

KOMPRESJA DANYCH, WPPT 2006

23.02.2006

Lista nr 1

1. Uogólnić nierówność Krafta (tj. algorytm budowania kodowania oraz twierdzenie o nierówności Krafta) dla kodowania przy pomocy k symboli (zamiast kodowania binarnego, jak było na wykładzie).
2. Prześledzić jaką postać mają kody Huffmana dla:
 - k jednakowo prawdopodobnych symboli
 - dla k symboli, takich że $p(a_i) = \frac{1}{2^i}$ dla $i < k$, $p(a_k) = \frac{1}{2^{k-1}}$.
3. W algorytmie Huffmana, symbole o jednakowym prawdopodobieństwie umieszczane są na uporządkowanej liście.
 - pokazać, że kolejność umieszczania nowo tworzonych symboli na takiej liście ma istotny wpływ na kształt kodowania: pokazać, że możliwe są kodowania o nieizomorficznych drzewach kodowania.
 - Jaka strategię należałoby użyć, aby odchylenie standardowe długości kodów było jak najmniejsze? (oczekiwana długość kodów będzie oczywiście jednakowa)
4. Pokazać, że oczekiwana długość kodów Huffmana i Shannon-Fano może różnić się o co najwyżej 1. Wskazać przykład, dla którego osiągana jest możliwie duża różnica pomiędzy długościami kodów Shannon-Fano i Huffmana.
5. Pokazać, że entropia $H(p, 1 - p)$ ma maksimum dla $p = \frac{1}{2}$.
6. Niech S oznacza określone źródło symboli, zaś S^k źródło, w którym otrzymywane są bloki k symboli z S , z tym że
 - poszczególne elementy bloku generowane są niezależnie (w sensie rachunku prawdopodobieństwa),
 - każdy element bloku generowany jest zgodnie z rozkładem prawdopodobieństwa określonym dla S .

Pokazać, że $H(S^k) = H(S) \cdot k$. Pokazać, że równość ta może być fałszywa, jeśli pominiemy założenie o niezależności.
7. Kodowanie blokowe Huffmana pozwala zbliżyć się ze średnią długością kodu do entropii. Zanalizować jakie ma to jednak konsekwencje dla wielkości drzewa kodowania, wymagań pamięciowych i czasowych kodowania i dekodowania oraz preprocessingu (tj. budowy drzewa kodowania). Zaproponować kodowanie *hybrydowe*, tzn. łączące w sobie elementy różnych strategii, dla zapewnienia lepszej efektywności kodowania i dekodowania, jednocześnie przy dość dobrym aproksymowaniu entropii.

/-/ Mirosław Kutylowski