

Popularnonaukowe streszczenie

Jednym z najczęściej wykorzystywanych obiektów w informatyce jest drzewo – struktura danych, która pozwala reprezentować hierarchiczne zależności między obiektami. Drzewa spotykamy na każdym kroku, na przykład jako strukturę organizacyjną firmy czy drzewo genealogiczne rodziny i dlatego tak istotne jest, by badać problemy, które ich dotyczą. **Głównym celem tego projektu jest konstrukcja wydajnych algorytmów oraz analiza trudności problemów na drzewach.** Zamierzam badać poszczególne problemy na drzewach i analizować je w dwóch aspektach: ograniczeń górnych i dolnych.

Pierwszy kierunek badań ma na celu **znalezienie jak najefektywniejszego rozwiązania** dla rozważanych problemów. Najczęściej naszym celem jest, by algorytm wyznaczał poprawną odpowiedź, a przy tym działał jak najszybciej. Trudno wtedy porównywać efektywność algorytmów bez ich faktycznej implementacji, dlatego interesuje nas tzw. złożoność, czyli miara wydajności, która - z pewnym przybliżeniem - pozwala określić, dla jak dużych danych nasz algorytm znajdzie rozwiązanie np. w ciągu minuty. Dzięki nim wiemy, czy przy pomocy naszego algorytmu możemy przetwarzać dane rzędu dziesiątek, tysięcy, czy milionów elementów. Taka informacja jest dla nas bardzo istotna, bo gdy jesteśmy biologami lub chemikami i pracujemy z olbrzymimi danymi, to może się okazać, że nasze dokładne rozwiązanie nie jest w stanie ich przetworzyć nawet w rok! Wtedy musimy szukać jeszcze wydajniejszego podejścia, które być może nie znajdzie najlepszego rozwiązania, ale na pewno będzie działało szybko. Musimy jednak równocześnie zagwarantować, że znalezione rozwiązanie jest niewiele gorsze od optymalnego, powiedzmy o 20%. Jeśli więc algorytm jest szybki i mamy gwarancję, że znalezione przez niego rozwiązanie nie jest dużo gorsze od optymalnego, to nazywamy go aproksymacyjnym. Takimi algorytmami też będę zajmował się w tym projekcie.

Drugim kierunkiem badań jest **analiza ograniczeń dolnych** trudności problemów. Takie spojrzenie jest znacznie trudniejsze, gdyż wymaga pokazania, że żadne rozwiązanie danego problemu nie będzie odpowiednio szybkie. Trudno jest jednak wykazywać własności wszystkich możliwych rozwiązań, gdyż muszą one uwzględniać również te jeszcze nieznanne podejścia. W związku z tym, w ciągu ostatnich lat zaczęto analizować zależności pomiędzy problemami i pokazywać, że dany problem jest nie trudniejszy od innego. Zakładając teraz, że pewnych klasycznych algorytmów, nad którymi wielu naukowców pracowało latami, nie da się już lepiej rozwiązać możemy stawiać tzw. warunkowe ograniczenia dolne (ang. **conditional lower bounds**). Są to hipotezy mówiące, że danego problemu nie da się szybko rozwiązać, bo wtedy dałoby się przyspieszyć rozwiązanie dla innego problemu, a to wydaje się niemożliwe. Takie stwierdzenia pozwalają nam lepiej zrozumieć trudność problemów i zależności pomiędzy nimi i dlatego są tak istotne. Z tego powodu chciałbym dalej rozwijać tę dziedzinę.

Jednym z problemów, jakimi chciałbym się zająć na początku projektu, jest **Odległość Edycyjna Pomiedzy Drzewami**, w której – mając dane dwa drzewa – pytamy, ile minimalnie operacji dodania lub usunięcia wierzchołka trzeba wykonać, by z jednego drzewa otrzymać drugie.

Podsumowując, celem tego projektu jest systematyczna analiza problemów na drzewach. Badając wydajność algorytmów będę skupiał się na pesymistycznych przypadkach, więc uzyskane przeze mnie wyniki będą miały znaczenie głównie teoretyczne. Niemniej jednak jestem przekonany, że pozwolą nam one lepiej zrozumieć wiele problemów rozwiązywanych w praktyce i usprawnić ich rozwiązania.