Information Theory. 2nd Chapter Problems

Boris Kudryashov

ITMO University

October 30, 2016

Problems

Table: Universal coding algorithm comparison

Algorithm	Number of traverses	Asymptotic redundancy	codeword length for text (1)
2-traverse coding, Huffman code	2	$1 + K_1/n$	302
Enumerative coding	2	$\frac{M \log n + K_3}{2n}$	283
Adaptive coding (A)	1	$\frac{M \log n + K_4}{2n}$	291
Adaptive coding (D)	1	$\frac{M \log n + K_5}{2n}$	283

IF_WE_CANNOT_DO_AS_WE_WOULD_WE_SHOULD_DO_AS_WE_CAN

(1)

Problems

1. Chose a sequence and build the table like 1. The list of suggested sequences:

```
who chatters to you will chatter about you
шел козел с косой козой, шла коза с босым козлом
либо дождик, либо снег, либо будет, либо нет
на острую косу много и покосу!  покоси-ка, коса!
два щенка щека к щеке грызли щетку в уголке
корабли лавировали, лавировали, да не вылавировали!
не узнавай друга в три дня, узнавай в три года
better late than never but better never late
men make houses but women make homes
кукушка хвалит петуха за то, что хвалит он кукушку
четыре чертенка чертили черными чернилами чертеж
can you can a can as a canner can can a can?
early to bed and early to rise makes a man wise
от умного научишься, от глупого разучишься
do not trouble trouble until trouble troubles you!
не имей сто рублей, а имей сто друзей
```

2 Try to use the estimation like:

$$\hat{p}_n(a) = \frac{\tau_n(a) + \alpha}{n + M\alpha},$$

for the chosen text, instead of (2). Adjust α to minimize codeword length (use your PC). Compare result with A and D algorithms.

$$\hat{p}_n(a) = \frac{\tau_n(a) + 1/2}{n + M/2} = \frac{2\tau_n(a) + 1}{2n + M}. \quad (2)$$

Problems

3 Approach, described in this chapter, is oriented on the DMS, but after some modifications it can be used for sources with memory. What should be changed in A and D algorithms?