

Algorytmy i Struktury Danych (2023)

Lista zadań 11 - kody Huffmana

1. **Regularne** drzewo binarne to takie, które nie ma węzłów z jednym dzieckiem. Udowodnij, że drzewo binarne, które **nie jest regularne**, nie może odpowiadać optymalnemu kodowi prefiksowemu.
2. Czy kody Huffmana są wyznaczone jednoznacznie dla każdego tekstu? Dlaczego?
3. Dla podanego tekstu "bababacacadaaasadaca".
 - (a) Zasymuluj działanie algorytmu generującego kody Huffmana, narysuj otrzymane drzewo kodów, wypisz kody poszczególnych znaków, oraz zakodowany tekst.
 - (b) Oblicz, o ile bitów otrzymana reprezentacja tekstu będzie krótsza od reprezentacji otrzymanej za pomocą kodów o stałej długości.
 - (c) Mając dane drzewo kodów i zakodowany tekst wykonaj dekodowanie (zaznaczaj w ciągu bitów kreską gdzie kończą się kody poszczególnych znaków).
 - (d) Zobacz, jak deszyfrowany tekst zmieni się, gdy zmienisz pierwszy bit zaszyfrowanej wersji na przeciwny.
4. Jaki jest optymalny kod Huffmana, dla zbioru częstości opartego na początkowych $n = 8$ liczbach Fibonacciego: a:1, b:1, c:2, d:3, e:5, f:8, h:21?
Uogólnij odpowiedź na przypadek dowolnego n .
5. Jaki jest optymalny kod Huffmana, dla zbioru częstości opartego na liczbach: a:11, b:12, c:13, d:14, e:15, f:18, h:19? Porównaj długość zakodowanego tekstu z długością tekstu zakodowanego przy pomocy kodów stałej długości.
Uogólnij odpowiedź na przypadek $n = 2^k$ liter w alfabecie, gdzie dodatkowo maksymalna ilość wystąpień jest mniejsza od dwukrotności minimalnej ilości wystąpień.
6. Udowodnij, że długość (ilość bitów) zaszyfrowanego tekstu, jest sumą liczb występujących w wewnętrznych węzłach drzewa kodów (nie liściach).
7. Narysuj na kartce sieć sortującą n liczb dla $n = 2, 4, 8, 16$. Powinna to być opisana na wykładzie sieć implementująca równoległą wersję algorytmu **mergesort**, działająca w czasie $O((\log n)^2)$. Prześledź działanie sieci o $n = 8$ dla ciągu wejściowego: 8 4 2 3 7 5 6 1, rysując jakie liczby wchodzi i wychodzą z każdego komparatora.