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

**ALGORYTMY I STRUKTURY DANYCH**

**Ćwiczenie 4**

**Algorytmy z powracaniem**

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

# Cel

Porównanie działania algorytmów wielomianowych i algorytmów z powracaniem na podstawie poszukiwania cyklu Eulera oraz Hamiltona. Czas działania, zalety, wady, problemy tych algorytmów dla różnych wielkości i gęstości grafów.

# Pomiary i wykresy

## Zestawienie czasów poszukiwania cyklu Eulera (tE), pojedynczego cyklu Hamiltona (tH1) i wszystkich cykli Hamiltona (tHA) dla gęstości grafu wynoszącej 0,6

|  |  |  |  |
| --- | --- | --- | --- |
| **Liczba wierzchołków n** | **Czas poszukiwania cyklu Eulera [s]** | **Czas poszukiwania pojedynczego cyklu Hamiltona [s]** | **Czas poszukiwania wszystkich cykli  Hamiltona [s]** |
| 11 | 0,0000153 | 0,0000074 | 0,0019721 |
| 12 | 0,0000175 | 0,0003662 | 0,0051625 |
| 13 | 0,0000247 | 0,0006996 | 0,0504545 |
| 14 | 0,0000277 | 0,0000057 | 0,1591009 |
| 15 | 0,0000295 | 0,0000114 | 1,3720964 |
| 16 | 0,0000335 | 0,0000096 | 7,2815627 |
| 21 | 0,0000479 | 0,0006864 |  |
| 26 | 0,0000706 | 0,0010996 |  |
| 31 | 0,0001008 | 0,0000398 |  |
| 36 | 0,0001566 | 0,0000715 |  |
| 41 | 0,0001817 | 0,0000691 |  |
| 46 | 0,0002330 | 0,0004020 |  |
| 51 | 0,0002825 | 0,0004253 |  |
| 56 | 0,0003441 | 0,0001304 |  |
| 61 | 0,0004059 | 0,0001539 |  |
| 66 | 0,0004793 | 0,0001790 |  |

## Zestawienia czasu poszukiwania cyklu Eulera w zależności od gęstości grafu i liczby wierzchołków

|  |  |  |
| --- | --- | --- |
| **Gęstość d** | **Liczba wierzchołków n** | **Czas poszukiwania cyklu Eulera [s]** |
| 0,2 | 100 | 0,000417 |
| 200 | 0,001569 |
| 300 | 0,003641 |
| 400 | 0,006907 |
| 500 | 0,010742 |
| 600 | 0,015000 |
| 700 | 0,021420 |
| 800 | 0,029348 |
| 900 | 0,035727 |
| 1000 | 0,044909 |
| 0,6 | 100 | 0,001135 |
| 200 | 0,004748 |
| 300 | 0,011200 |
| 400 | 0,020137 |
| 500 | 0,032592 |
| 600 | 0,048191 |
| 700 | 0,071167 |
| 800 | 0,096506 |
| 900 | 0,122442 |
| 1000 | 0,160438 |

## Zestawienie czasu poszukiwania pojedynczego cyklu Hamiltona, wszystkich cykli Hamiltona i liczby cykli Hamiltona dla różnych gęstości grafu

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gęstość d** | **Liczba wierzchołków n** | **Czas poszukiwania pojedynczego cyklu Hamiltona [s]** | **Czas poszukiwania wszystkich cykli  Hamiltona [s]** | **Liczba cykli  Hamiltona [s]** |
| 0,2 | 11 | 0,0000073 | 0,0000077 | 0 |
| 12 | 0,0000108 | 0,0000115 | 0 |
| 13 | 0,0000280 | 0,0000276 | 0 |
| 14 | 0,0000353 | 0,0000360 | 0 |
| 15 | 0,0000855 | 0,0000830 | 0 |
| 16 | 0,0000221 | 0,0001351 | 2 |
| 21 | 0,2158274 |  |  |
| 26 | 0,6708023 |  |  |
| 31 | 2,0333483 |  |  |
| 36 | 0,0716998 |  |  |
| 41 | zbyt długi czas |  |  |
| 46 | 0,0430433 |  |  |
| 51 | 0,0218058 |  |  |
| 56 | 0,0780594 |  |  |
| 61 | 0,0005354 |  |  |
| 66 | 0,0004168 |  |  |
| 0,6 | 11 | 0,0000074 | 0,0019721 | 160 |
| 12 | 0,0003662 | 0,0051625 | 252 |
| 13 | 0,0006996 | 0,0504545 | 8312 |
| 14 | 0,0000057 | 0,1591009 | 50410 |
| 15 | 0,0000114 | 1,3720964 | 248042 |
| 16 | 0,0000096 | 7,2815627 | 991046 |
| 21 | 0,0006864 |  |  |
| 26 | 0,0010996 |  |  |
| 31 | 0,0000398 |  |  |
| 36 | 0,0000715 |  |  |
| 41 | 0,0000691 |  |  |
| 46 | 0,0004020 |  |  |
| 51 | 0,0004253 |  |  |
| 56 | 0,0001304 |  |  |
| 61 | 0,0001539 |  |  |
| 66 | 0,0001790 |  |  |

# Wnioski

Eulerowskie grafy, na których wykonywane jest ćwiczenie, generowane są za pomocą autorskiego pomysłu. Na początku tworzona jest baza składająca się z niewielkich „cykli”, w których każdy wierzchołek jest połączony z poprzednim i następnym, np. w obrębie n od 1 do 6 połączenia wyglądałyby tak: 1-2-3-4-5-6-1. Podczas dodawania każdego kolejnego cyklu wybrany zostaje losowy wierzchołek – „łącznik”, będący już częścią grafu. Wierzchołek ten będzie częścią nowego cyklu, np. podczas dodawania nowego cyklu w obrębie n od 7 do 11 przy łączniku „3” nowe połączenia wyglądałyby tak: 3-7-8-9-10-11-3, w ten sposób zachowana zostaje spójność i gwarantowane jest istnienie cyklu Eulera. Na tak wykonanym grafie eulerowskim dodawane są kolejne „trójkąty”, czyli pełne połączenia między trzema wybranymi wierzchołkami aż do osiągnięcia zadanej gęstości. „Trójkąty” nie niszczą właściwości grafu, gdyż zwiększają każdemu wybranemu wierzchołkowi stopień o 2, przez co pozostaje on parzysty, gdzie warunkiem koniecznym i wystarczającym istnienia cyklu Eulera w grafie jest parzysty stopień każdego z jego wierzchołków i spójność.

Poszukiwanie cyklu Eulera należy do klasy problemów P, czyli rozwiązywanych przez deterministyczną maszynę Turinga (DTM) w czasie wielomianowym. Taka klasyfikacja jest możliwa dzięki istnieniu algorytmu, który rozwiązuje problem w czasie wielomianowym. Złożoność algorytmu poszukującego cykl Eulera opartego na DFS dla grafu przedstawionego jak podczas ćwiczenia za pomocą listy sąsiadów wynosi O(m), m – liczba krawędzi. Dla innych reprezentacji algorytm działałby dłużej, gdyż oferują one gorszy dostęp do zbioru sąsiadów wierzchołka. Algorytm analizuje tylko krawędzie, wierzchołki praktycznie nie mają dla niego znaczenia, przez co w liście sąsiadów, gdzie dostęp do danego wierzchołka jak i pierwszego sąsiada na liście jest swobodny, kluczowe pod względem czasowym stają się operacje na krawędziach, przez co złożoność O(m). W innych reprezentacjach powstałby dodatkowy narzut w postaci poszukiwania dowolnego sąsiada, w macierzy sąsiedztwa na przykład, trzeba by było analizować za każdym razem część wiersza/kolumny, a więc większy wpływ zaczęłaby mieć liczba krawędzi (większy rozmiar macierzy do analizy). Na podstawie wzoru widać, iż przy stałej gęstości grafu, m zależy od n, przy gęstości równej 1 osiągając n2-n, skąd wniosek, że algorytm poszukujący cykl Eulera ma kwadratową złożoność obliczeniową O(n2), co dodatkowo zostało uwidocznione na wykresie zatytułowanym „Poszukiwanie cyklu Eulera”.

Poszukiwanie cyklu Hamiltona należy do klasy problemów NP-zupełnych. Taka klasyfikacja jest możliwa, dzięki dowodowi transformacji wielomianowej do niego przez wszystkie problemy klasy NP. Problemów tej klasy nie da się rozwiązywać na DTM w czasie wielomianowym; nie da się sformułować wielomianowego algorytmu. Algorytmy na problem typu NP-zupełnego (w tym szukanie cyklu Hamiltona) mają złożoność wykładniczą, gdyż bazują na generowaniu wszystkich możliwości i sprawdzaniu, czy któreś jest poprawnym rozwiązaniem. Jednym ze sposobów na optymalizację rozwiązywania problemu NP-zupełnego jest wykorzystanie algorytmu z powracaniem. Ideą algorytmu z powracaniem jest generowanie tylko „sensownych” możliwości na rozwiązanie problemu. Algorytm z powracaniem wykorzystuje się nie tylko do poszukiwania cykli Hamiltona, ale też do innych problemów NP, w których nie można wykorzystać lepszych sposobów rozwiązywania jak np. programowania dynamicznego. Generowanie wszystkich możliwych rozwiązań w celu poszukania cyklu Hamiltona jest generowaniem permutacji, skąd złożoność obliczeniowa O(n!). Algorytm z powracaniem nie analizuje części niemożliwych przypadków dzięki nieprzeglądaniu połączeń między wierzchołkami, które nie istnieją jak i połączeń do wierzchołków, które zostały już odwiedzone w danym rekurencyjnym zejściu poszukiwania. Mimo ograniczenia analizowanych przypadków w algorytmie z powracaniem, pesymistyczna złożoność poszukiwania jednego, jak i wszystkich cykli Hamiltona pozostaje O(n!). Pesymistyczny przypadek istnieje dla grafu pełnego. Poszukiwanie pojedynczego cyklu Hamiltona również ma złożoność O(n!), ze względu na możliwość braku cykli Hamiltona w grafie, gdzie przetwarzanie kończy się po pełnej analizie. Taką sytuację można zauważyć przy gęstości 0.2, gdzie nie zważając na delikatne błędy pomiarowe czas poszukiwania jednego i wszystkich cykli Hamiltona były sobie równe, gdy w grafie nie występowały cykle Hamiltona. Oczywiście im większa gęstość, a co się z tym łączy – liczba cykli Hamiltona w grafie (więcej krawędzi to większe możliwości chodzenia innymi wierzchołkami z powrotem do wierzchołka bazowego), tym większe jest prawdopodobieństwo na znalezienie jednego z nich i zakończenie działania algorytmu w krótkim czasie, co jest widoczne przy poszukiwaniu pojedynczego cyklu Hamiltona dla d=0,6. Na tym samym wykresie widać również ogromną niestabilność czasową poszukiwania pojedynczego cyklu Hamiltona, mające swoje uzasadnienie w możliwości zakończenia przetwarzania zarówno w pierwszym zejściu rekurencji do długości n, jak i po analizie całego grafu. Poszukiwanie wszystkich cykli Hamiltona ma za to dość stabilny czas wykonywania. Gęstość ma również odwrotny wpływ w porównaniu z szukaniem pojedynczego cyklu – mniejsza gęstość zapewnia dużo krótszy czas wykonywania. Reprezentacja grafu ma spore znaczenie, ponieważ przy każdym wywołaniu rekurencyjnym trzeba wyznaczyć zbiór sąsiadów wierzchołka, przez co została wykorzystana lista sąsiadów z najlepszym dostępem do nich. Sposób reprezentacji nie wpływa jednak na samo określenie złożoności obliczeniowej, gdyż narzut powstały przez źle wybraną reprezentację grafu jest zbyt mały przy złożoności O(n!). Liczba wierzchołków wpływa na zagłębienie rekurencji, które osiąga wartość n, gdy zostaje znaleziony cykl, zaś liczba krawędzi wpływa na „szerokość” drzewa przeszukiwań problemu, gdyż konieczna jest analiza większej liczby możliwości przejścia z jednego wierzchołka na drugi, co skutkuje większą liczbą „odnóg” w drzewie. Każda kolejna krawędź jest z kolei analizowana również w innych odnogach drzewa, przez co większa liczba krawędzi lawinowo wpływa na rozmiar problemu. Mała liczba krawędzi powoduje zaś, że powrót z rekurencji częściej odbywa się już przy mniejszym jej zagłębieniu, co jeszcze bardziej zmniejsza problem z każdą krawędzią mniej. Liczba krawędzi jak już było powiedziane przy okazji cykli Eulera jest rzędu n2. W tym wypadku duży wpływ na wielkość problemu ma zarówno liczba wierzchołków, jak też liczba krawędzi, choć liczba krawędzi wpływa na to w znacznie większym stopniu. Przy stałej gęstości jednak, na liczbę krawędzi bezpośrednio w stopniu kwadratowym wpływa liczba wierzchołków, przez co przy danej gęstości wielkość problemu zależy bezpośrednio od liczby wierzchołków, nie zaś od liczby krawędzi.