

Κώδικες Huffman

Μέθοδος συμπίεσης δεδομένων:

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

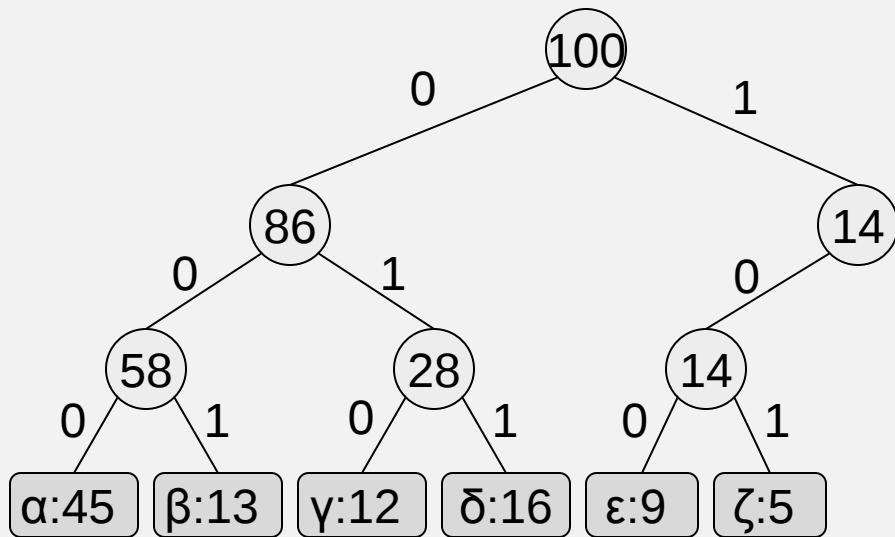
Π.χ. Αρχείο που αποτελείται από τους χαρακτήρες α,β,γ,δ,ε και ζ

	<u>α</u>	<u>β</u>	<u>γ</u>	<u>δ</u>	<u>ε</u>	<u>ζ</u>
συχνότητα ($\times 10^3$)	45	13	12	16	9	5
κώδικας σταθερού μήκους	000	001	010	011	100	101
κώδικας μεταβλητού μήκους	0	101	100	111	1101	1100

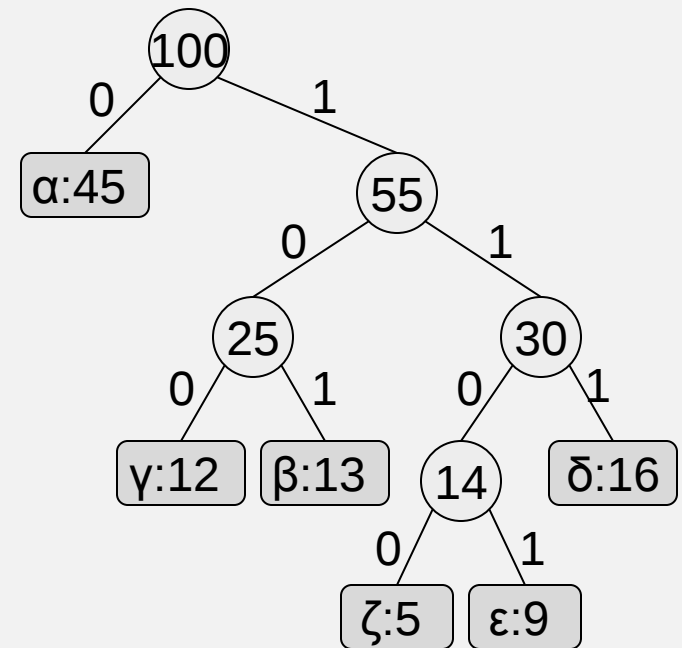
Απροθηματικός κώδικας: κανένας κωδικός δεν αποτελεί πρόθημα άλλου κωδικού

Κώδικες Huffman

	<u>α</u>	<u>β</u>	<u>γ</u>	<u>δ</u>	<u>ε</u>	<u>ζ</u>
συχνότητα ($\times 10^3$)	45	13	12	16	9	5
κώδικας σταθερού μήκους	000	001	010	011	100	101
κώδικας μεταβλητού μήκους	0	101	100	111	1101	1100



$$3 \cdot 100 = 300$$



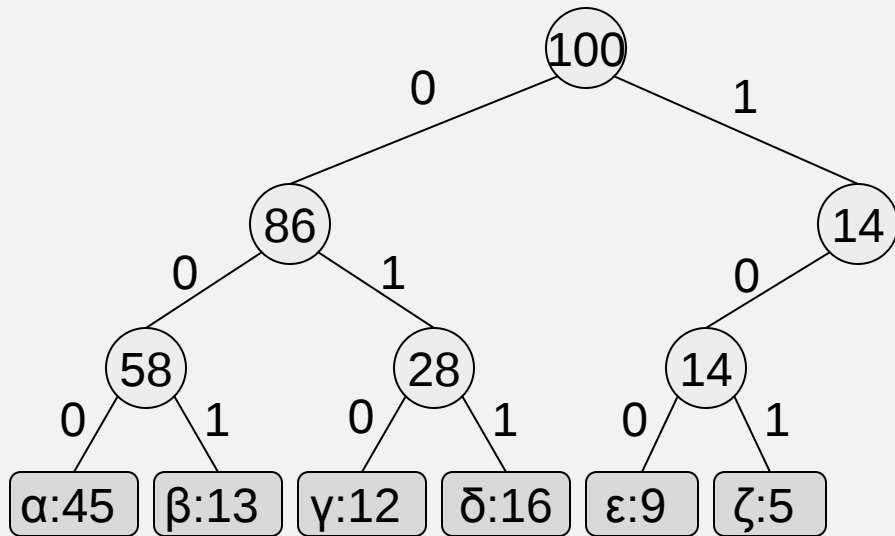
$$1 \cdot 45 + 3 \cdot 13 + 3 \cdot 12 + 3 \cdot 16 + 4 \cdot 9 + 4 \cdot 5 = 224$$

Κώδικες Huffman

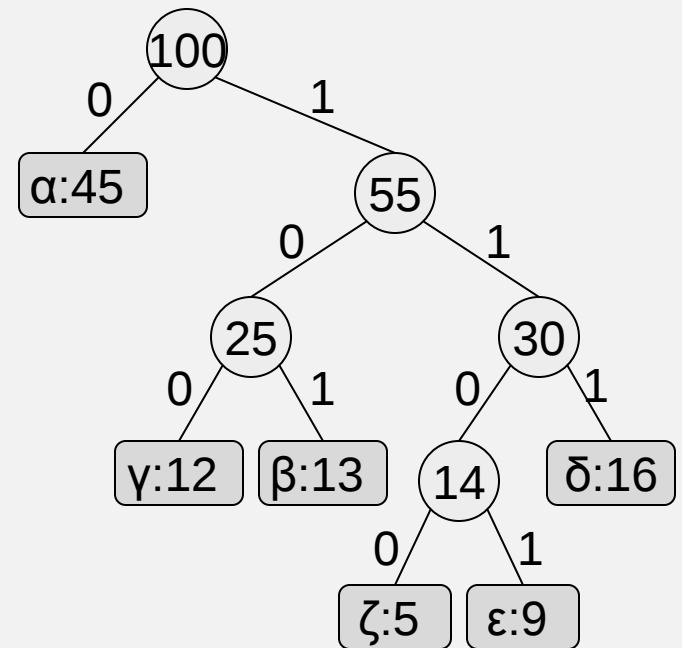
$f(c)$ συχνότητα του χαρακτήρα c

$d_T(c)$ βάθος του c στο δένδρο T

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



$$3 \cdot 100 = 300$$



$$1 \cdot 45 + 3 \cdot 13 + 3 \cdot 12 + 3 \cdot 16 + 4 \cdot 9 + 4 \cdot 5 = 224$$

Κώδικες Huffman

Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$

α:45	β:13	γ:12	δ:16	ε:9	ζ:5
------	------	------	------	-----	-----

Κώδικες Huffman

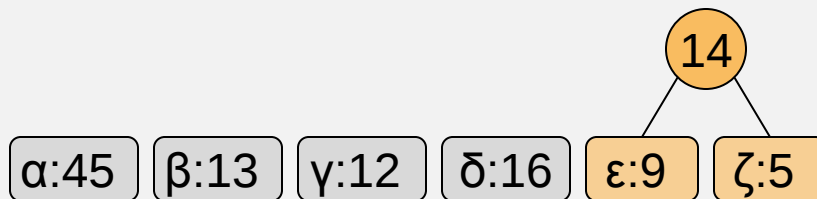
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

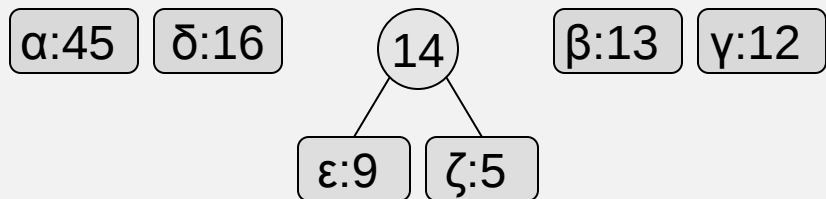
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

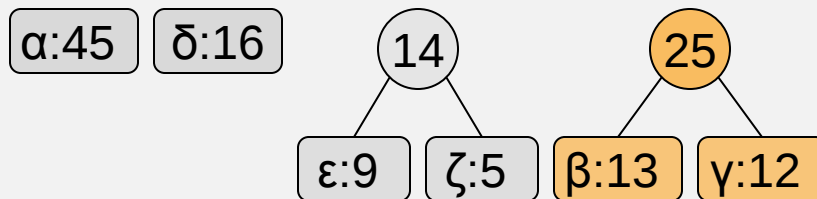
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

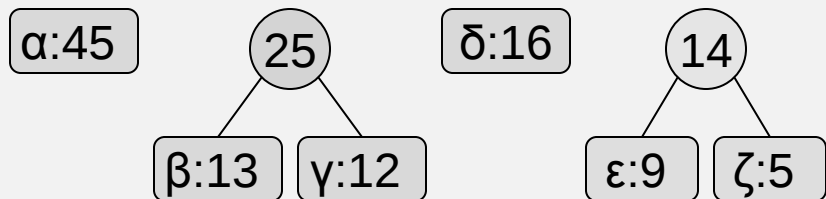
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

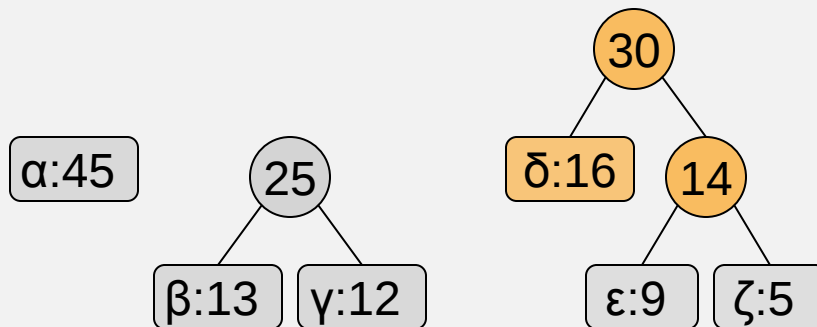
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

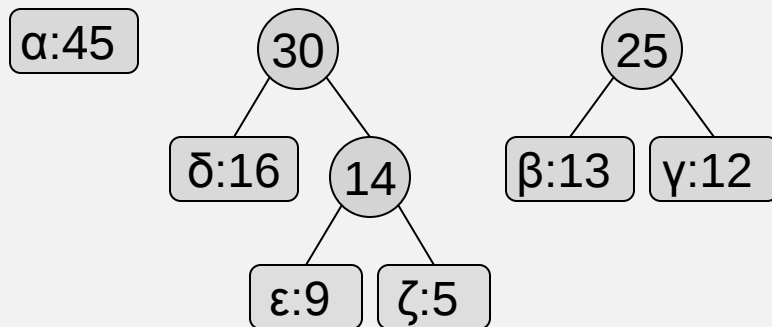
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

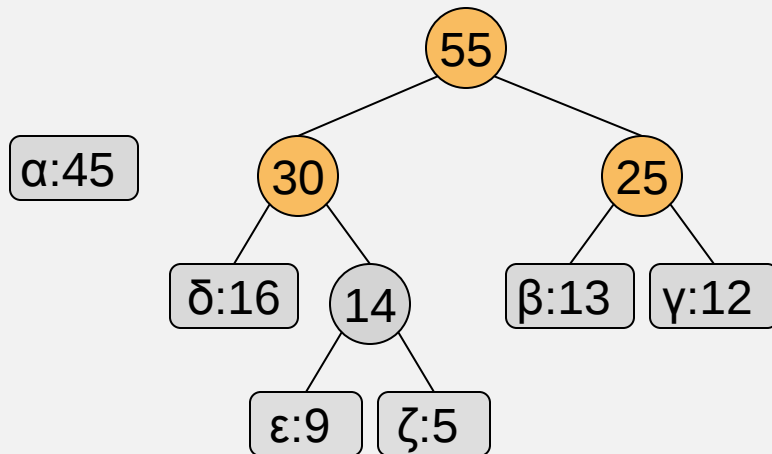
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

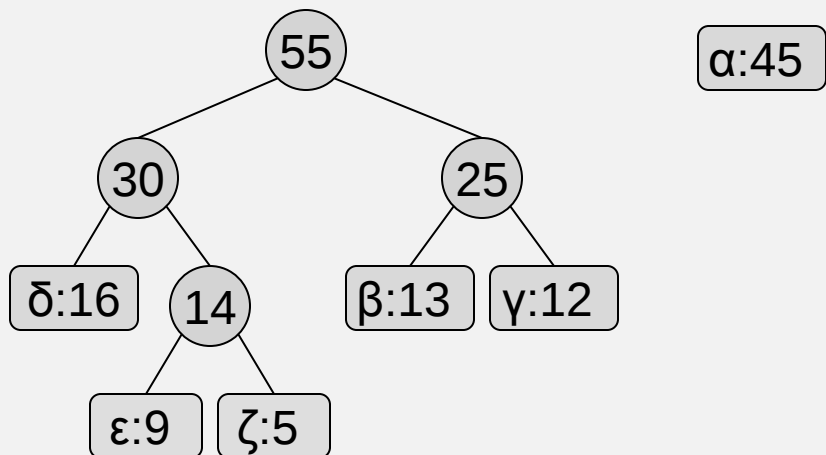
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

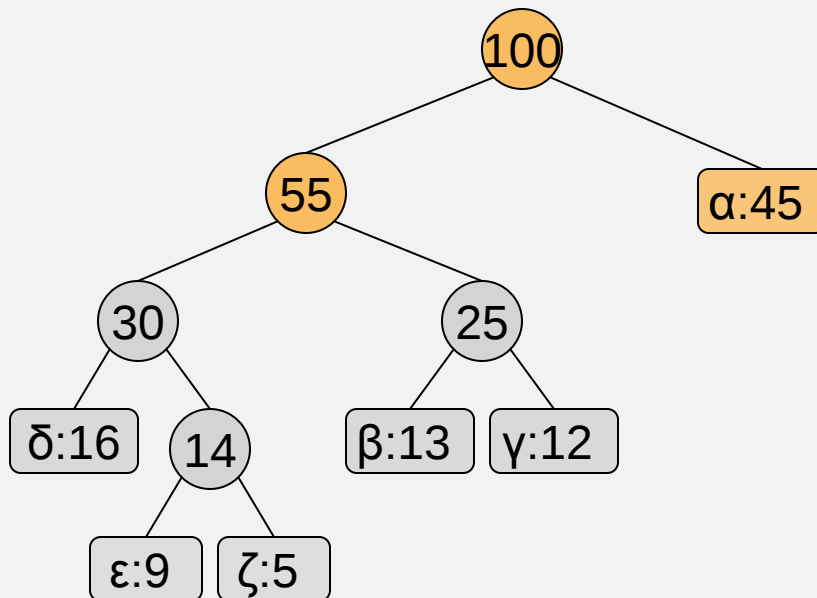
Κατασκευή του κώδικα

Σε κάθε βήμα συνδυάζουμε 2 χαρακτήρες με την **ελάχιστη συχνότητα**.

Έστω x και y δύο τέτοιοι χαρακτήρες.

Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα

$$f(z) \leftarrow f(x) + f(y)$$



Κώδικες Huffman

Υλοποίηση με ουρά προτεραιότητας

Σε κάθε βήμα :

- Συνδυάζουμε 2 χαρακτήρες, έστω x και y , με την ελάχιστη συχνότητα
 $\Rightarrow$ 2 εξαγωγές ελάχιστου
- Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα $f(z) \leftarrow f(x) + f(y)$
 $\Rightarrow$ εισαγωγή στοιχείου

Πλήθος των εξαγωγών ελάχιστου = πλήθος εισαγωγών

Κώδικες Huffman

Υλοποίηση με ουρά προτεραιότητας

Σε κάθε βήμα :

- Συνδυάζουμε 2 χαρακτήρες, έστω x και y , με την ελάχιστη συχνότητα
 $\Rightarrow$ 2 εξαγωγές ελάχιστου
- Αντικαθιστούμε τους x και y από νέο χαρακτήρα z με συχνότητα $f(z) \leftarrow f(x) + f(y)$
 $\Rightarrow$ εισαγωγή στοιχείου

Πλήθος των εξαγωγών ελάχιστου = πλήθος εισαγωγών

= πλήθος κόμβων του δένδρου = $2n - 1$

όπου n ο αριθμός των διαφορετικών χαρακτήρων

Συνολικός χρόνος κατασκευής = $O(n \log n)$

