

KOMPRESJA DANYCH, WPPT 2006

12.03.2006

Lista nr 2

1. Oszacuj długość kodu liczby n w kodowaniu Fibonnaciego.
2. Kody Golomba są tworzone w następujący sposób: liczbę n dzieli się z resztą przez q : tj. $n = s \cdot q + r$. Następnie koduje się s unarnie, zaś r binarnie. Porównaj długości takich kodów z kodami Eliasa.
3. Kodowanie Hamminga przyporządkowuje literom ciągi różnej długości w zależności od ich częstości występowania. Kody Tunstalla tworzone są w odmienny sposób: zamiast pojedynczych liter koduje się być może bloki różnej długości. Czyni się to w taki sposób, aby każdy kodowany symbol lub blok miał mniej więcej takie samo prawdopodobieństwo.
Poszukaj tabel częstości występowania liter i bloków liter w języku polskim; na tej podstawie zaprojektuj odpowiednie kodowanie tego rodzaju.
4. Sformułować i udowodnić własność skalowania E_3 pozwalającą na kodowanie jedynie przeskalowań typu E_1 oraz E_2 .
5. Prześledzić kodowanie arytmetyczne dla ciągu $abacab$ przy prawdopodobieństwach symboli, odpowiednio, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{1}{6}$. Wygenerować odpowiedni kod binarny. Prześledzić pracę dekodera.
6. Napisać pseudokod kodowania i dekodowania arytmetycznego.
7. Przeanalizować jaki wpływ na kodowanie arytmetyczne ma precyzja arytmetyki w komputerze – prawdopodobieństwa są wyrażone z pewnym błędem!
8. Prześledź działanie dynamicznego kodowania Huffmana wybierz w tym celu losowo słowa o długości 10 utworzone z symboli alfabetu 4 literowego.
9. Napisać schemat blokowy dla dynamicznego kodowania Huffmana.
10. Przeanalizować złożoność czasową dynamicznego kodowania Huffmana. Pokazać, że uaktualnienie wynikające z pojawienia się nowego symbolu zajmuje czas $O(l)$, gdzie l oznacza wysokość drzewa kodowania po aktualizacji.

/-/ Mirosław Kutylowski