

Εισαγωγή στην πληροφορική και τις εφαρμογές της

ΠΑΝΑΓΙΩΤΗΣ ΠΑΠΑΖΟΓΛΟΥ

Website: microdev.gr



Επιστημονικές Εκδόσεις
ΤΖΙΟΛΑ

ΚΕΦΑΛΑΙΟ 4

Δεδομένα υπολογιστή



Αποτύπωση δεδομένων

όλα τα δεδομένα «αποτυπώνονται» στον υπολογιστή σε δυαδική μορφή

(α) Πώς θα υπολογίσουμε τον όγκο των δεδομένων που αποθηκεύονται σε ένα αρχείο που φιλοξενεί τις καρτέλες μαθητών ενός σχολείου;

(β) Ποια είναι η απαιτούμενη χωρητικότητα του επικοινωνιακού καναλιού για τη ζωντανή μετάδοση ενός σεμιναρίου;

(γ) Πώς μπορούμε να βελτιώσουμε την απόδοση (με όρους ταχύτητας) ενός προγράμματος, που διαχειρίζεται μεγάλο όγκο δεδομένων;

(δ) Πώς θα βελτιστοποιήσουμε τη χρήση της διαθέσιμης μνήμης για προγράμματα και δεδομένα;

(ε) Ποιος είναι ο όγκος του αρχείου για την αποθήκευση μιας έγχρωμης εικόνας;

(στ) Πώς θα κωδικοποιήσουμε την κατάσταση εξωτερικών συσκευών, ώστε να τις ελέγχουμε μέσω του υπολογιστή;

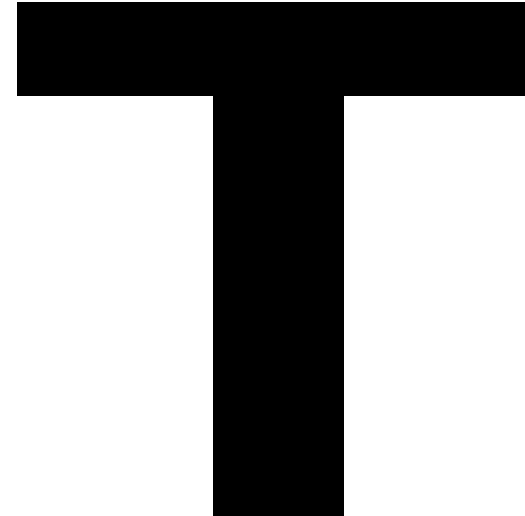
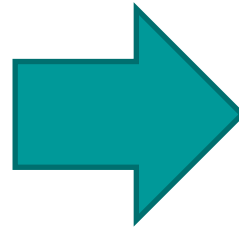
(ζ) Με ποιο τρόπο γίνεται η συμπίεση δεδομένων και σε τι εξυπηρετεί;

όλα τα παραπάνω σχετίζονται με την «αποτύπωση» των δεδομένων στον υπολογιστή

Εικόνα - Σύμβολο

Παράδειγμα

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0



Κωδικοποίηση συμβόλου

- 1 : pixel αναμμένο (μαύρο χρώμα)
- 0 : pixel σβηστό (λευκό, όπως το φόντο)
- Απαιτούνται συνολικά 96 bit (12 γραμμές X 8 στήλες)
- Δεδομένου ότι 1 byte = 8 bit → κάθε γραμμή περιέχει 1 byte
- 12 γραμμές = 12 byte
- {00, FF, FF, 18, 18, 18, 18, 18, 18, 18, 18, 18} → κάθε γραμμή στο δεκαεξαδικό

Εφαρμογή

Ετικέτες προϊόντων

Ετικέτες προϊόντων	
Ετικέτα	Προϊόν
0	Καραβόπανο
1	Δερματίνη
2	Βισκόζη
3	Πολυεστέρα
4	Ζέρσεϊ
5	Λινό
6	Βαμβάκι
7	Νάιλον
8	Ραιγιόν
9	Microfiber

Παράδειγμα με το δεκαδικό σύστημα αρίθμησης

- Βιοτεχνία κατασκευάζει δέκα (10) διαφορετικά είδη υφασμάτων
- Κάθε προϊόν (τύπος υφάσματος) θα έχει μια μοναδική ετικέτα
- Ετικέτες με ένα ψηφίο

$$10^1=10$$

1 ψηφίο (εκθέτης) = αναπαράσταση με μοναδικό τρόπο, 10 διαφορετικών προϊόντων

Για 150 διαφορετικά υφάσματα, πόσα ψηφία απαιτούνται για κάθε ετικέτα;

$$\text{Ψηφία}=\text{Εκθέτης}=2$$

$$10^2=100$$

2 ψηφία = ετικέτες για 100 διαφορετικά υφάσματα

$$\text{Ψηφία}=\text{Εκθέτης}=3$$

$$10^3=1000$$

3 ψηφία = ετικέτες για 1000 διαφορετικά υφάσματα

Γενικός τύπος

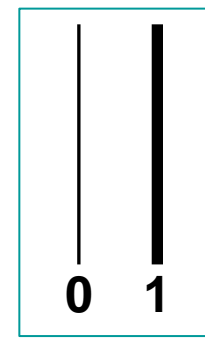
$$10^N \geq K$$

για **K** διαφορετικά προϊόντα απαιτούνται **N** ψηφία του δεκαδικού συστήματος για τις αντίστοιχες ετικέτες



Εφαρμογή

Προσαρμοσμένο bar code



- Κατάστημα διαθέτει 140 συγκεκριμένα μοντέλα κινητών τηλεφώνων
- bar code = ηλεκτρονική καταγραφή προϊόντων
- Το bar code χρησιμοποιεί μόνο δύο τύπους κατακόρυφων γραμμών που αντιπροσωπεύουν το 0 και 1
- Για $N=8$ θα ισχύει $2^8=256$, που υπερκαλύπτει το πλήθος των προϊόντων

Προϊόν 34₍₁₀₎ = 00100010₍₂₎
Προϊόν 51₍₁₀₎ = 00110011₍₂₎
Προϊόν 80₍₁₀₎ = 01010000₍₂₎
Προϊόν 96₍₁₀₎ = 01100000₍₂₎



Εφαρμογή (1)

1 ή 8bit;

Παράδειγμα

- Εικόνα 300x200 pixel, βάθος χρώματος 8bit
- Συνολικά από $300 \times 200 \times 8 \text{bit} = 480000 \text{bit} = 60000 \text{byte}$
- $1 \text{Kbyte} = 1024 \text{byte}$, $60000 \text{byte} / 1024 = 58.59 \text{Kbyte}$ (περίπου 59Kbyte)



```
image=imread('omonoia.jpg');  
image2=rgb2gray(image);  
figure; imshow(image2);
```



Μετατροπή σε κλίμακα
του γκρι (MATLAB)



256 αποχρώσεις
του γκρι

Εφαρμογή (2)

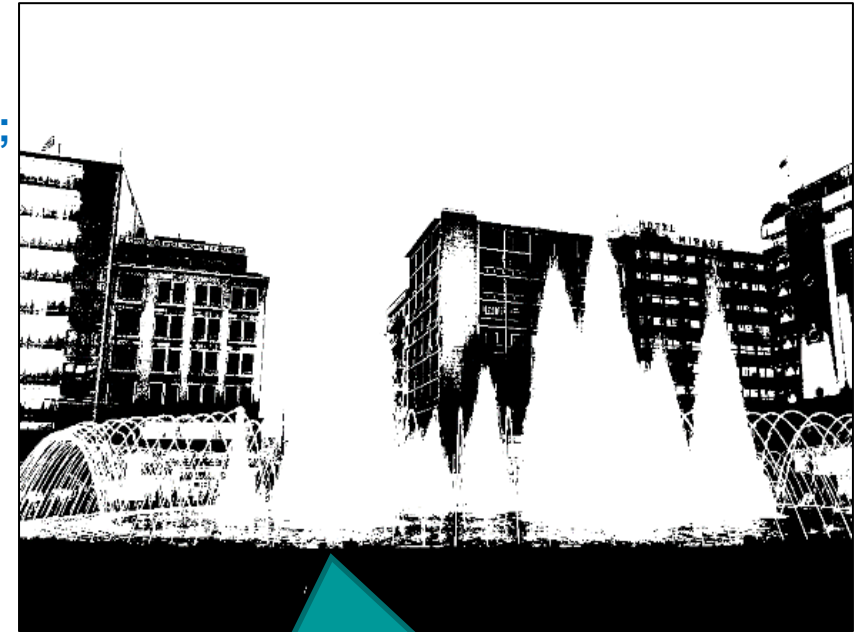
1 ή 8bit;



```
image1bit = (image2 / 128)*255;  
figure; imshow(image1bit);
```



Μετατροπή σε 1bit
βάθος χρώματος
(MATLAB)



2 αποχρώσεις

Εφαρμογή

Ζωντανή μετάδοση σεμιναρίου

- Προβολή 25 εικόνων το δευτερόλεπτο = 25 fps
- 1 εικόνα = 1 πλαίσιο
- Εικόνα = 640x480 pixels, 256 αποχρώσεις

Πόση χωρητικότητα καναλιού απαιτείται;

- 1 Εικόνα (πλαίσιο) $\rightarrow 640 \times 480 \times 8 \text{bit}$ ($2^8=256$) $\rightarrow 2457600 \text{bit}$
- $640 \times 480 \times 8 \times 25 = 61440000 \text{bit}$ ανά δευτερόλεπτο
- $61440000 \text{bit per second} = 61440000 \text{bps}$
- Χοντρικά 61 Mbps



Σύνολα συμβόλων (αλφάβητα)

ASCII 7 bit (1)

ASCII 7bit				
10-δικό	16-δικό	2-δικό (7bit)	Σύμβολο	Περιγραφή
0	00	0000000	NUL	Null char
1	01	0000001	SOH	Start of Heading
2	02	0000010	STX	Start of Text
3	03	0000011	ETX	End of Text
4	04	0000100	EOT	End of Transmission
5	05	0000101	ENQ	Enquiry
6	06	0000110	ACK	Acknowledgment
7	07	0000111	BEL	Bell
8	08	0001000	BS	Back Space
9	09	0001001	HT	Horizontal Tab
10	0A	0001010	LF	Line Feed
11	0B	0001011	VT	Vertical Tab
12	0C	0001100	FF	Form Feed
13	0D	0001101	CR	Carriage Return
14	0E	0001110	SO	Shift Out / X-On
15	0F	0001111	SI	Shift In / X-Off
16	10	0010000	DLE	Data Line Escape
17	11	0010001	DC1	Device Control 1 (oft. XON)
18	12	0010010	DC2	Device Control 2
19	13	0010011	DC3	Device Control 3 (oft. XOFF)
20	14	0010100	DC4	Device Control 4
21	15	0010101	NAK	Negative Acknowledgement
22	16	0010110	SYN	Synchronous Idle
23	17	0010111	ETB	End of Transmit Block
24	18	0011000	CAN	Cancel
25	19	0011001	EM	End of Medium
26	1A	0011010	SUB	Substitute
27	1B	0011011	ESC	Escape
28	1C	0011100	FS	File Separator
29	1D	0011101	GS	Group Separator
30	1E	0011110	RS	Record Separator
31	1F	0011111	US	Unit Separator
32	20	0100000		Space
33	21	0100001	!	
34	22	0100010	"	
35	23	0100011	#	
36	24	0100100	\$	
37	25	0100101	%	
38	26	0100110	&	

2⁷ → 128 σύμβολα

Ενδεικτικά

65	41	1000001	A
66	42	1000010	B
67	43	1000011	C
68	44	1000100	D
69	45	1000101	E
70	46	1000110	F
71	47	1000111	G
72	48	1001000	H
73	49	1001001	I
74	4A	1001010	J
75	4B	1001011	K
76	4C	1001100	L
77	4D	1001101	M
78	4E	1001110	N
79	4F	1001111	O
80	50	1010000	P
81	51	1010001	Q
82	52	1010010	R
83	53	1010011	S
84	54	1010100	T
85	55	1010101	U
86	56	1010110	V
87	57	1010111	W
88	58	1011000	X
89	59	1011001	Y
90	5A	1011010	Z
91	5B	1011011	[
92	5C	1011100	\
93	5D	1011101]
94	5E	1011110	^
95	5F	1011111	_
96	60	1100000	`
97	61	1100001	a
98	62	1100010	b
99	63	1100011	c
100	64	1100100	d
101	65	1100101	e
102	66	1100110	f
103	67	1100111	g
104	68	1101000	h



Σύνολα συμβόλων (αλφάβητα)

ASCII 7 bit (2)

«Good morning»

Κωδικοποίηση κειμένου		
Σύμβολο	16-δικό	2-δικό
G	47	1000111
o	6F	1101111
o	6F	1101111
d	64	1100100
	20	0100000
m	6D	1101101
o	6F	1101111
r	72	1110010
n	6E	1101110
i	69	1101001
n	6E	1101110
g	67	1100111

16-δικό

47 6F 6F 64 20 6D 6F 72 6E 69 6E 67

2-δικό

1000111 1101111 1101111 1100100 0100000
1101101 1101111 1110010 1101110 1101001
1101110 1100111

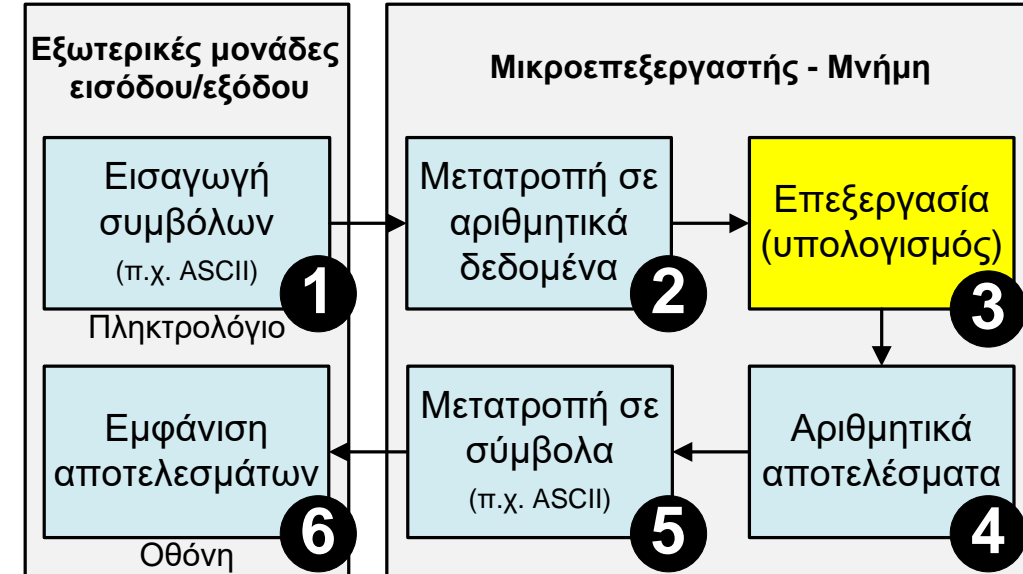


Σύνολα συμβόλων (αλφάβητα)

Ανάγνωση «αριθμητικών» ψηφίων από το πληκτρολόγιο

Παράδειγμα: πρόσθεση 1+3

- **Χρήστης** : θα πατήσει τα πλήκτρα <1> και <3>
- **Πλήκτρο <1>** : εισάγεται το σύμβολο ASCII “1” και όχι η αριθμητική αξία 1
- **Μετατροπή συμβόλου** “1” σε αριθμητική αξία
- **Αριθμός 3** : η ίδια διαδικασία
- **Πίνακας ASCII** : πλήκτρο <1> → εισάγεται το δεκαεξαδικό $31_{(16)}$
- **Αριθμητική αξία 1** : θα αφαιρέσουμε από το $31_{(16)}$ το $30_{(16)}$
- **Πράξη** : $31_{(16)} - 30_{(16)} = 1$ ή “1” - $30_{(16)} = 1$
- **Όλες οι πράξεις** : “1” - $30_{(16)} = 1$, δηλαδή $31_{(16)} - 30_{(16)} = 1$ και “3” - $30_{(16)} = 3$, δηλαδή $33_{(16)} - 30_{(16)} = 3$
- **Τελική εκτέλεση της πρόσθεσης 1+3**



Εφαρμογή

Αριθμητικός υπολογισμός

Κώδικας (υλοποίηση λογικής της προηγούμενης διαφάνειας)

```
#include <stdio.h>
int main()
{
    char a,b,c;
    int na,nb,nc;

    printf("Numbers A,B:");
    scanf("%c%c",&a,&b);
    na=a-0x30; nb=b-0x30;
    nc=na+nb;
    printf("Sum(1)=%d",nc);
    c=nc+0x30;
    printf("\nSum(2)=%c",c);
    return 0;
}
```



Αριθμοί με περισσότερα ψηφία (1)

Παράδειγμα

Ζητούμενη πράξη **1732+324**

Έστω N1, N2 πίνακες συμβόλων
για τους αριθμούς 1732 και 324

$N1[1]="1"$, $N1[2]="7"$, $N1[3]="3"$, $N1[4]="2"$ ή $N1=\{"1", "7", "3", "2"\}$

και

$N2[1]="3"$, $N2[2]="2"$, $N2[3]="4"$ ή $N2=\{"3", "2", "4"\}$

Μετατροπή σε αριθμητική αξία

$$NA1[1]=N1[1]-30_{(16)}$$

$$NA1[2]=N1[2]-30_{(16)}$$

$$NA1[3]=N1[3]-30_{(16)}$$

$$NA1[4]=N1[4]-30_{(16)}$$

$$NA2[1]=N2[1]-30_{(16)}$$

$$NA2[2]=N2[2]-30_{(16)}$$

$$NA2[3]=N2[3]-30_{(16)}$$

Για το 1732

Num1 =

$$= NA1[1]*10^3+NA1[2]*10^2+NA1[3]*10^1+NA1[4]*10^0 =$$

$$= NA1[1]*1000+NA1[2]*100+NA1[3]*10+NA1[4]*1 =$$

$$= 1*1000+7*100+3*10+2*1 =$$

$$= 1000+700+300+2 =$$

$$= 1732$$



Αριθμοί με περισσότερα ψηφία (2)

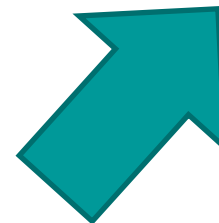
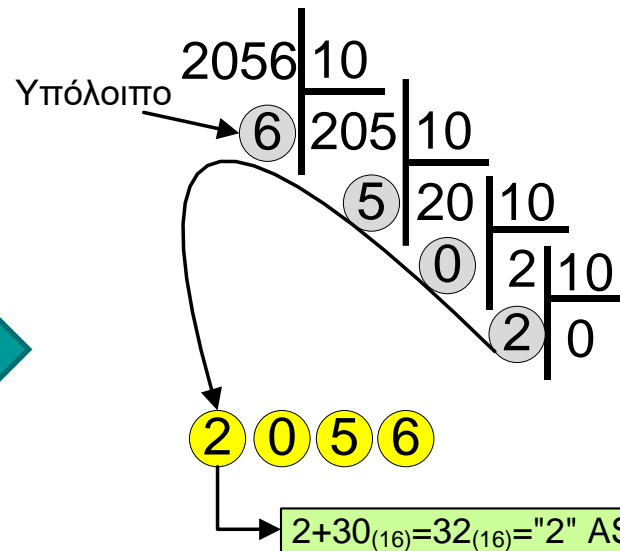
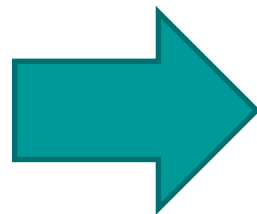
Παράδειγμα

Ζητούμενη πράξη **1732+324**

$$\begin{aligned} \text{Num2} &= 3 \cdot 10^2 + 2 \cdot 10^1 + 4 \cdot 10^0 = 3 \cdot 100 + 2 \cdot 10 + 4 \cdot 1 = 300 + 20 + 4 = 324 \\ \text{Sum} &= \text{Num1} + \text{Num2} = 1732 + 324 = 2056 \end{aligned}$$

Το αριθμητικό άθροισμα 2056 πρέπει να μετατραπεί σε ακολουθία ASCII για να εμφανιστεί στην οθόνη

Διαχωρισμός
ψηφίων



$$\begin{aligned} 2+30_{(16)} &= 32_{(16)} = '2' \text{ (ASCII)} \\ 0+30_{(16)} &= 30_{(16)} = '0' \text{ (ASCII)} \\ 5+30_{(16)} &= 35_{(16)} = '5' \text{ (ASCII)} \\ 6+30_{(16)} &= 36_{(16)} = '6' \text{ (ASCII)} \end{aligned}$$

Μετατροπή πεζών σε κεφαλαία

Παράδειγμα

h	'h'-32=104-32=72='H'	H
e	'e'-32=101-32=69='E'	E
l	'l'-32=108-32=76='L'	L
l	'l'-32=108-32=76='L'	L
o	'o'-32=111-32=79='O'	O

Εφαρμογή

```
#include <stdio.h>
int main()
{
    char word[7];
    printf("Word:");
    scanf("%s",word);
    for (int i=0;i<7;i++)
    {
        word[i]=word[i]-32;
        printf("%c",word[i]);
    }
}
```



Σύνολα συμβόλων (αλφάβητα)

ASCII 8 bit – Extended (1)

$2^8 \rightarrow 256$ σύμβολα

Δεν «καθιερώθηκε»

ASCII 8bit (σύμβολα στις θέσεις 128-255)			
10-δικό	16-δικό	2-δικό	Σύμβολο
128	80	10000000	A
129	81	10000001	B
130	82	10000010	Γ
131	83	10000011	Δ
132	84	10000100	E
133	85	10000101	Z
134	86	10000110	H
135	87	10000111	Θ
136	88	10001000	I
137	89	10001001	K
138	8A	10001010	Λ
139	8B	10001011	M
140	8C	10001100	N
141	8D	10001101	Ξ
142	8E	10001110	O
143	8F	10001111	Π
144	90	10010000	P
145	91	10010001	Σ
146	92	10010010	T
147	93	10010011	Υ
148	94	10010100	Φ
149	95	10010101	Χ
150	96	10010110	Ψ
151	97	10010111	Ω
152	98	10011000	α
153	99	10011001	β
154	9A	10011010	γ
155	9B	10011011	δ
156	9C	10011100	ε
157	9D	10011101	ζ

Ενδεικτικά

158	9E	10011110	η
159	9F	10011111	θ
160	A0	10100000	ι
161	A1	10100001	κ
162	A2	10100010	λ
163	A3	10100011	μ
164	A4	10100100	ν
165	A5	10100101	ξ
166	A6	10100110	ο
167	A7	10100111	π
168	A8	10101000	ρ
169	A9	10101001	σ
170	AA	10101010	ς
171	AB	10101011	τ
172	AC	10101100	υ
173	AD	10101101	φ
174	AE	10101110	χ
175	AF	10101111	ψ
176	B0	10110000	
177	B1	10110001	
178	B2	10110010	
179	B3	10110011	
180	B4	10110100	
181	B5	10110101	
182	B6	10110110	
183	B7	10110111	
184	B8	10111000	
185	B9	10111001	
186	BA	10111010	
187	BB	10111011	
188	BC	10111100	
189	BD	10111101	
190	BE	10111110	
191	BF	10111111	
192	C0	11000000	



Σύνολα συμβόλων (αλφάβητα)

ASCII 8 bit – Extended (2)

Παράδειγμα

Σχεδίαση γραφικών σε οθόνη κειμένου
(δεκαετίες 1980-1990)



Σύμβολα σχεδίασης		
Κώδικας ASCII (10-δικό)	16-δικό	Σύμβολο
201	C9	⌌
205	CD	=
187	BB	⌌
186	BA	⌌
178	B2	⌌
199	C7	⌌
196	C4	—
182	B6	⌌
32	20	κενό

Σύνολα συμβόλων (αλφάβητα)

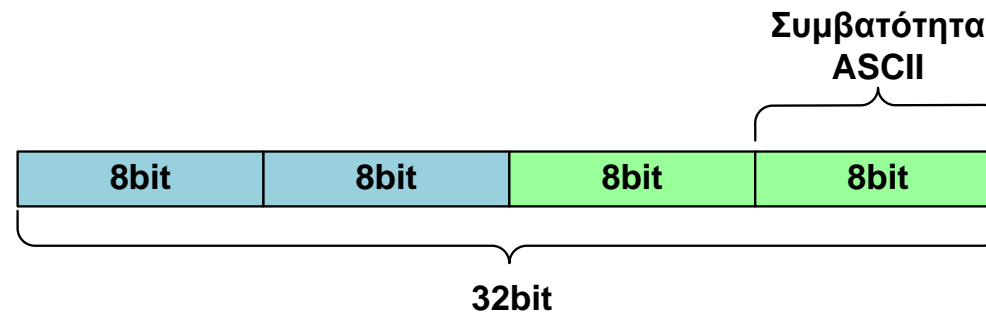
Unicode (1)

Ιδέα για ένα νέο αλφάβητο που θα περιέχει:

- α) τα σύμβολα όλων των γλωσσών
- β) ειδικά σύμβολα
- γ) σύμβολα για τη διατήρηση συμβατότητας με παλιά σύνολα συμβόλων (ASCII)

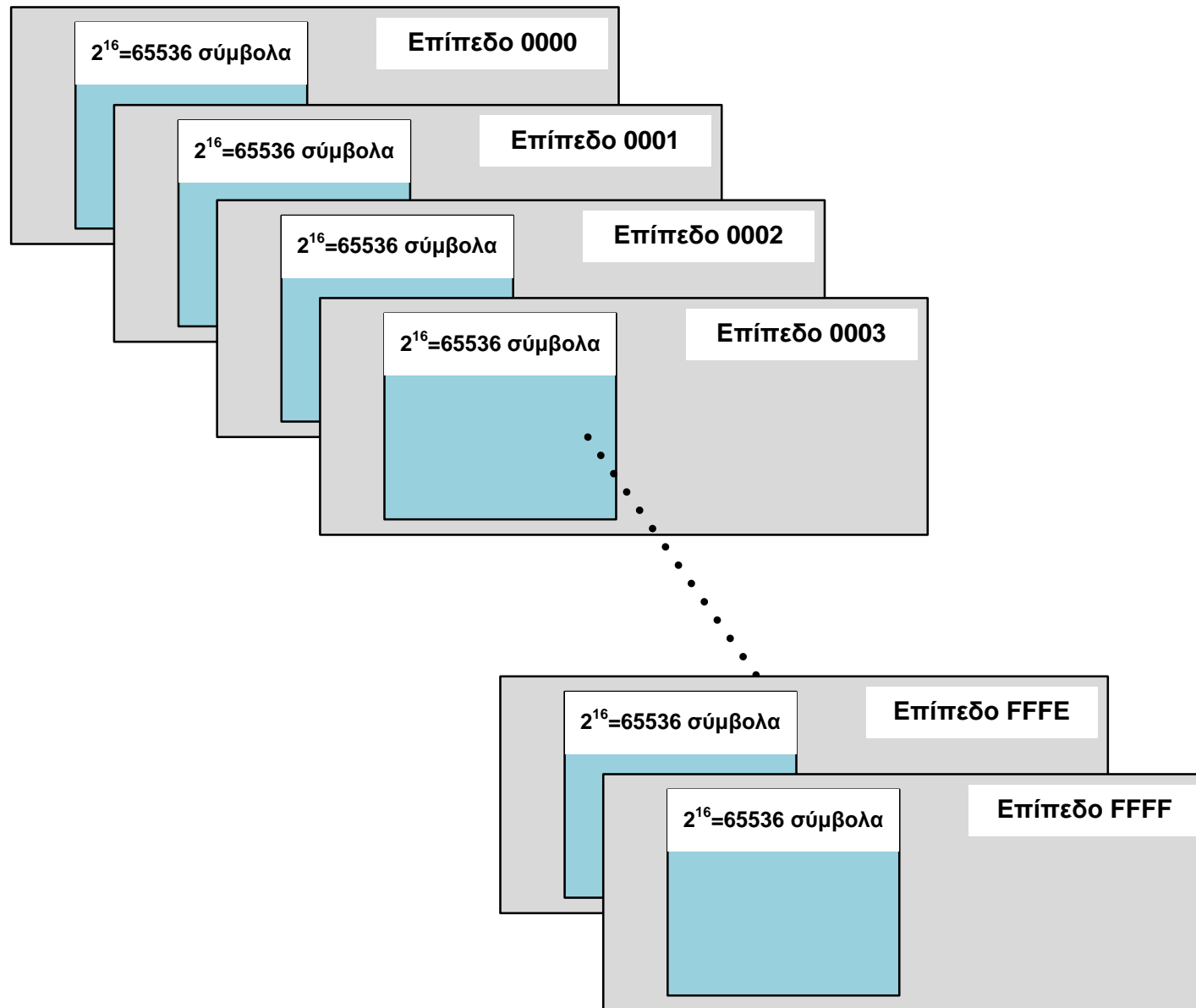
UNICODE (Universal Character Encoding)

UNICODE (ASCII συμβατότητα)	
	32bit
24bit	8bit (ASCII/ASCII extended)
000000000000000000000000	00000000 (αρχική τιμή)
000000000000000000000000	11111111 (τελική τιμή)



Σύνολα συμβόλων (αλφάβητα)

Unicode (2)



Τύποι δεδομένων (1)

Παράδειγμα

ΑΡΧΗ

Μεταβλητές Α, Β, Γ, ΑΘΡΟΙΣΜΑ

Διάβασε Α,Β,Γ

ΑΘΡΟΙΣΜΑ=Α+Β+Γ

Εμφάνισε ΑΘΡΟΙΣΜΑ

ΤΕΛΟΣ

Ο προγραμματιστής
επιλέγει τον τύπο
δεδομένων για κάθε
μεταβλητή

Συνηθισμένοι τύποι μεταβλητών στη γλώσσα C

Τύπος	Μέγεθος (bytes)	Μέγεθος (bits)	Εύρος τιμών
int	4	32	-2,147,483,648 ... 2,147,483,647
unsigned int	4	32	0 ... 4,294,967,295
short int	2	16	-32,768 ... 32767
unsigned short int	2	16	0 ... 65535
float	4	32	$3.4 \cdot 10^{-38}$... $3.4 \cdot 10^{+38}$
double	8	64	$1.7 \cdot 10^{-308}$... $1.7 \cdot 10^{+308}$
char	1	8	-128 ... 127
unsigned char	1	8	0 ... 255



Τύποι δεδομένων (2)

Παράδειγμα

ΑΡΧΗ

Μεταβλητές Α, Β, Γ, ΑΘΡΟΙΣΜΑ

Διάβασε Α, Β, Γ

ΑΘΡΟΙΣΜΑ = Α + Β + Γ

Εμφάνισε ΑΘΡΟΙΣΜΑ

ΤΕΛΟΣ

Ο προγραμματιστής
επιλέγει τον τύπο
δεδομένων για κάθε
μεταβλητή

Ποια πραγματική διαφορά έχουν οι παρακάτω
επιλογές τύπων δεδομένων;

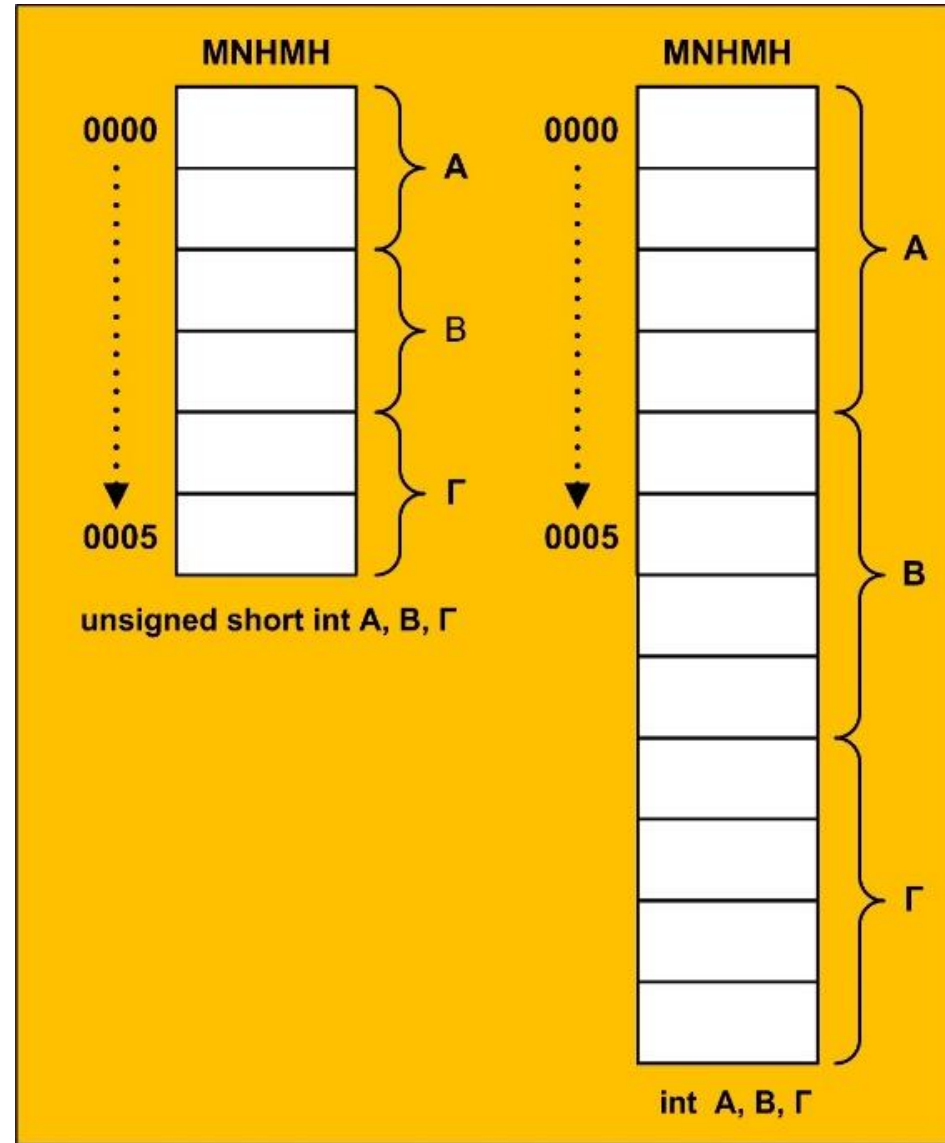
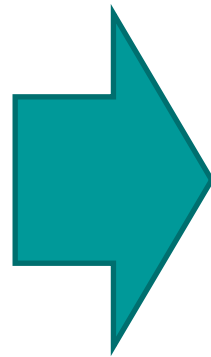
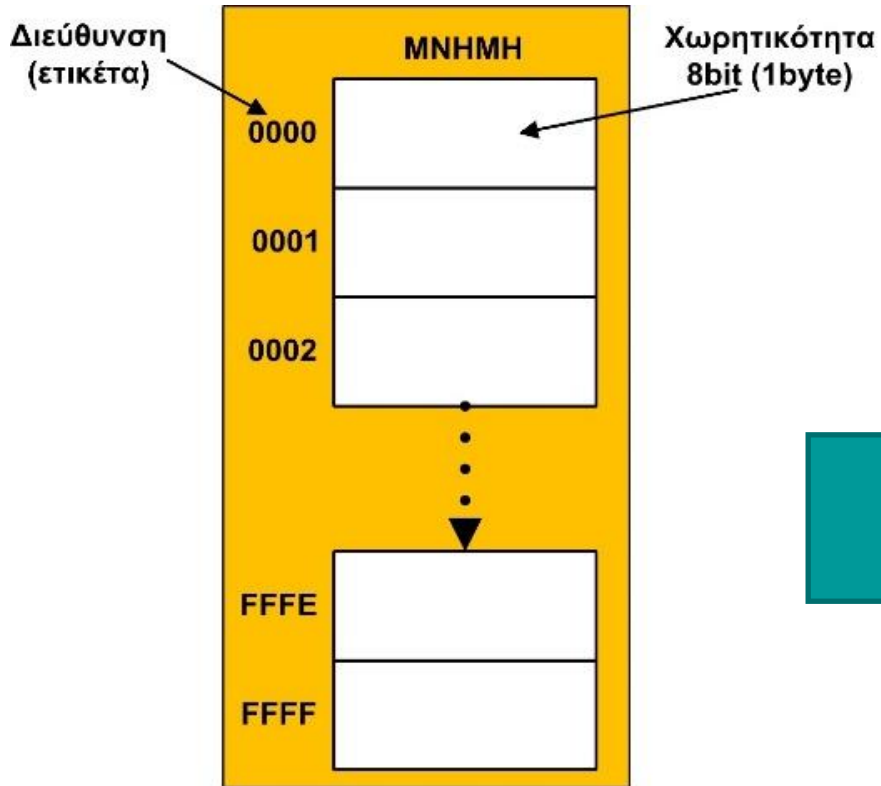
(α) unsigned short int Α, Β, Γ

(β) int Α, Β, Γ

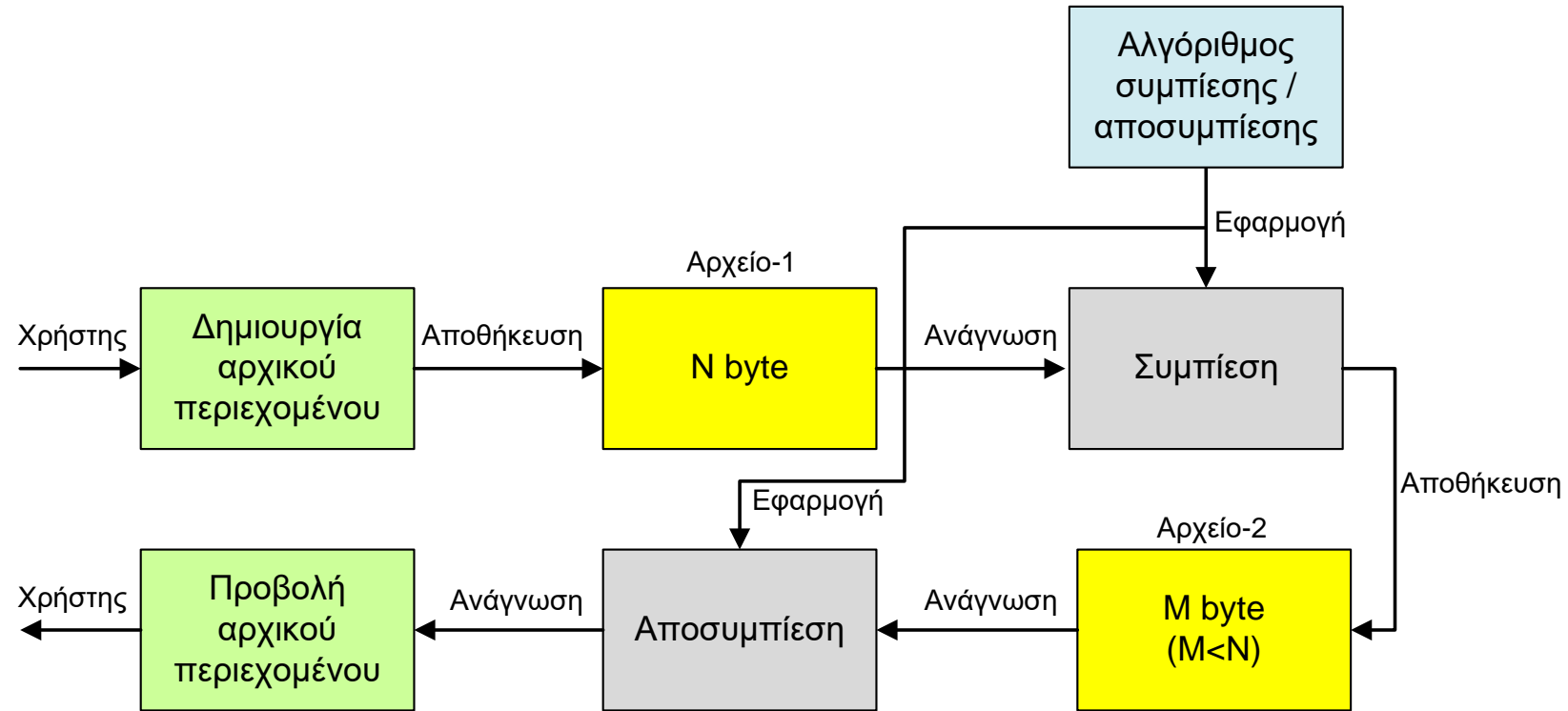
Τύποι δεδομένων (3)

Παράδειγμα

Γενικό μοντέλο μνήμης



Συμπίεση δεδομένων (1)



Τύποι συμπίεσης

- (α) συμπίεση χωρίς απώλειες
- (β) συμπίεση με απώλειες

Συμπίεση δεδομένων (2)

Συμπίεση χωρίς απώλειες

Αλγόριθμος κωδικοποίησης RLC (Run-Length Compression)

Περίπτωση 1

Έστω αρχείο με περιεχόμενο **DDDDDDDDFFFFFFAAABBBRRRR**

Πλήθος εμφάνισης συμβόλων

Σύμβολο D: συνεχόμενη εμφάνιση 7 φορές (κωδικοποίηση σε D07)

Σύμβολο F: συνεχόμενη εμφάνιση 6 φορές (κωδικοποίηση σε F06)

Σύμβολο A: συνεχόμενη εμφάνιση 3 φορές (κωδικοποίηση σε A03)

Σύμβολο B: συνεχόμενη εμφάνιση 3 φορές (κωδικοποίηση σε B03)

Σύμβολο R: συνεχόμενη εμφάνιση 4 φορές (κωδικοποίηση σε R04)

Η ακολουθία **DDDDDDDDFFFFFFAAABBBRRRR**, γίνεται **D07F06A03B03R04**

Συμπίεση $15/23 = .65 = 65\%$ του αρχικού, δηλαδή $100\% - 65\% = 35\%$

Συμπίεση δεδομένων (3)

Συμπίεση χωρίς απώλειες

Αλγόριθμος κωδικοποίησης RLC (Run-Length Compression)

Περίπτωση 2

Αν το αρχικό περιεχόμενο ήταν **ΚΚΚΑΑΑΓΓΓ**

Με την ίδια λογική το συμπιεσμένο αρχείο θα είχε τη μορφή **Κ03Α03Γ03**

Συμπίεση = 0%.

Περίπτωση 3

Αν το αρχικό περιεχόμενο ήταν **ΚΚΑΓΦΔ**

Το συμπιεσμένο θα είχε περιεχόμενο **Κ02Α01Γ01Φ01Δ01**

Συμπιεσμένο αρχείο > αρχικό (9 περισσότερα σύμβολα)

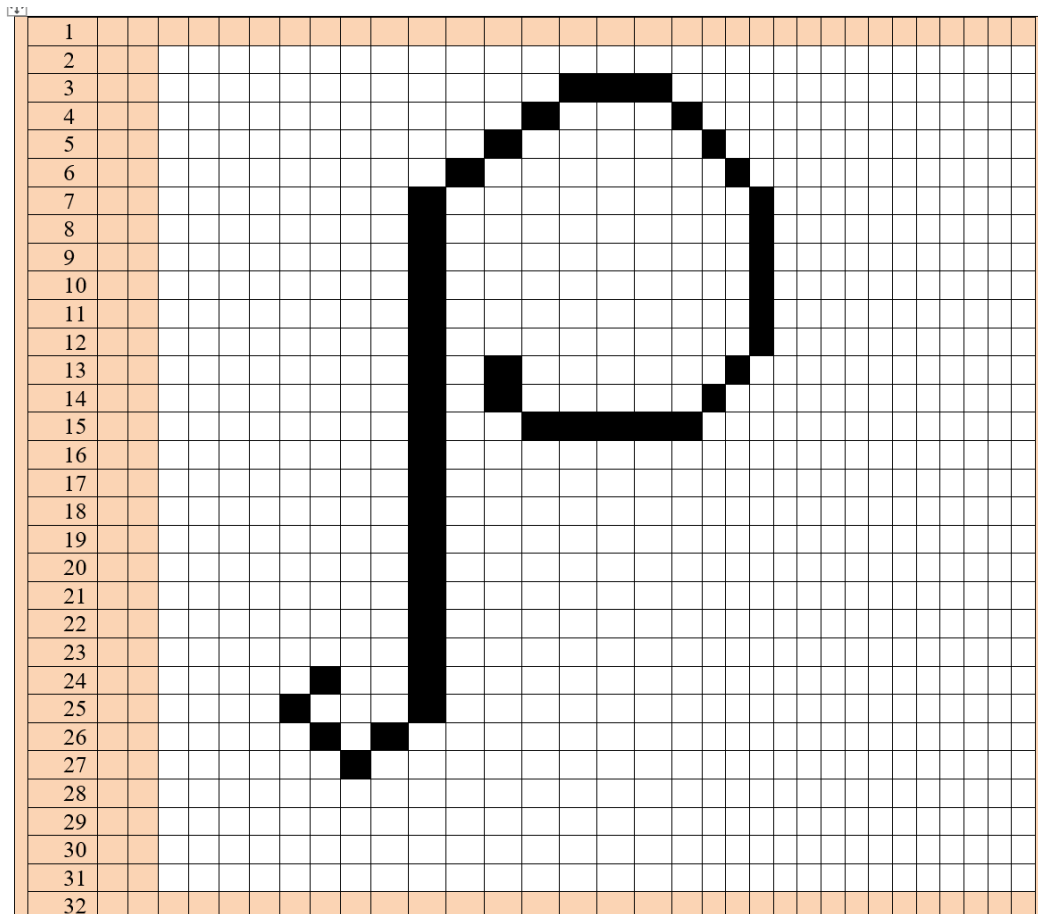


Συμπίεση δεδομένων (4)

Συμπίεση χωρίς απώλειες

Αλγόριθμος κωδικοποίησης RLC (Run-Length Compression)

Μελετήστε τη συμπίεση του συμβόλου που ακολουθεί



Συμπύεση δεδομένων (5)

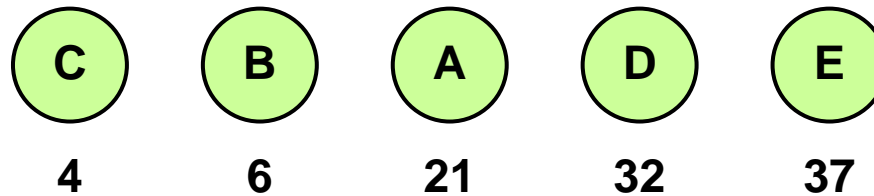
Κωδικοποίηση Huffman

Έστω κείμενο 1500 συμβόλων με ποσοστά (%) εμφάνισης ανά σύμβολο:

A	B	C	D	E
21	6	4	32	37

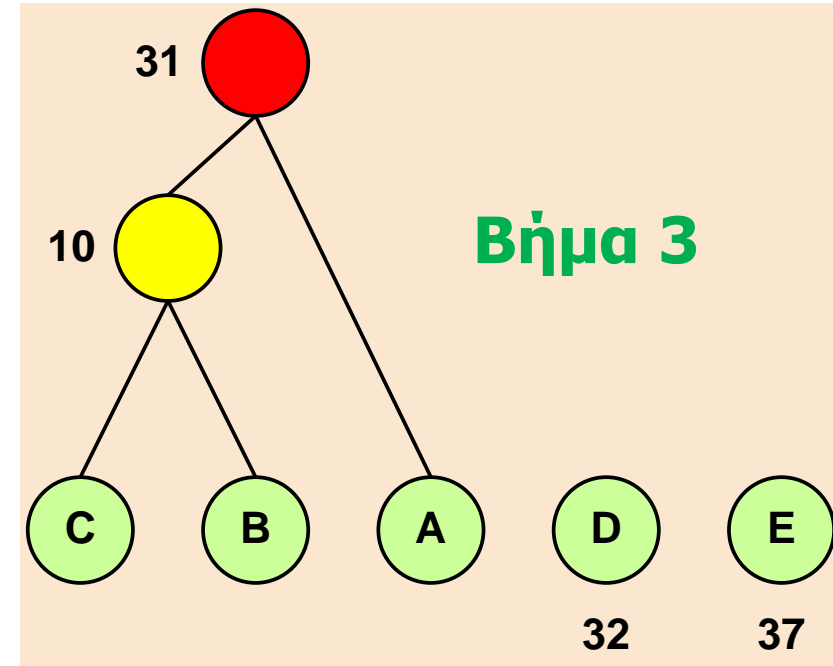
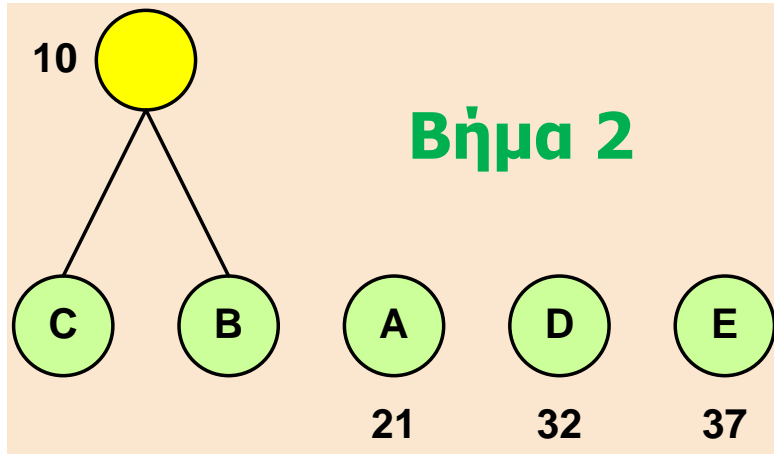
- Ο μικρότερος κώδικας θα αντιστοιχεί στο συχνότερα εμφανιζόμενο σύμβολο
- Επιλέγουμε κάθε φορά δύο κόμβους με τα μικρότερα ποσοστά

Βήμα 1



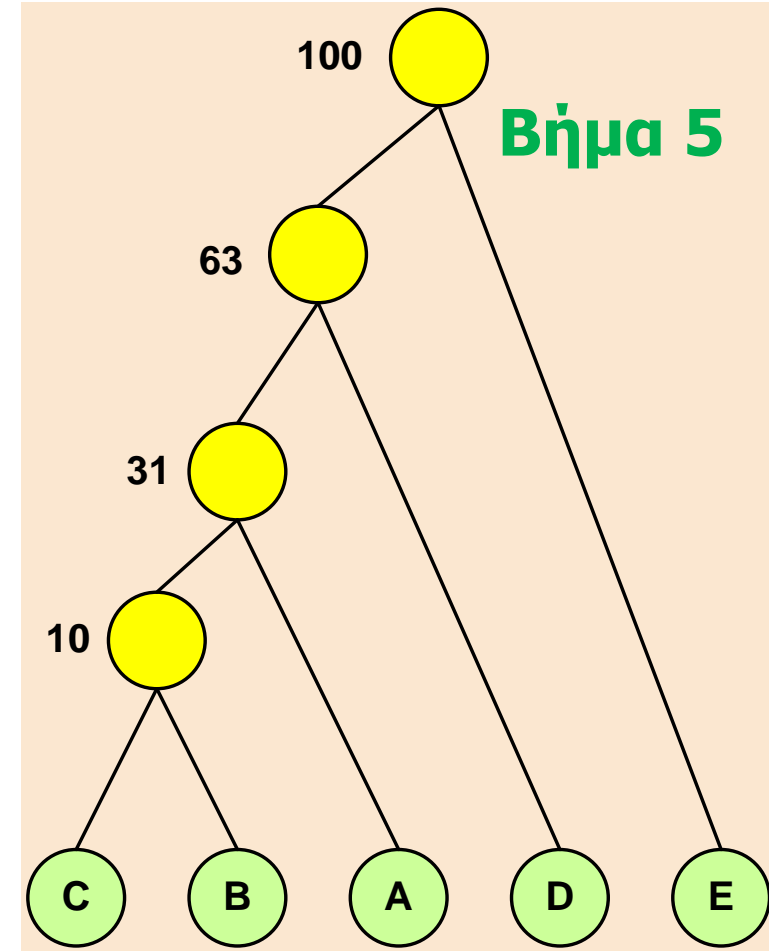
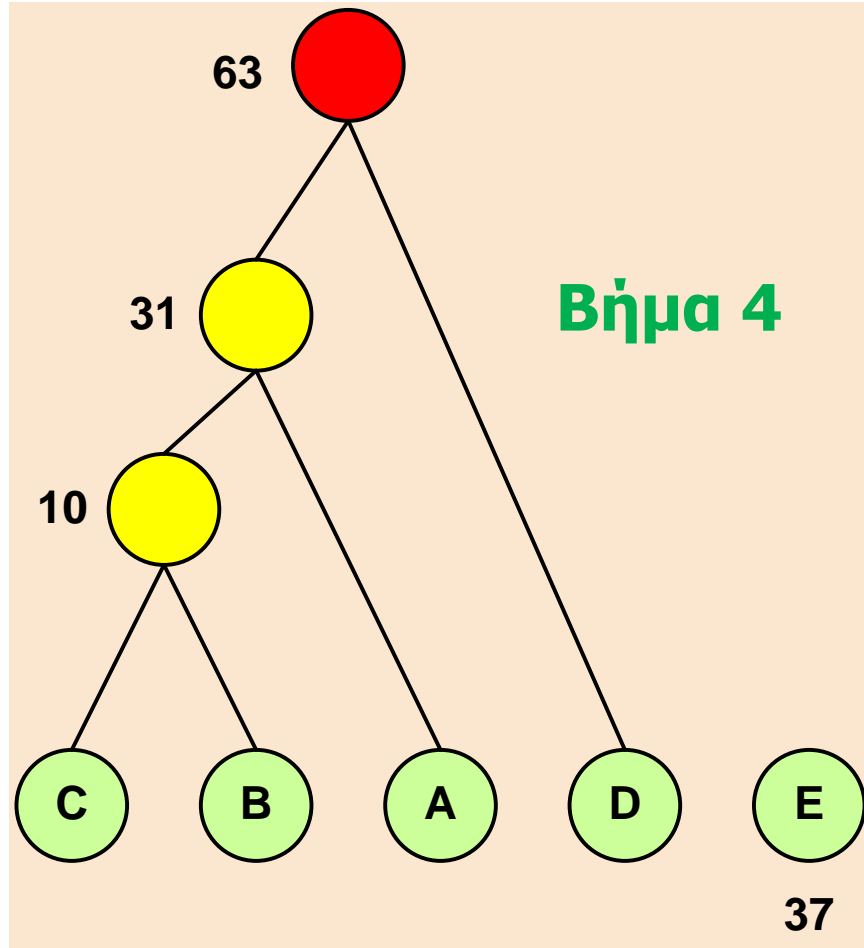
Συμπύεση δεδομένων (6)

Κωδικοποίηση Huffman



Συμπύεση δεδομένων (7)

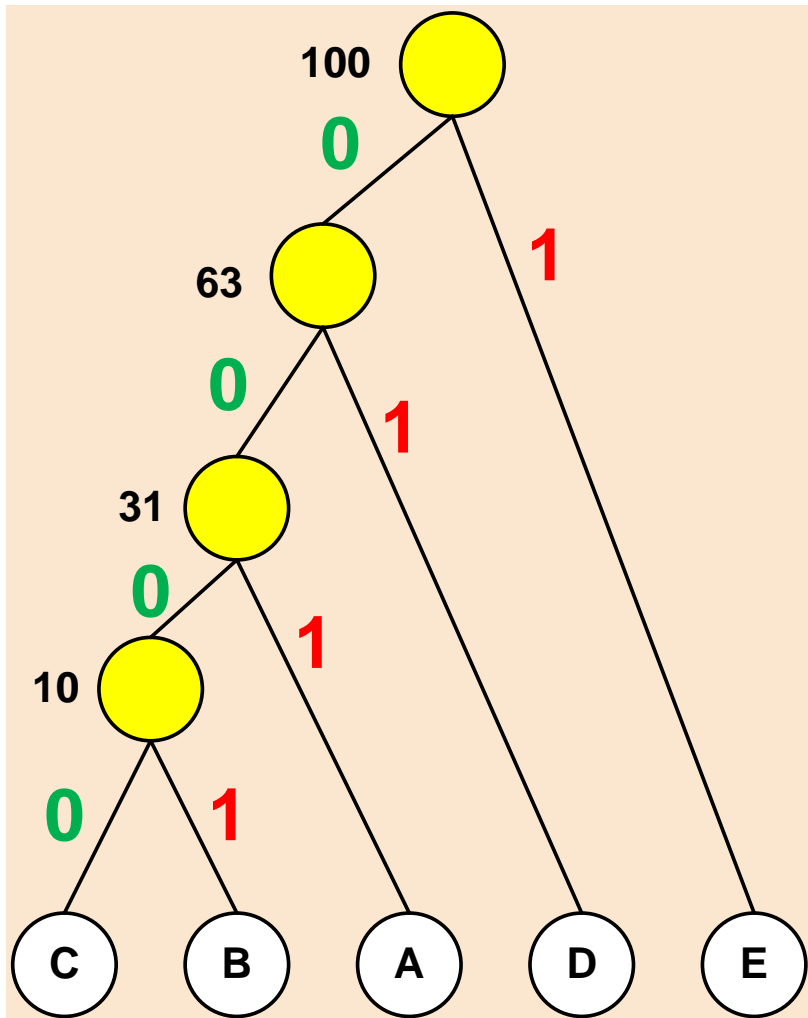
Κωδικοποίηση Huffman



Συμπίεση δεδομένων (8)

Κωδικοποίηση Huffman

Αρίθμηση των κλάδων



Τελική κωδικοποίηση συμβόλων

	C	B	A	D	E
%	4	6	21	32	37
Κωδικοποίηση	0000	0001	001	01	1



Ολοκλήρωση κεφαλαίου
Δείτε τις ασκήσεις από το βιβλίο

