przypadków wzrasta wraz ze wzrostem miejsca w plecaku. Zależność czasu od liczby paczek jest liniowa w przypadku programowania dynamicznego. Heurystyka GH4 cechuje się złożonością logarytmiczną, zaś metody siłowe wykładniczą. Pojemność plecaka nie ma wpływu na szybkość działania algorytmów GH4 oraz BF1.

Algorytmy heurystyczne to takie algorytmy, które nie gwarantują podania optymalnego rozwiązania , a jedynie przybliżone w stopniu zależnym od danej metody. Algorytmy zachłanne to metody, które dokonują wyborów najlepszych jedynie w danym momencie. Algorytmy listowe wykorzystują sortowanie danych, by następnie przeglądać je według zadanej kolejności. Poza GH1 wszystkie zastosowane w ćwiczeniu metody heurystyczne są zachłanne oraz listowe.

Zauważalny jest spadek błędów metod heurystycznych wraz ze zwiększeniem miejsca w plecaku, natomiast liczba elementów nie ma wpływu na wielkość błędu.

Z ostatniej tabeli można wnioskować, że reguła ma wpływ na jakość rozwiązania. Największym błędem cechuje się metoda doboru losowego, a najmniejszym metoda doboru po stosunku wartości do rozmiaru. Mimo tego, że metoda ta bierze pod uwagę oba parametry elementów w doborze kombinacji, zachłanność tej metody czyni ją metodą heurystyczną, ponieważ optymalne rozwiązanie lokalne nie musi być optymalnym rozwiązaniem globalnym. Rozwiązania wybierane przez ten algorytm są jednak obarczone bardzo niewielkim błędem względem wartości optymalnej, może być ona więc z powodzeniem wykorzystywana, jeśli najlepsze rozwiązanie nie jest ściśle wymagane. Metoda GH2 dokonuje wyboru jedynie na podstawie wielkości elementów, co sprawia, że jej błąd jest większy niż błąd metody GH3 dokonującej wyboru na podstawie wartości. Dzieje się tak dlatego, że mała wielkość nie gwarantuje wysokiej wartości. Gdy celem jest zmaksymalizowanie wartości, metoda dobierająca elementy po najwyższych wartościach ma w związku z tym większe szanse na dobór kombinacji o łącznej wyższej wartości. Dla metody GH2 korzystne dane wejściowe będą takie, w których elementy o niewielkich rozmiarach będą miały wysoką wartość. Dla metody GH3 najkorzystniejsze dane to takie, dla których elementy o najwyższej wartości będą miały niewielki rozmiar. Metoda GH4 będzie zaś działać najlepiej dla danych wejściowych o zbliżonych rozmiarach, co zminimalizuje szansę niezmieszczenia się wartościowego elementu w plecaku.

Metody zachłanne wykorzystuje się również do rozwiązywania innych problemów niż plecakowy (Np. do wyznaczania najkrótszej ścieżki między wierzchołkami grafu), nie zawsze jest też metodą heurystyczną, ponieważ w niektórych przypadkach optymalność lokalna rozwiązania będzie wiązać się z optymalnością globalną.