

22/05/2019

Κώδικας Huffman

Μέθοδος αλφάνηξης δεδομένων

Μας δίνεται ένα αρχείο με  $n$  διαφορετικούς χαρακτήρες. Θέλουμε να αντιστοιχίσουμε τον κάθε χαρακτήρα σε ένα δυαδικό κώδικα, έτσι ώστε να ελαχιστοποιηθεί ο συνολικός αριθμός των bits όλου του αρχείου.

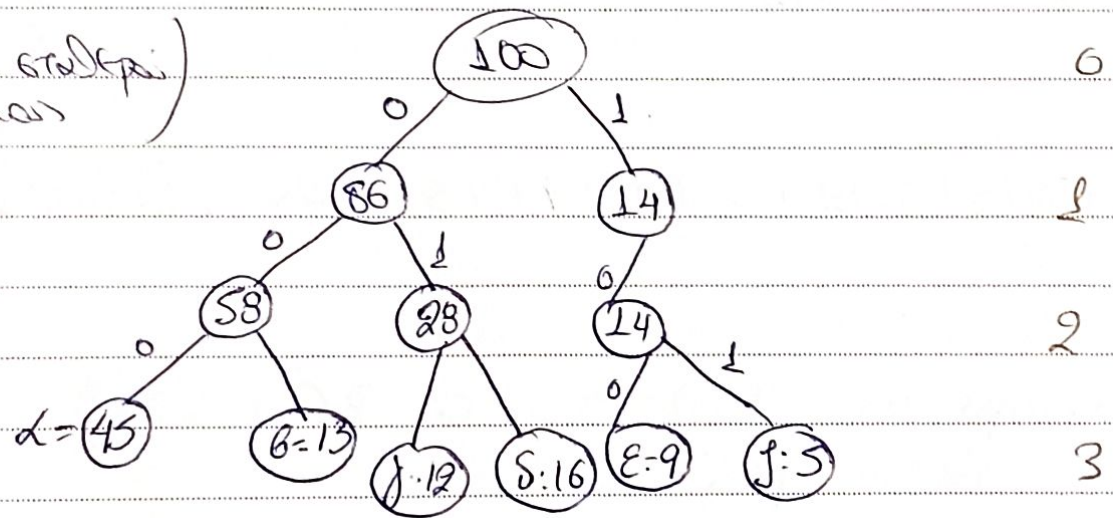
π.χ. Αρχείο που αποτελείται από τον χαρακτήρες α, β, γ, δ, ε, ζ.

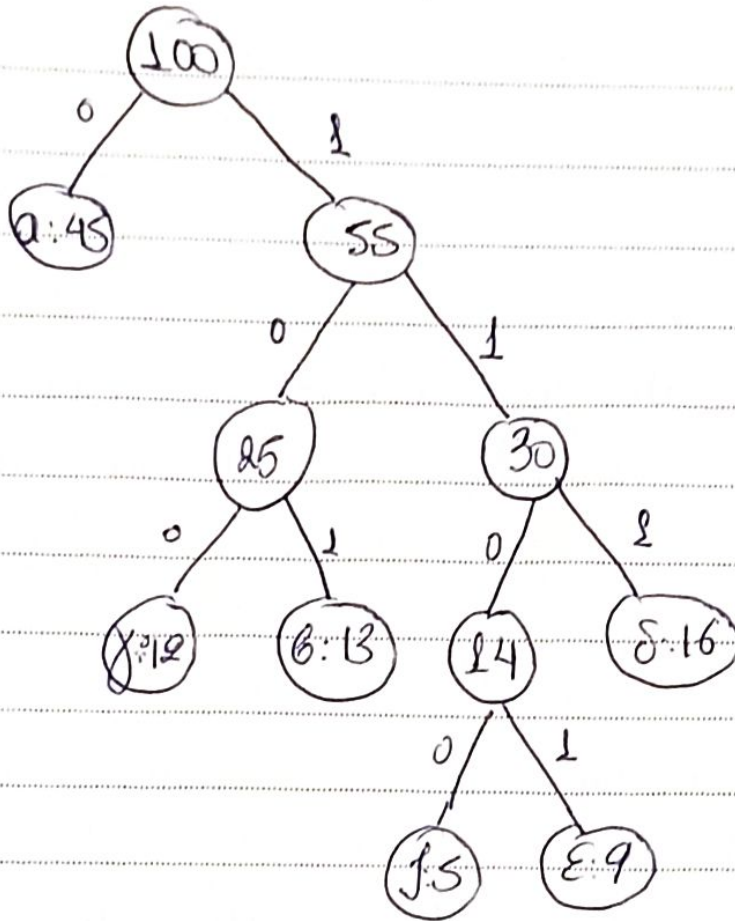
συχνότητα ( $f(c)$ )      α      β      γ      δ      ε      ζ      (=100)

κώδικας σταθερού μήκους      000      001      010      011      100      101

κώδικας μεταβλητού μήκους      0      101      100      111      1101      1100

(κώδικας σταθερού μήκους)





(κωδίνων με 0/1 bits)  
 μήκους)

$f(c)$ : συχνότητα του χαρακτήρα  $c$

$d_T(c)$ : βάθος του  $c$

$B(T) = \sum_c f(c) \cdot d_T(c)$  συνολικός αριθμός bits κωδικοποιήσεων

$$B(T) = 3 \cdot 45 + 3 \cdot 13 + 3 \cdot 12 + 3 \cdot 9 + 3 \cdot 5 = 3 \cdot 100 = 300 \quad \left( \begin{array}{l} \text{για κώδικα} \\ \text{σταθερού} \\ \text{μήκους} \end{array} \right)$$

$$B(T) = 1 \cdot 45 + 3 \cdot 13 + 3 \cdot 12 + 3 \cdot 16 + 4 \cdot 9 + 4 \cdot 5 = 224 \quad \left( \begin{array}{l} \text{για κώδικα} \\ \text{με μεταβλητό} \\ \text{μήκος} \end{array} \right)$$

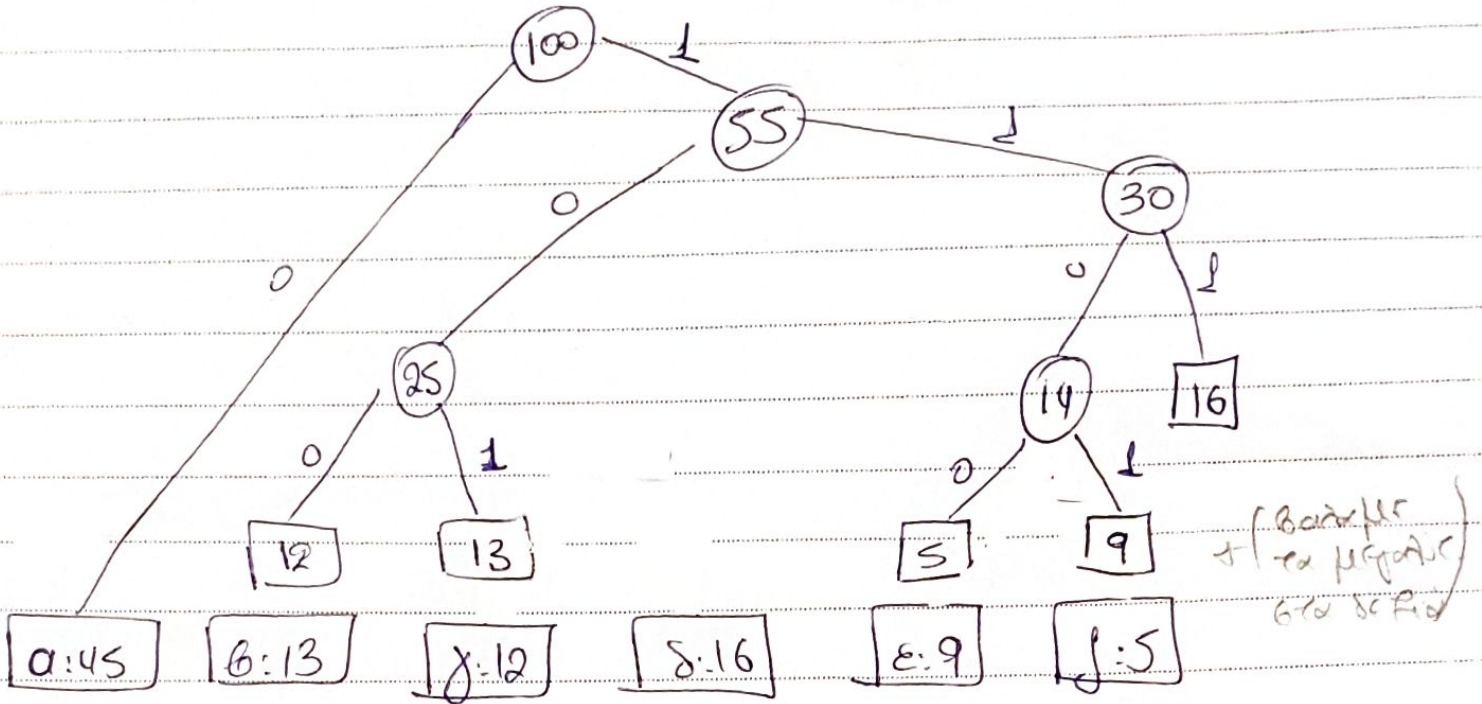
Από τη σχέση από  $B(T) = 300$  σε  $B(T) = 224$   
 το κώδικα βελτιώθηκε.

# Κατασκευή κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την ελάχιστη συχνότητα.

Εστω  $x$  και  $y$  δύο τυχαίοι χαρακτήρες

Αναδιατάσσουμε τους  $x$  και  $y$  από νέο χαρακτήρα  $z$  με συχνότητα:  $f(z) \leftarrow f(x) + f(y)$



Εστω ότι: a b g s e

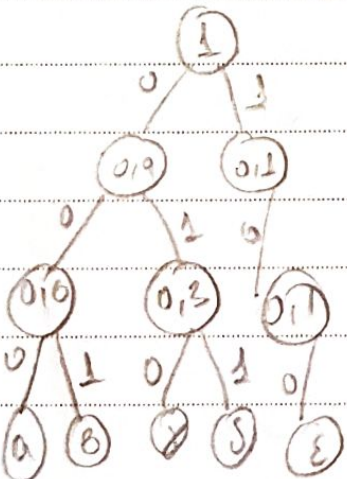
0,4 0,2 0,2 0,1 0,1

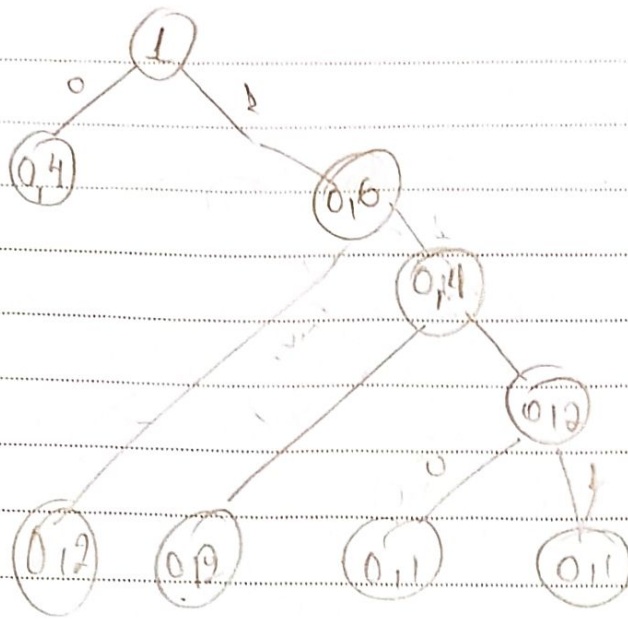
000 001 010 011 100

0 10 110 1110 1111

(συνεχώς μικρό)

$$B(T) = 3 \cdot (0,4 + 0,2 + 0,2 + 0,1 + 0,1) = 3 \cdot 1 = 3$$





μεταθάνισ

$$T(B) = 1 \cdot 0.4 + 2 \cdot 0.2 + 3 \cdot 0.9 + 4 \cdot 0.1 + 4 \cdot 0.1 = 2.2$$

SOS ⇒ Αποδείξτε με Master theory  
πολιτιστικά

συμπεριεχόμενα και προαιρετικά μαθήματα  
απόφοιτοι ταξινόμησης