

Κεφάλαιο 3^ο

Αναπαράσταση και Δομές Δεδομένων – Βάσεις Δεδομένων

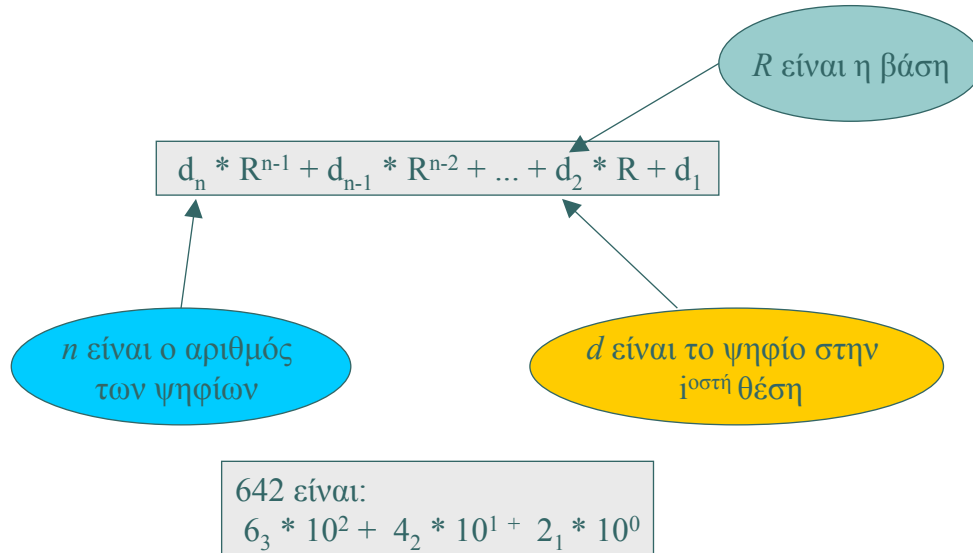
3.1 Συστήματα Αρίθμησης

Βασικό σύστημα αρίθμησης για την αναπαράσταση των δεδομένων στους Ηλεκτρονικούς Υπολογιστές είναι το δυαδικό σύστημα (binary arithmetic system). Η επιλογή δεν είναι τυχαία, αφού τα ηλεκτρονικά κυκλώματα που χρησιμοποιούνται στους Ηλεκτρονικούς Υπολογιστές μπορούν να ευρισκονται σε μία από δύο καταστάσεις: Ανοιχτό – κλειστό / Αληθές – ψευδές / Αγωγή ρεύματος - Διακοπή ρεύματος / Υψηλή τάση – Χαμηλή τάση / «0» ή «1». Τα ηλεκτρονικά κυκλώματα που μπορούν να βρίσκονται σε δύο καταστάσεις ονομάζονται ψηφιακά (digital). Η κατάλληλη διασύνδεση ψηφιακών κυκλωμάτων επιτρέπει την αποθήκευση και επεξεργασία δεδομένων σε ψηφιακή μορφή, τα οποία ονομάζονται δυαδικά ψηφία (binary digits) ή bits.

Το πιο γνωστό από τα συστήματα αρίθμησης είναι το δεκαδικό (decimal), το οποίο έχει ως βάση (base) το «10» και ως ψηφία (digits) τα 0,1,2,3,4,5,6,7,8 και 9. Για παράδειγμα ο αριθμός 642, στο δεκαδικό σύστημα αναλύεται ως εξής:

$$\begin{aligned}
 &+ 6 \times 10^2 = 6 \times 100 = 600 \\
 &+ 4 \times 10^1 = 4 \times 10 = 40 \\
 &+ 2 \times 10^0 = 2 \times 1 = 2 \quad = 642
 \end{aligned}$$

Ο γενικός τύπος, με βάση τον οποίο αναλύονται οι αριθμοί σε όλα τα συστήματα αρίθμησης είναι ο εξής:



Εικόνα 3.1. Σημειολογία θέσης στα συστήματα αρίθμησης

Τα συνηθέστερα αριθμητικά συστήματα εκτός από το δεκαδικό είναι τα εξής:

- **Δυαδικό σύστημα** (Βάση: το $R=2$, Σύμβολα: 0,1)

$$\begin{aligned}
 (10001)_2 &= 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 \\
 \text{(σε δεκαδικό)} &= 16 + 0 + 0 + 0 + 1 \\
 &= (17)_{10}
 \end{aligned}$$
- **Οκταδικό σύστημα** (Βάση: το $R=8$, Σύμβολα: 0,1,2,3,4,5,6,7)

$$\begin{aligned}
 (762)_8 &= 7 \times 8^2 + 6 \times 8^1 + 2 \times 8^0 \\
 \text{(σε δεκαδικό)} &= 448 + 48 + 2 \\
 &= (498)_{10}
 \end{aligned}$$
- **Δεκαεξαδικό σύστημα** (Βάση: το $R=16$, Σύμβολα: 0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F, όπου $A=10$, $B=11$, $C=12$, $D=13$, $E=14$ και $F=15$)

$$\begin{aligned}
 (3FB)_{16} &= 3 \times 16^2 + 15 \times 16^1 + 11 \times 16^0 \\
 \text{(σε δεκαδικό)} &= 768 + 240 + 11 \\
 &= (1019)_{10}
 \end{aligned}$$

Η τεχνική (αλγόριθμος) της μετατροπής ενός αριθμού από το δεκαδικό στο δυαδικό σύστημα έχει ως εξής: διαίρεση του αριθμού δια του 2 (τη βάση του δυαδικού συστήματος) και συνεχείς ακέραιες διαιρέσεις των παραγόμενων πηλίκων δια του 2 μέχρι την τελική διαίρεση 1 : 2 με πηλίκιο 0 και υπόλοιπο 1. Η δυαδική μορφή του αριθμού σχηματίζεται από τα υπόλοιπα των διαιρέσεων με αριστερότερο ψηφίο του αριθμού το τελευταίο υπόλοιπο και δεξιότερο το πρώτο.

Διαιρούμε συνεχώς το δεκαδικό νούμερο (π.χ. 34) με το **δύο** μέχρι το πηλίκιο να γίνει **0**.

- 34:2 → Πηλίκιο 17, Υπόλοιπο 0
- 17:2 → Πηλίκιο 8, Υπόλοιπο 1
- 8:2 → Πηλίκιο 4, Υπόλοιπο 0
- 4:2 → Πηλίκιο 2, Υπόλοιπο 0
- 2:2 → Πηλίκιο 1, Υπόλοιπο 0
- 1:2 → Πηλίκιο 0, Υπόλοιπο 1

Σχηματίζουμε τον αριθμό γράφοντας τα **υπόλοιπα** από το **τέλος** προς την **αρχή**:

Ο αριθμός $(34)_{10}$ είναι ο $(100010)_2$

Η αντίστροφη μετατροπή ενός αριθμού από δυαδική μορφή σε δεκαδική γίνεται προσθέτοντας τα δυαδικά ψηφία του αριθμού πολλαπλασιασμένα όμως με την αντίστοιχη δύναμη του 2. Οι θέσεις των ψηφίων ενός δυαδικού αριθμού αναπαριστούν μονάδες, δυνάδες, τετράδες, οκτάδες, δεκαεξάδες κλπ., όπως αντίστοιχα ενός δεκαδικού αριθμού αναπαριστούν μονάδες, δεκάδες, εκατοντάδες, χιλιάδες, κλπ.

Για το προηγούμενο παράδειγμα η δεκαδική μορφή του προκύπτει ως εξής:

$$\begin{aligned}
 (100010)_2 &= 1 \cdot 2^5 + 0 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 \\
 (\text{σε δεκαδικό}) &= 32 + 0 + 0 + 0 + 2 + 0 \\
 &= (34)_{10}
 \end{aligned}$$

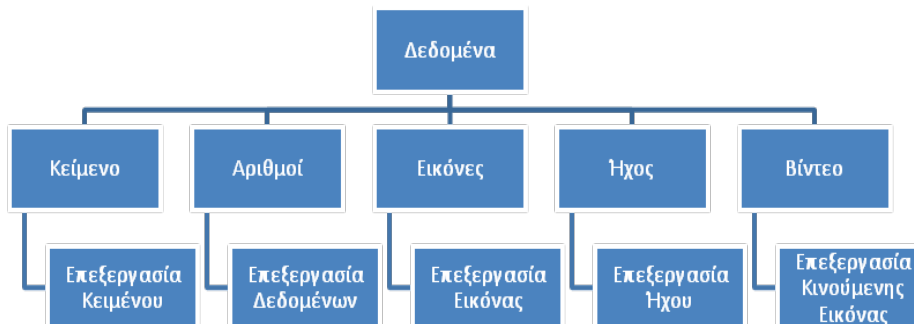
Για την μετατροπή δεκαδικών σε οκταδικό ή δεκαεξαδικό σύστημα είναι προτιμότερη η ενδιάμεση μετατροπή σε δυαδικό πρώτα και η εφαρμογή του κανόνα των 3 και 4 bits αντίστοιχα, όπως απεικονίζεται στην εικόνα 3.2.

$$\begin{array}{cccc}
 (101|111|011|100)_2 & & & \\
 \updownarrow & \updownarrow & \updownarrow & \updownarrow \\
 (5 & 7 & 3 & 4)_8 & & & \\
 & & & & & & \uparrow \\
 (1110|1000|1100)_2 & & & \\
 \updownarrow & \updownarrow & \updownarrow & \\
 (E & 8 & C)_{16} & & & &
 \end{array}$$

Εικόνα 3.2. Μετατροπή από δυαδικό σε οκταδικό και δεκαεξαδικό.

3.2 Τύποι Δεδομένων

Οι Ηλεκτρονικοί Υπολογιστές χρησιμοποιούνται για την επεξεργασία των δεδομένων, τα οποία συνήθως προέρχονται από το φυσικό κόσμο. Οι βασικοί τύποι δεδομένων με τους οποίους «ασχολούνται» οι Ηλεκτρονικοί Υπολογιστές ταξινομούνται στις κατηγορίες του παρακάτω σχήματος. Για την κάθε κατηγορία δεδομένων έχουν αναπτυχθεί μια σειρά από τεχνικές και τεχνολογίες για την επεξεργασία τους.



Εικόνα 3.3. Κατηγορίες Δεδομένων

3.3 Λογικές Πράξεις

Πολύ σημαντικό ρόλο στη λειτουργία ενός Ηλεκτρονικού Υπολογιστή, πέρα από την υιοθέτηση του δυαδικού συστήματος αρίθμησης, διαδραματίζουν οι λογικές

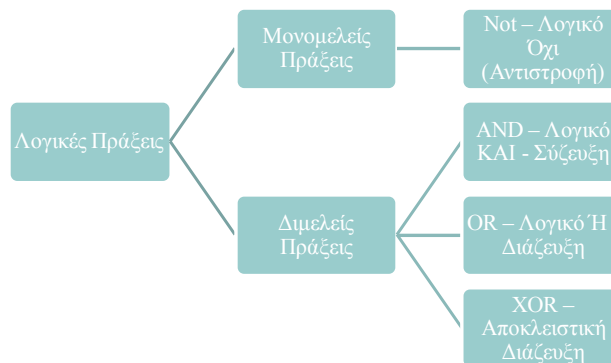
πράξεις ή δυαδική λογική και η υλοποίηση της μέσα από ηλεκτρονικά κυκλώματα. Όπως είναι γνωστό η τιμή ενός bit μπορεί να είναι 0 ή 1. Συνήθως το μηδέν «0» χαρακτηρίζεται ως η λογική τιμή ψευδής (false) και το ένα «1» ως η λογική τιμή αληθής (true). Με τον τρόπο αυτό μπορούμε να χρησιμοποιήσουμε τα bit και να εφαρμόσουμε λογικές πράξεις.



Εικόνα 3.4. Μονομελείς και Διμελείς πράξεις

Οι λογικές πράξεις με τη σειρά τους κατηγοριοποιούνται ανάλογα με το πλήθος των bit τα οποία δέχονται ως είσοδο. Πιο συγκεκριμένα, οι μονομελείς πράξεις (*unary operators*) αφορούν πράξεις με μία είσοδο (εφαρμογή σε ένα 1 bit), ενώ οι διμελείς πράξεις (*binary operators*) αφορούν πράξεις με δύο εισόδους (εφαρμογή σε 2 bit).

Οι πιο συνήθεις μονομελείς και διμελείς πράξεις είναι το «λογικό όχι» ή αντιστροφή (NOT), το «λογικό και» ή σύζευξη (AND), το «λογικό ή» ή «διάζευξη» (OR) και η αποκλειστική διάζευξη (XOR).

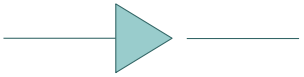


Εικόνα 3.5. Περιπτώσεις μονομελών και διμελών πράξεων

Για κάθε μια από τις λογικές πράξεις δίδεται στην συνέχεια ο πίνακας αλήθειας της, ο οποίος αναφέρει όλους τους δυνατούς συνδυασμούς ως δεδομένα εισόδου και τα αντίστοιχα δεδομένα εξόδου.

Πίνακας 3.1. Πράξη Αντιστροφή - Not

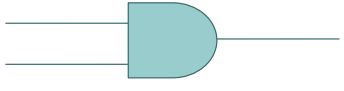
x	NOT x
1	0
0	1



- Έχει μόνο μια είσοδο
- Αντιστρέφει την είσοδο στην έξοδο

Πίνακας 3.2. Πράξη Και - And


x	y	x AND y
0	0	0
1	0	0
0	1	0
1	1	1



- Έχει δύο εισόδους
- Η έξοδος είναι λογικό 1, μόνο όταν και οι δυο εισοδοι είναι λογικό 1
- Έστω και μια είσοδος να βρίσκεται σε λογικό 0, τότε η έξοδος θα είναι σε λογικό 0

Πίνακας 3.3. Πράξη Η - Or

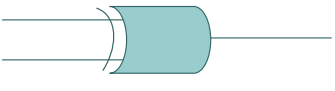
x	y	x OR y
0	0	0
1	0	1
0	1	1
1	1	1



- Έχει δύο εισόδους
- Η έξοδος είναι λογικό 0, μόνο όταν και οι δυο εισοδοι είναι λογικό 0
- Έστω και μια είσοδος να βρίσκεται σε λογικό 1, τότε η έξοδος θα είναι σε λογικό 1

Πίνακας 3.4. Πράξη XOR (eXclusive OR – αποκλειστικό Ή)

x	Y	x XOR y
0	0	0
1	0	1
0	1	1
1	1	0



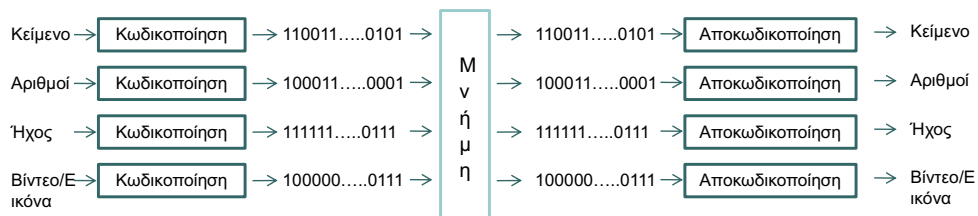
- Έχει δύο εισόδους
- Η έξοδος είναι λογικό 1, μόνο όταν οι δυο εισοδοί έχουν διαφορετική τιμή
- Η έξοδος είναι λογικό 0, μόνο όταν οι δυο εισοδοί έχουν την ίδια τιμή

Η δυαδική λογική είναι πολύ χρήσιμη κατά τη διάρκεια των αναζητήσεων στα διάφορα πληροφοριακά συστήματα, αφού επιτρέπει την επεξεργασία των κριτηρίων αναζήτησης και τον καθορισμό των επιθυμητών αποτελεσμάτων κάθε φορά.

3.4 Εσωτερική αναπαράσταση δεδομένων

Η μετάβαση των δεδομένων από το φυσικό κόσμο στον κόσμο των Ηλεκτρονικών Υπολογιστών πραγματοποιείται μέσα από τη διαδικασία της ψηφιοποίησης (*digitization*) ή της δυαδικής κωδικοποίησης της πληροφορίας, εφόσον στηρίζεται στην αναπαράσταση των πληροφοριών με την χρήση του δυαδικού κώδικα (*binary code*).

Η ελάχιστη μονάδα δυαδικής πληροφορίας είναι το λεγόμενο δυαδικό ψηφίο ή bit (*binary digit*) ενώ η μονάδα αποθήκευσης της δυαδικής πληροφορίας είναι η λεγόμενη ψηφιολέξη ή byte, η οποία αποτελείται από 8 bits (*bit pattern*).



Εικόνα 3.6. Κωδικοποίηση δεδομένων για την αποθήκευσή τους σε Ηλεκτρονικό Υπολογιστή

Η αναπαράσταση των δεδομένων δεν μπορεί να πραγματοποιηθεί με τη χρήση ενός μόνο bit. εξαιτίας της μεγάλης ποσότητας πληροφοριών που αποτελούν εικόνες, ήχο, κείμενο κ.λπ. Η λύση είναι η χρήση των Bit Patterns – Σχημάτων Bits. Ως bit pattern ορίζουμε τις ακολουθίες ή συμβολοσειρές από bit. Οι

συμβολοσειρές αποθηκεύονται στη μνήμη χωρίς να είναι γνωστό εκ των προτέρων ο τύπος των δεδομένων που αντιπροσωπεύουν. Η ερμηνεία του τύπου των δεδομένων είναι δουλειά των συσκευών εισόδου/εξόδου ή των προγραμμάτων / λογισμικών.

Πίνακας 3.5. Σχέση πλήθους συμβόλων και μήκος σχήματος bit (bit pattern)

Πλήθος συμβόλων ($\Pi=2^n$)	Αναπαράσταση σε δύναμη του 2 (2^n)	Μήκος σχήματος bit (n bits)
2	2^1	1
4	2^2	2
8	2^3	3
16	2^4	4
32	2^5	5
64	2^6	6
128	2^7	7
256	2^8	8
512	2^9	9
....	
16.384	2^{14}	14
32.768	2^{15}	15
65.536	2^{16}	16

Επιπλέον, ο όγκος των πληροφοριών και των δεδομένων μετρείται ,όπως έγινε αντιληπτό στην περίπτωση την Κύριας Μνήμης ή των περιφερειακών μέσω αποθήκευσης, σε πολλαπλάσια του bit ή του byte. Πιο συγκεκριμένα:

- 1 byte = 8 bits
- 1 KiloByte (KB) = 2^{10} Bytes = 1024 Bytes
- 1 MegaByte (MB) = 2^{10} KB = 2^{20} Bytes = 1.048.576 Bytes
- 1 Gigabyte (GB) = 2^{10} MB = 2^{20} KB = 2^{30} Bytes = 1.073.741.824 Bytes

- 1 TeraByte (TB) = 2^{10} GB = 2^{20} MB = 2^{30} KB = 2^{40} Bytes = 10.99.511.627.776 Bytes

Για λόγους απλοποίησης κατά τη διάρκεια των μετατροπών ισχύουν οι παρακάτω παραδοχές:

- 1 KiloByte (KB) \cong 1000 Bytes
- 1 MegaByte (MB) \cong 1000 KB
- 1 Gigabyte (GB) \cong 1000 MB
- 1 TeraByte (TB) \cong 1000 GB

3.4.1. Αναπαράσταση κειμένου

Για την αναπαράσταση των συμβόλων κειμένου έχουν σχεδιαστεί μια σειρά από σχήματα bit.

Κώδικας είναι μία αμφιμονοσήμαντη αντιστοίχιση συμβόλων ενός αλφαβήτου σε ακολουθία από δυαδικά ψηφία.

Ο σκοπός κάθε δυαδικού κώδικα κειμένου είναι να θέσει κάθε χαρακτήρα ενός κειμένου (γράμμα, αριθμό, σημείο στίξης) σε δυαδική μορφή. Ο κώδικας περιέχει εκτός από τις δυαδικές μορφές των χαρακτήρων και κάποιες επιπλέον δυαδικές μορφές για συγκεκριμένες λειτουργίες ελέγχου, π.χ. καμπανάκι (BEL – bell), εκτέλεση (CR - carriage return ή enter), αλλαγή γραμμής (LF - line feed), διαγραφή (DEL - delete), μια -θέση - πίσω (BS - backspace) κλπ. Αυτές οι επιπλέον δυαδικές μορφές ονομάζονται και μη-εκτυπώσιμοι (*nonprintable*) χαρακτήρες. Γενικά ένας κώδικας με N bits ανά χαρακτήρα έχει τη δυνατότητα να παραστήσει 2^N χαρακτήρες (βλέπε σχέση μήκους bit pattern με πλήθος συμβόλων). Επομένως όσο μεγαλύτερο είναι το μήκος της δυαδικής παράστασης τόσο περισσότερους χαρακτήρες έχει τη δυνατότητα ο κώδικας να χειριστεί.

Προκειμένου όλοι οι υπολογιστές να έχουν μια κοινή αναπαράσταση δεδομένων, δηλαδή κάθε αλφαριθμητικός χαρακτήρας να έχει ως αναπαράσταση συγκεκριμένη ακολουθία δυαδικών ψηφίων, δημιουργήθηκε ένα κοινό πρότυπο και συγκεκριμένα ο κώδικας ASCII (American Standard Code for Information Interchange – Αμερικάνικος Πρότυπος Κώδικας για την ανταλλαγή πληροφοριών) από το Αμερικανικό Ινστιτούτο Εθνικών Προτύπων.

Character	Bit pattern	Byte number	Character	Bit pattern	Byte number
A	01000001	65	Ο	10111100	188
B	01000010	66	.	00101110	46
C	01000011	67	:	00111010	58
a	01100001	97	\$	00100100	36
b	01100010	98	\	01011100	92
o	01101111	111	~	01111110	126
p	01110000	112	1	00110001	49
q	01110001	113	2	00110010	50
r	01110010	114	9	00111001	57
x	01111000	120	©	10101001	169
y	01111001	121	>	00111110	62
z	01111010	122	‰	10001001	137

Εικόνα 3.7. Κώδικας ASCII

Στον κώδικα ASCII χρησιμοποιούνται 7 ή 8 δυαδικά ψηφία (ανάλογα αν πρόκειται για τον απλό ή διευρυμένο κώδικα ASCII). Στους χαρακτήρες που κωδικοποιούνται με 7 δυαδικά ψηφία κατά ASCII περιλαμβάνονται χαρακτήρες ελέγχου, σημεία στίξης, τελεστές λογικών ή αριθμητικών πράξεων, πεζά και κεφαλαία γράμματα και αριθμοί. Με τον τρόπο αυτό κωδικοποιούνται $2^7 = 128$ διαφορετικοί χαρακτήρες. Επειδή ο αριθμός αυτός είναι πολύ περιορισμένος και δεν επαρκεί για να περιλάβει και τα γράμματα μιας επιπλέον γλώσσας (π.χ. ελληνικής ή κάποιας άλλης ευρωπαϊκής), χρησιμοποιείται ένα ακόμη δυαδικό ψηφίο, οπότε είναι δυνατή η κωδικοποίηση ακόμη 128 χαρακτήρων. Ο κώδικας ASCII των 8 ψηφίων έχει προτυποποιηθεί ως ISO-8859 και είναι γνωστός ως διευρυμένος κώδικας ASCII (*extended ASCII*). Η πιο ολοκληρωμένη προσπάθεια που αφορά την παγκόσμια κωδικοποίηση που θα καλύπτει ένα μεγάλο εύρος των πιο διαδομένων γραπτών γλωσσών, έγινε με το λεγόμενο κώδικα UNICODE, ο οποίος απαιτεί 2 δυαδικές λέξεις (bytes) ή 16 δυαδικά ψηφία για κάθε χαρακτήρα. Ο κώδικας αυτός έχει τυποποιηθεί ως ISO-10646. Επιπλέον, για την ολοκληρωτική επίλυση του θέματος κωδικοποίησης ο Διεθνής Οργανισμός Προτυποποίησης (*International Standard Organization - ISO*) έχει σχεδιάσει ένα κώδικα με 32 bit, οποίος μπορεί να αναπαραστήσει 4.294.967.296 συνολικά χαρακτήρες. Τέλος μια άλλη διαδεδομένη κωδικοποίηση του παρελθόντος, που αναπτύχθηκε από την IBM είναι το σύστημα EBCDIC (Extended Binary Code for Data Interchange). Ο κώδικας EBCDIC χρησιμοποιεί 8 bits ανά χαρακτήρα και επομένως μπορεί να παραστήσει $2^8=256$ χαρακτήρες.

Πίνακας 3.6. Συνοπτική παρουσίαση κωδίκων αναπαράστασης κειμένου

ASCII	Extended ASCII	UNICODE
<ul style="list-style-type: none"> • Χρήση 7 bits για την αναπαράσταση κάθε χαρακτήρα (Western European). Ξεκινά με το σύμβολο 0000000 (κενός χαρακτήρας), μέχρι το σύμβολο 1111111 (χαρακτήρας διαγραφής del) • Υπάρχουν 39 μη εκτυπώσιμοι χαρακτήρες ή όπως είναι γνωστοί χαρακτήρες ελέγχου • Οι αριθμητικοί χαρακτήρες είναι τοποθετημένοι πριν από τα γράμματα • Υπάρχουν πολλοί εκτυπώσιμοι χαρακτήρες • Τα γράμματα A ως Z (κεφαλαία γράμματα στην αγγλική γλώσσα) είναι πριν από τα πεζά γράμματα (a...z). • Τα κεφαλαία γράμματα ξεχωρίζουν από τα πεζά κατά ένα μόνο χαρακτήρα (το έκτο bit), π.χ. A=1000001, a=1100001, B=1000010, b=1100010 • Μεταξύ πεζών και κεφαλαίων υπάρχουν 6 ειδικοί χαρακτήρες • Σύνολο χαρακτήρων: $2^7 = 128$ χαρακτήρες 	<ul style="list-style-type: none"> • Προσθήκη ενός επιπλέον bit από τα αριστερά. Αύξηση του αριθμού των χαρακτήρων που μπορεί να αναπαρασταθούν κατά ένα συντελεστή 2 (256). • Κάθε κατασκευαστής χρησιμοποίησε το επιπλέον bit για να ορίσει μια σειρά από νέους χαρακτήρες, με αποτέλεσμα να «χαθεί» η έννοια της τυποποίησης και προτυποποίησης. • Στην Ελλάδα χρησιμοποιήθηκε για να οριστεί το ελληνικό αλφάβητο. ΕΛΟΤ 928: αποτελεί επέκταση του ASCII στο 8 bit από τον Ελληνικό Οργανισμό Τυποποίησης 	<ul style="list-style-type: none"> • Υπήρχε η ανάγκη για την ύπαρξη ενός κώδικα που θα υποστήριζε με ενιαίο τρόπο γλώσσες και αλφάβητα πέρα από την αγγλική • Ο κώδικας θα έπρεπε να έχει αυξημένη χωρητικότητα • Ο UNICODE ήταν το αποτέλεσμα της συνεργασίας των κατασκευαστών H/Y και αποτελείται από 16 bit • Μπορεί να αναπαραστήσει 65.536 (2^{16}) διαφορετικούς χαρακτήρες • Οι πρώτοι 128 χαρακτήρες στο Unicode πρότυπο αντιστοιχούν ακριβώς με το ASCII πρότυπο • Οι ASCII χαρακτήρες αντιστοιχούν στις τιμές 000016 έως 007F16 του Unicode • Αναπαρίσταται από 4 δεκαεξαδικά (Hex) ψηφία

3.4.2 Αναπαράσταση αριθμητικών δεδομένων

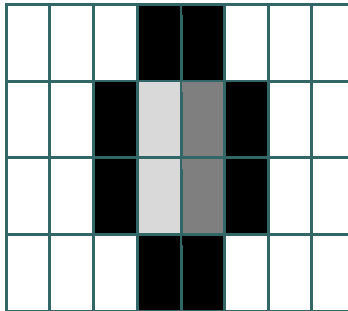
Στην περίπτωση κωδικοποίησης αριθμητικών ποσοτήτων είναι πιο εύλογο να χρησιμοποιηθεί κατ' ευθείαν η δυαδική παράσταση. Για την αναπαράσταση κατά ASCII π.χ. του αριθμού 65.535 απαιτούνται 5 bytes ή 40 bits, ενώ για τη δυαδική του παράσταση απαιτούνται μόνο 2 bytes ή 16 bits. Οι αριθμητικές αυτές ποσότητες έχουν εισαχθεί είτε με πληκτρολόγηση είτε αποτελούν αποτελέσματα επεξεργασίας κάποιου προγράμματος ή τέλος είναι οι διακριτές τιμές ενός ψηφιακού σήματος που προέκυψαν από τη δειγματοληψία και ψηφιοποίηση ενός αναλογικού σήματος (π.χ. έντασης φωνής, χρωματικής σύνθεσης και έντασης εικονοστοιχείων μιας εικόνας κλπ.). Αυτά τα αριθμητικά δεδομένα αποθηκεύονται σε δυαδική μορφή σε ένα αρχείο, που συνήθως ονομάζεται αρχείο δεδομένων (data file). Ακριβέστερη αναπαράσταση της τιμής του σήματος σημαίνει μεγαλύτερο πλήθος επιτρεπτών διακριτών τιμών και συνεπώς μεγαλύτερο αριθμό bits ανά δείγμα. Συνήθως χρησιμοποιούνται 16 bits (2 bytes) για τη δυαδική παράσταση κάθε δείγματος, συνεπώς 32768 θετικές και άλλες τόσες αρνητικές διακριτές τιμές για την ψηφιοποίηση του σήματος.

3.4.3 Αναπαράσταση εικόνας

Η αναπαράσταση των εικόνων γίνεται με δυο τρόπους. Ο πρώτος τρόπος καλείται ψηφιογραφική (*bitmap graphics*) και ο δεύτερος τρόπος διανυσματικός (*vector graphics*). Με την ψηφιογραφική μέθοδο κάθε εικόνα διαιρείται σε μια διάταξη pixel (picture elements – εικονοστοιχεία ή στοιχεία εικόνας). Το μέγεθος του pixel καθώς και ο αριθμός τους (στις δυο διαστάσεις) καθορίζονται από έναν παράγοντα που αποκαλείται ανάλυση εικόνας (resolution).

Παράδειγμα αναπαράστασης εικόνας με την Ψηφιογραφική Μέθοδο

Έστω η εικόνα του παρακάτω σχήματος, η οποία αποτελείται από 32 pixels και σε κάθε pixel αντιστοιχεί μια χρωματική αναπαράσταση από τις διαθέσιμες (άσπρο, ανοικτό γκρι, σκούρο γκρι και μαύρο).



Εικόνα 3.8. Εικόνα 32 pixels, 4 χρωματικές αποχρώσεις

Για κάθε pixel απαιτούνται 2 ψηφία (bits), ώστε να αναπαρασταθούν οι 4 συνδυασμοί (00=άσπρο, 01=ανοικτό γκρι, 10=σκούρο γκρι και 11=μάυρο). Επομένως για την αναπαράσταση της εικόνας του σχήματος θα πραγματοποιηθεί η παρακάτω κωδικοποίηση (Μέγεθος εικόνας: 32 εικονοστοιχεία * 2 bits ανά εικονοστοιχείο = 64bits ή 8 bytes).

	1	2	3	4	5	6	7	8
1	00	00	00	11	11	00	00	00
2	00	00	11	01	10	11	00	00
3	00	00	11	01	10	11	00	00
4	00	00	00	11	11	00	00	00

Εικόνα 3.9. Αναπαράσταση δεδομένων εικόνας με την ψηφιογραφική μέθοδο

Αν τώρα η αναπαράσταση αφορά έγχρωμες εικόνες τότε κάθε εικονοστοιχείο απαιτεί πολύ περισσότερα bits για την κωδικοποίηση του χρώματος του. Συνήθως κάθε εικονοστοιχείο περιγράφεται μέσα από τα τρία βασικά χρώματα και την ένταση του καθενός από αυτά (R=Red, G=Green, B=Blue). Για παράδειγμα αν χρησιμοποιούνται 8 bit για την αναπαράσταση του κάθε χρώματος τότε απαιτούνται 24 bit για κάθε pixel. Άρα για την εικόνα των προηγούμενων παραδειγμάτων θα απαιτούνται 32 pixels * 24 bits/pixel = 768 bits ή 96 bytes.

Το πρόβλημα με την ψηφιογραφική μέθοδο εμφανίζεται στην αλλαγή του μεγέθους των εικόνων. Αντίθετα η διανυσματική μέθοδος αναπαράστασης των εικόνων απεικονίζει τα σχήματα με γεωμετρικό τρόπο, με τη βοήθεια γραμμών, ορθογώνιων, ελλείψεων ή τόξων. Το μέγεθος των διανυσματικών αρχείων εξαρτάται από την πολυπλοκότητα των σχημάτων. Η συγκεκριμένη μέθοδος χρησιμοποιείται κυρίως για τη δημιουργία εικόνων (δυο και τριών διαστάσεων κυρίως), ενώ δεν συνίσταται για την διαδικασία ψηφιοποίησης εικόνων.

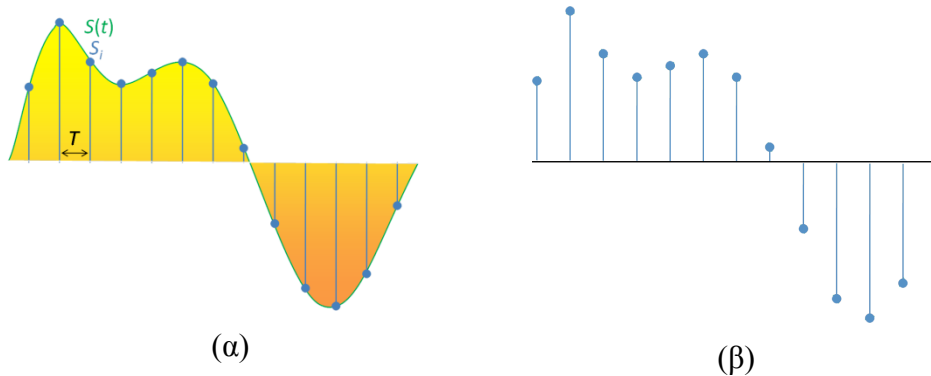
3.4.4 Αναπαράσταση ήχου

Για την αποθήκευση πραγματικών σημάτων στον Η/Υ είναι απαραίτητες δύο διαδικασίες, αυτή της δειγματοληψίας και εν συνεχεία αυτή της ψηφιοποίησης. Για τη δειγματοληψία απαιτείται ειδικός εξοπλισμός που συναντάται στις κάρτες ήχου, ο λεγόμενος αναλογικο-ψηφιακός μετατροπέας (A/D converter). Με τη βοήθεια αυτού του μετατροπέα, λαμβάνονται σε τακτά διαστήματα δείγματα του σήματος

(τιμές του συγκεκριμένου φυσικού μεγέθους, π.χ. ένταση του ήχου). Όσο μεγαλύτερη είναι η συχνότητα δειγματοληψίας τόσο πιστότερη είναι η αναπαράσταση του σήματος από τα δείγματά του, αλλά βέβαια τόσο μεγαλύτερος είναι και ο απαιτούμενος χώρος αποθήκευσης αλλά και ο ρυθμός μετάδοσης στην περίπτωση ψηφιακής μεταφοράς του σήματος από Η/Υ σε Η/Υ. Γι' αυτό απαιτείται ένας συμβιβασμός ανάμεσα στην ποιότητα και το κόστος. Πρακτικά εφαρμόζεται ο λεγόμενος νόμος του Shannon, σύμφωνα με τον οποίο η συχνότητα δειγματοληψίας πρέπει να είναι τουλάχιστον διπλάσια της μεγαλύτερης συχνότητας του σήματος. Στη συνέχεια με την ψηφιοποίηση μετατρέπεται κάθε δείγμα – αριθμός σε ένα δυαδικό αριθμό καθορισμένου μήκους. Και πάλι είναι προφανές ότι όσο περισσότερα bits δίνονται για την αναπαράσταση των δειγμάτων τόσο μεγαλύτερη είναι η ακρίβεια αναπαράστασης. Και πάλι όμως αύξηση των bits ανά δείγμα έχει ως συνέπεια την αύξηση των αποθηκευτικών χώρων και των ρυθμών μετάδοσης.

Μέσα από το παράδειγμα που ακολουθεί αναπαρίστανται με γραφικό τρόπο τα βήματα για την αποθήκευση του ήχου, την αναπαράσταση δεδομένων και την αναπαραγωγή του.

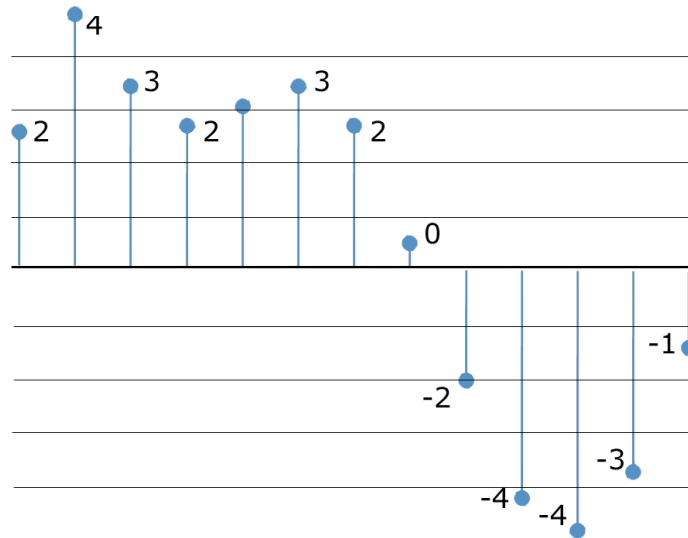
Βήμα 1. Δειγματοληψία του αναλογικού σήματος του ήχου. Η Δειγματοληψία (sampling) είναι η μέτρηση της τιμής του σήματος σε ίσα χρονικά διαστήματα. Αυτό το σταθερό χρονικό διάστημα αποτελεί τη λεγόμενη περίοδο δειγματοληψίας, το αντίστροφο της οποίας είναι η συχνότητα δειγματοληψίας.



Εικόνα 3.10. Δειγματοληψία αναλογικού σήματος ήχου. Για κάθε $s_i \rightarrow$ λαμβάνουμε μια τιμή. Περίοδος δειγματοληψίας T

Βήμα 2. Κβάντωση του δείγματος. Γενικά Κβάντωση (quantization) ενός φυσικού μεγέθους ονομάζεται η ιδιότητα του μεγέθους αυτού να λαμβάνει καθορισμένες και όχι τυχαίες τιμές. Ένα χαρακτηριστικό παράδειγμα αποτελούν οι τροχιές των ηλεκτρονίων γύρω από τον πυρήνα του ατόμου ή των πλανητών γύρω από τον Ήλιο. Οι τροχιές αυτές είναι κβαντισμένες δηλαδή ένα ηλεκτρόνιο ή ένας πλανήτης δεν μπορεί να βρεθεί σε τυχαία θέση μέσα στο χώρο που περιβάλλει τον

πυρήνα ή τον Ήλιο. Η διαδικασία λοιπόν της Κβάντωσης ενός δείγματος αφορά την αντιστοίχιση μιας τιμής του δείγματος πάνω σε ένα καθορισμένο σύνολο τιμών. Στο συγκεκριμένο παράδειγμα η κβάντωση αφορά ένα σύνολο τιμών από -4 ως και 4 με βήμα 1 (-4, -3, -2, -1, 0, 1, 2, 3, 4).



Εικόνα 3.11. Διαδικασία κβάντωσης

Βήμα 3. Οι κβαντισμένες τιμές στην συνέχεια μετατρέπονται σε δυαδικά σχήματα. Τα δυαδικά σχήματα αποθηκεύονται στην περιφερειακή μνήμη και μπορούν να ανασυνθέσουν το ηχητικό σήμα μέσω της αντίστροφης διαδικασίας.

Για την διαδικασία μετατροπής υιοθετείται το παρακάτω σχήμα bit, όπου το πρώτο bit δείχνει το πρόσημο και τα υπόλοιπα 3 την τιμή. Επομένως:

$$0000 = 0$$

$$0001 = 1 \quad 0010 = 2 \quad 0011 = 3 \quad 0100 = 4$$

$$1001 = -1 \quad 1010 = -2 \quad 1011 = -3 \quad 1100 = -4$$

Επομένως, οι τιμές S_i μετατρέπονται ως εξής:

$$S_1 = 2 = 0010, S_2 = 4 = 0100, S_3 = 3 = 0011, S_4 = 2 = 0010, S_5 = 3 = 0011, S_6 = 3 = 0011, S_7 = 2 = 0010, S_8 = 0 = 0000, S_9 = -2 = 1010, S_{10} = -4 = 1100, S_{11} = -4 = 1100, S_{12} = -3 = 1011, S_{13} = -1 = 1001$$

Τυπικές τιμές για μεν τη δειγματοληψία των ηχητικών σημάτων είναι 44,1 kHz (44.100 δείγματα ανά sec), για δε τη ψηφιοποίηση 16 bits ανά δείγμα. Σύμφωνα με τις τιμές αυτές, για την αποθήκευση 40 λεπτών ενός στερεοφωνικού ηχητικού σήματος απαιτούνται :

$44100 \text{ δείγματα} / \text{sec} \times 16 \text{ bits} / \text{δείγμα} \times 2 \text{ κανάλια} = 1.411.200 \text{ bits}$ ή 176.400 bytes για κάθε sec, συνεπώς $176400 \times 60 \times 40 = 423.360.000 \text{ bytes}$ ή 413.437 kbytes για τα 40 λεπτά.

3.4.5 Αναπαράσταση βίντεο

Παρόμοια μπορεί να υπολογιστεί και ο απαιτούμενος χώρος αποθήκευσης για ψηφιακό video. Στην περίπτωση του video τα δείγματα είναι πλαίσια (frames). Μια τυπική συχνότητα δειγματοληψίας είναι 30 fps (frames per sec). Μια τυπική ανάλυση πλαισίων video σε μορφή PAL είναι 768 x 576 pixels ανά frame. Αν θεωρηθεί βάθος χρώματος 24 bits (true color), τότε σύμφωνα με τις τιμές αυτές, για την αποθήκευση 2 ωρών video απαιτείται η εξής χωρητικότητα:

$30 \text{ frames} / \text{sec} \times 768 \times 576 \text{ pixels} / \text{frame} \times 24 \text{ bits} / \text{pixel} = 418.037.760 \text{ bits}$ ή $52.254.720 \text{ bytes}$ περίπου 50 Mbytes για κάθε sec, συνεπώς $50 \text{ M} \times 3600 \times 2 = 360 \text{ Gbytes}$ για τις 2 ώρες. Μια τέτοια χωρητικότητα είναι απαγορευτική ακόμη και για τους δίσκους DVD, γι' αυτό απαιτείται οπωσδήποτε μια προηγούμενη συμπίεση του σήματος και στη συνέχεια αποθήκευση.

3.5 Δομές Δεδομένων

Δομή Δεδομένων ΔΔ (Data Structure) είναι κάθε σύνολο δεδομένων με συγκεκριμένη οργάνωση. Η ανάκτηση της πληροφορίας γίνεται μέσω κάποιου μηχανισμού, ο οποίος κάνει χρήση του συγκεκριμένου αυτού τρόπου οργάνωσης με ταχύτητα και αποτελεσματικότητα.

Παραδείγματα ΔΔ είναι

- οι *πίνακες (tables)*, συλλογές μεταβλητών του ίδιου τύπου (πχ. φυσικών αριθμών)
- οι *εγγραφές (records / structures)*, περιγραφές με βάση κάποια χαρακτηριστικά ενός συγκεκριμένου τύπου αντικειμένων (πχ. βιβλίων)
- τα *αρχεία (files)*, συλλογές εγγραφών ομοειδών αντικειμένων (πχ. κατάλογος βιβλίων)
- οι *Βάσεις Δεδομένων (BD) – Database*, δομές συνδεδεμένων μεταξύ τους αρχείων (πχ. περιβάλλον βιβλιοθήκης με καταλόγους βιβλίων, χρηστών και δανεισμών)

3.5.1 Μεταβλητές

Όπως φάνηκε στο Κεφάλαιο 2 τα δεδομένα των προγραμμάτων αποθηκεύονται προσωρινά σε θέσεις μνήμης. Οι θέσεις μνήμης στις γλώσσες υψηλού επιπέδου αντιστοιχούν στις λεγόμενες μεταβλητές. Η μεταβλητή ενός προγράμματος έχει ένα μοναδικό όνομα που αντιστοιχεί στη διεύθυνση της θέσης μνήμης και ένα περιεχόμενο (πχ. αριθμός, χαρακτήρας, αλυσίδα χαρακτήρων κ.λπ.). Το όνομα της μεταβλητής μπορεί να είναι ένα γράμμα ή αλυσίδα αλφαριθμητικών ψηφίων, στην οποία όμως το πρώτο είναι γράμμα και από την οποία αποκλείονται αριθμητικοί τελεστές (+, -, *, /) και σημεία στίξης, π.χ. A, B, X, A1, NAME, LENGTH κ.λπ. Το περιεχόμενο μιας μεταβλητής μπορεί να είναι ένα απλό αντικείμενο (αριθμός, χαρακτήρας, αλυσίδα χαρακτήρων κλπ), αλλά και μια ολόκληρη δομή, όπως ένας πίνακας ή μια εγγραφή.

3.5.2 Πίνακες

Πίνακας είναι μια σειρά διαδοχικών θέσεων μνήμης, στις οποίες μπορούν να αποθηκευτούν ομοειδή δεδομένα (αριθμοί, χαρακτήρες, αλυσίδες χαρακτήρων κ.λπ.), ενώ ο χρόνος πρόσβασης (*access time*) σε οποιοδήποτε στοιχείο του πίνακα είναι σταθερός, ανεξάρτητα από τη θέση του. Το όνομα του πίνακα ακολουθεί τις γενικές συμβάσεις των ονομάτων των μεταβλητών. Ο προσδιορισμός ενός συγκεκριμένου στοιχείου του πίνακα γίνεται μέσω του ονόματος του πίνακα και της τάξης του στοιχείου μέσα σε παρενθέσεις ή αγκύλες, π.χ. A[5] ή A(5) δηλώνει το 5ο στοιχείο του πίνακα A. Οι πίνακες μπορεί να είναι μιας διάστασης (στήλη ή γραμμή), δυο διαστάσεων (MxN) ή και περισσότερων διαστάσεων.



Εικόνα 3.12. Πίνακας μιας διάστασης N θέσεων

3.5.3 Εγγραφές (Records)

Μια εγγραφή (*record*) είναι μια συλλογή από σχετικά μεταξύ τους στοιχεία πιθανώς διαφορετικών τύπων, η οποία έχει μοναδικό όνομα. Κάθε στοιχείο μιας εγγραφής ονομάζεται πεδίο (*field*). Το πεδίο είναι το μικρότερο στοιχείο ενός επωνύμου συνόλου δεδομένων με κάποια σημασία. Στα πεδία μπορούν να αντιστοιχίζονται τιμές (διαφορετικών τύπων) και να προσπελούνται όταν είναι απαραίτητο. Το πεδίο εκείνο που χαρακτηρίζει με μοναδικό τρόπο μια εγγραφή ονομάζεται πεδίο – κλειδί (*key*).

Η βασική διαφορά μεταξύ μιας εγγραφής και ενός πίνακα είναι ότι στο πίνακα οι τιμές πρέπει να είναι του ίδιου τύπου, ενώ στις εγγραφές μπορεί να είναι του ίδιου τύπου ή διαφορετικού.

```
Record Book {      Title : string
                  Author : string
                  Pages: int
                  YearOfPub : int
                  }
book1 as Book;
book1.Title = “Η Φόνισσα” , book1.Author = “Αλέξανδρος
Παπαδιαμάντης”, book1.Pages = 200, book1.YearOfPub=1950
```

3.5.4 Αρχεία

Ενώ οι μεταβλητές και οι πίνακες αναφέρονται στην κεντρική μνήμη και αφορούν στην προσωρινή αποθήκευση δεδομένων και αποτελεσμάτων κατά την εκτέλεση προγραμμάτων, τα αρχεία αποτελούν δομές δεδομένων διαφόρων βαθμών πολυπλοκότητας τα οποία αποθηκεύονται μόνιμα σε κάποια περιφερειακή μνήμη (μαγνητικό δίσκο κ.λπ.).

Το αρχείο είναι ανοικτό σε προσθήσεις, διαγραφές και τροποποιήσεις εγγραφών. Άλλες λειτουργίες αρχείων είναι η ταξινόμηση ενός αρχείου με βάση ένα ή περισσότερα πεδία του, καθώς και η συγχώνευση αρχείων. Τα αρχεία είναι οργανωμένα σειριακά ή άμεσα, οπότε υφίστανται αντίστοιχου τύπου πρόσβαση. Τα σειριακά (*sequential*) αρχεία είναι έτσι οργανωμένα ώστε η πρόσβαση σε κάποια εγγραφή να γίνεται μετά από την υποχρεωτική ανάγνωση όλων των προηγούμενων εγγραφών. Χρησιμοποιείται κυρίως για μικρό όγκο πληροφορίας ή στις περιπτώσεις που γίνεται συνήθως ανάγνωση ενός μεγάλου ποσοστού συνεχόμενων εγγραφών. Τα άμεσα (*random*) αρχεία είναι έτσι οργανωμένα ώστε η πρόσβαση σε κάποια εγγραφή να γίνεται άμεσα με την κατ' ευθείαν ανεύρεση του αύξοντα αριθμού της εγγραφής. Χρησιμοποιείται κυρίως για μεγάλο όγκο πληροφορίας, ιδίως στις περιπτώσεις που απαιτείται γρήγορη επικοινωνία. Τέλος υπάρχει μια ειδική οργάνωση αρχείων, σειριακή με δείκτη (*indexed sequential*) που βασίζεται στη λεγόμενη ευρετηρίαση. Κάποιο από τα πεδία χρησιμοποιείται ως πεδίο ταξινόμησης των εγγραφών, παράγοντας ένα παράλληλο μικρότερο αρχείο, με μοναδικά πεδία το ευρετηριακό πεδίο και τον αύξοντα αριθμό της εγγραφής. Η αναζήτηση της εγγραφής γίνεται με βάση το ευρετηριακό πεδίο, το οποίο παραπέμπει στη σχετική εγγραφή.

Ευρετήριο		Α/Α	Αριθμός Λογαριασμού	Όνομα	Ποσό	...
002010	01	01	002010	Μαρία Καλτσά	300,00	...
020101	02	02	020101	Νίκος Σχοινάς	1200,00	...
022020	03	03	022020	...		
212121	04	04	212121			
334411	05	05	334411			
.....			
555555	119	119	555555			

Εικόνα 3.13. Αρχείο και ευρετήριο

3.6 Βάσεις δεδομένων – Databases

Βάση δεδομένων είναι ένα σύνολο αλληλοσυνδεδεμένων αρχείων στο οποίο αποθηκεύεται, οργανώνεται και συντηρείται ένα σύνολο πληροφοριών και δεδομένων.

Μια βάση δεδομένων (data base) είναι μια συλλογή δεδομένων τα οποία σχετίζονται με λογικό, αλλά όχι απαραίτητα και με φυσικό τρόπο.

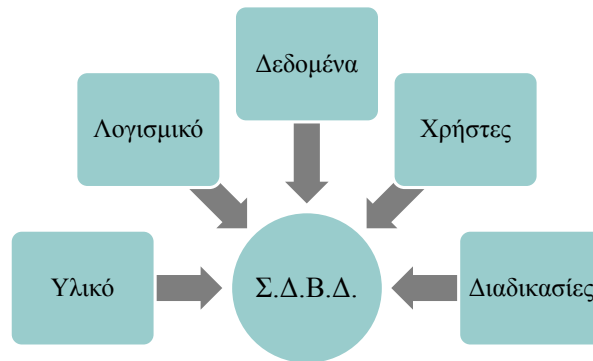
Μια βάση δεδομένων μπορεί να δημιουργηθεί και να συντηρηθεί είτε χειρόγραφα, είτε με χρήση ηλεκτρονικών υπολογιστών. Επιπλέον, μπορεί να έχει οποιοδήποτε μέγεθος και κυμαινόμενη πολυπλοκότητα. Για παράδειγμα μπορεί να είναι:

- Μια λίστα ονομάτων και διευθύνσεων με λίγες εκατοντάδες εγγραφές.
- Ένας κατάλογος βιβλιοθήκης με εκατομμύρια εγγραφές βιβλίων, αποθηκευμένες υπό διαφορετικές κατηγορίες (όνομα συγγραφέα, θέμα βιβλίου, τίτλο βιβλίου) ταξινομημένες κατά αλφαβητική σειρά σε κάθε κατηγορία.
- Η Βάση Δεδομένων της Εφορείας, όπου για 5 εκατομμύρια φορολογούμενους υπάρχουν πληροφορίες αρκετών δεκάδων Kilobytes.

3.6.1 Σύστημα Διαχείρισης Βάσης Δεδομένων

Το Σύστημα Διαχείρισης Βάσης Δεδομένων (DBMS – Database Management System) είναι το σύστημα εκείνο που μεσολαβεί μεταξύ του χρήστη και της βάσης,

παρέχοντας όλα τα απαραίτητα εργαλεία για την αποτελεσματική διαχείρισή της. Πιο συγκεκριμένα το Σ.Δ.Β.Δ. ορίζει, δημιουργεί και συντηρεί βάσεις δεδομένων. Αποτελείται από μια συλλογή προγραμμάτων που επιτρέπουν στους χρήστες να δημιουργήσουν και να συντηρήσουν μια βάση δεδομένων.



Εικόνα 3.14. Σύστημα Διαχείρισης Βάσης Δεδομένων

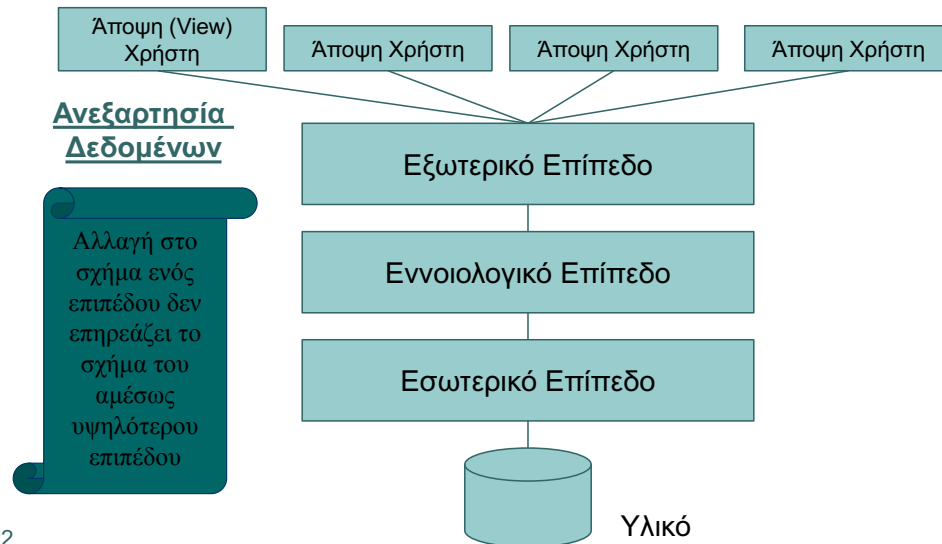
Όπως φαίνεται και στην πιο πάνω εικόνα οι έννοιες που συνθέτουν ένα Σύστημα Διαχείρισης Βάσης Δεδομένων είναι οι εξής:

- **Υλικό:** Το υλικό είναι το φυσικό υπολογιστικό σύστημα το οποίο επιτρέπει τη φυσική προσπέλαση των δεδομένων.
- **Λογισμικό:** Το λογισμικό είναι το πραγματικό πρόγραμμα που επιτρέπει στους χρήστες να προσπελάζουν, να συντηρούν, και να ενημερώνουν τα δεδομένα.
- **Δεδομένα:** Τα δεδομένα μιας βάσης δεδομένων αποθηκεύονται με φυσικό τρόπο σε συσκευές αποθήκευσης. Τα δεδομένα αποτελούν ξεχωριστή οντότητα από το λογισμικό που τα προσπελάζει..
- **Χρήστες:** Τελικοί Χρήστες (Διαχειριστές ΒΔ, Κανονικοί Χρήστες) – Προγράμματα Εφαρμογών.
- **Διαδικασίες:** Είναι το σύνολο των διαδικασιών ή κανόνων που πρέπει να έχουν οριστεί με σαφή τρόπο και να ακολουθούνται από τους χρήστες της βάσης δεδομένων.

Επιπλέον η αρχιτεκτονική ενός Συστήματος Διαχείρισης Βάσεων Δεδομένων οργανώνεται στα εξής επίπεδα:

- **Εσωτερικό Επίπεδο:** Το Εσωτερικό Επίπεδο (*internal level*) προσδιορίζει το που θα αποθηκευτούν τα δεδομένα στην αποθηκευτική συσκευή.

- **Εννοιολογικό Επίπεδο:** Το Εννοιολογικό Επίπεδο (*conceptual*) καθορίζει την λογική άποψη των δεδομένων. Σε αυτό το επίπεδο ορίζονται το μοντέλο δεδομένων και τα διαγράμματα ροής.
- **Εξωτερικό Επίπεδο:** Το Εξωτερικό Επίπεδο (*external level*) αλληλεπιδρά απευθείας με το χρήστη. Μετατρέπει τα δεδομένα που έρχονται από το εννοιολογικό επίπεδο σε μια μορφή που να είναι πιο οικεία στους χρήστες.



22

Εικόνα 3.15. Αρχιτεκτονική Συστήματος Διαχείρισης Βάσεων Δεδομένων

3.6.2 Μοντέλα Βάσεων Δεδομένων και ιστορική εξέλιξη

Κατά καιρούς και πάντα σε σχέση με την εξέλιξη της τεχνολογίας αλλά και των αναγκών προτάθηκαν μια σειρά από μοντέλα για την οργάνωση των Βάσεων Δεδομένων. Τα πιο γνωστά από αυτά είναι τα παρακάτω:

- **Ιεραρχικό Μοντέλο:** Για την οργάνωση των δεδομένων το Ιεραρχικό Μοντέλο (*hierarchical model*) ακολουθεί την τοπολογία ενός δένδρου.
- **Δικτυακό Μοντέλο:** Για την οργάνωση των δεδομένων το Δικτυακό Μοντέλο (*network model*) ακολουθεί την τοπολογία ενός γράφου¹.
- **Σχισιακό Μοντέλο:** Για την οργάνωση των δεδομένων το Σχισιακό Μοντέλο (*relation model*) κάνει χρήση πινάκων (ή σχέσεων) που **σχετίζονται μεταξύ τους**.

¹ Σύνολο κόμβων και βελών όπου οι κόμβοι απεικονίζουν τύπους αντικειμένων και τα βέλη τύπους σχέσεων.

- **Αντικειμενοστραφές Μοντέλο:** Στο Αντικειμενοστραφές Μοντέλο (*object oriented model*), με την βοήθεια των πινάκων ορίζονται οι οντότητες (αντικείμενα) και οι μεταξύ τους σχέσεις.

Τα παραπάνω μοντέλα ακολούθησαν την τεχνολογική εξέλιξη των δυνατοτήτων των Συστημάτων Διαχείρισης Βάσεων Δεδομένων, όπως αυτή παρουσιάζεται στον παρακάτω πίνακα.

Πίνακας 3.7. Χαρακτηριστικά Συστημάτων Διαχείρισης Βάσεων Δεδομένων ανά γενιά

Ονομασία	Περιγραφή χαρακτηριστικών
1950: Πρώτη Γενιά - Αρχεία σε Ταινίες	<ul style="list-style-type: none"> • Αποθήκευση σε κάρτες και ταινίες (σειριακή επεξεργασία)
1960: Δεύτερη Γενιά - Αρχεία σε Δίσκους	<ul style="list-style-type: none"> • Θεμελίωση του Επιστημονικού Πεδίου Βάσεων Δεδομένων - Συστήματα Δικτύου και Ιεραρχικά • Ανάπτυξη Συστημάτων Β.Δ. για Ιδιωτική Χρήση • Εμφάνιση του πρώτου γενικού-σκοπού Σ.Δ.Β.Δ.: <i>Integrated Data Store (GE)</i> • Ακολουθούν: <i>Information Management System (IMS) IBM, SABRE Airline Reservation System (AA+IBM)</i>
1970: Τρίτη Γενιά - Προ-Σχεσιακή	<ul style="list-style-type: none"> • Καθαρός διαχωρισμός μεταξύ «φυσικής» και «λογικής» οργάνωσης ΒΔ - Αναγνώριση του ρόλου της Μοντελοποίησης • Θεωρητικός ορισμός του Σχεσιακού Μοντέλου (Edgar Codd, IBM, San Jose)
1980-1990: Τέταρτη Γενιά - Σχεσιακό Μοντέλο και Συστήματα	<ul style="list-style-type: none"> • Υψηλού επιπέδου, ισχυρές γλώσσες επεξεργασίας, <i>SQL</i> • Εμφάνιση Αντικειμενοστραφών Συστημάτων, <i>Ευφρών Συστημάτων</i>
1990-2000: Πέμπτη Γενιά - Μετα-Σχεσιακή Εποχή	<ul style="list-style-type: none"> • Εμπορικά Αντικειμενοστραφή Συστήματα • Πολυμέσα, Συστήματα στο <i>Internet</i>

3.6.3 Πλεονεκτήματα – Μειονεκτήματα Βάσεων Δεδομένων

Η χρήση Βάσεων δεδομένων για την οργάνωση, αποθήκευση, προσπέλαση και συντήρηση δεδομένων παρουσιάζει κατά κύριο λόγο μόνο πλεονεκτήματα (βλέπε πίνακα που ακολουθεί). Ως μειονεκτήματα σημειώνονται το αρχικό κόστος απόκτησης και παραμετροποίησης ενός Σ.Δ.Β.Δ. και η μετέπειτα διαχείριση και συντήρησή του. Παρόλα αυτά, με τη σημερινή κυριαρχία του Διαδικτύου και του

Cloud Computing, η πρόσβαση σε ένα Σ.Δ.Β.Δ. για την ανάπτυξη μιας βάσης δεδομένων είναι εύκολη, ενώ απαιτεί σχετικά χαμηλό κόστος. Τα αρχεία θα μπορούσαν να προτιμηθούν ως μια εύκολη λύση για τη γρήγορη καταχώρηση δεδομένων, ωστόσο καθώς αυξάνει ο όγκος των πληροφοριών και των χρηστών που αποκτούν πρόσβαση σε αυτά θα παρουσιάζονται όλα τα μειονεκτήματα που παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 3.8. Πλεονεκτήματα – Μειονεκτήματα Βάσεων Δεδομένων

Πλεονεκτήματα	Μειονεκτήματα
<ul style="list-style-type: none"> • Ταυτόχρονη προσπέλαση από πολλούς χρήστες • Το Σ.Δ.Β.Δ. περιλαμβάνει λογισμικό για έλεγχο ταυτόχρονης προσπέλασης (concurrency control) και εξασφαλίζει σωστό αποτέλεσμα ενημέρωσης • Έλεγχος Πλεονασμών – Αποφυγή επαναλήψεων - Οικονομία χώρου και μόχθου - Συμβατότητα μεταξύ δεδομένων • Περιορισμός μη εξουσιοδοτημένης προσπέλασης - Το Σ.Δ.Β.Δ. περιλαμβάνει λογισμικό ασφάλειας και εξουσιοδότησης (security & authorisation) - Διαφορετικοί περιορισμοί προδιαγράφονται για διαφορετικούς χρήστες • Παροχή μηχανισμών τήρησης εφεδρικών αντιγράφων και ανάκαμψης (backup) - Αντιμετώπιση (ανάκαμψη) μετά από βλάβες υλικού ή λογισμικού • Αυτοπεριγραφή - Ένα Σ.Δ.Β.Δ. δεν περιέχει μόνο την ίδια τη βάση, αλλά και τον πλήρη ορισμό της (κατάλογος συστήματος) • Το λογισμικό του Σ.Δ.Β.Δ. είναι ανεξάρτητο από συγκεκριμένη εφαρμογή • Ανεξαρτησία Προγραμμάτων - Δεδομένων • Η δομή των αρχείων δεδομένων αποθηκεύεται ξεχωριστά από το λογισμικό προσπέλασης - Αλλαγή στη δομή των αρχείων δεν επηρεάζει το λογισμικό του Σ.Δ.Β.Δ. • Υποστήριξη Πολλαπλών Όψεων των Δεδομένων - Διαφορετικοί χρήστες απαιτούν διαφορετική προοπτική (όψη) της Β.Δ. - Μια όψη μπορεί να είναι ένα υποσύνολο της Β.Δ. ή να περιέχει εικονικά (virtual) δεδομένα που 	<ul style="list-style-type: none"> • Δαπανηρή προμήθεια και συντήρηση • Απαιτήσεις σε υλικό και προσωπικό • Καθυστέρηση στη διαχείριση δεδομένων
	Λόγοι που επιλέγονται τα Αρχεία για την οργάνωση των δεδομένων
	<ul style="list-style-type: none"> • Δεν αναμένονται αλλαγές στις εφαρμογές και στη δομή της Β.Δ. • Υπάρχουν πειστικές απαιτήσεις λειτουργίας σε πραγματικό χρόνο • Δεν απαιτείται προσπέλαση πολλών χρηστών στα δεδομένα
	Μειονεκτήματα Αρχείων
	<ul style="list-style-type: none"> • Πλεονασμός δεδομένων (data redundancy) • Ασυνέπεια δεδομένων (inconsistency) • Δύσκολη αναζήτηση (search) • Δυσκολία διαμοιρασμού (data sharing) • Αδυναμία ταυτόχρονης πρόσβασης (concurrent access) • Ανομοιομορφία (heterogeneity)

<p>παράγονται από τα αρχεία αλλά δεν αποθηκεύονται</p>	<ul style="list-style-type: none"> • Δυσκολία στην εφαρμογή κανόνων ακεραιότητας (<i>integrity rules</i>) • Δυσκολία στην τήρηση αντιγράφων ασφαλείας (<i>backup</i>)
--	---

3.6.4 Ρόλοι σε ένα περιβάλλον Σ.Δ.Β.Δ.

Οι βασικοί ρόλοι χρηστών σε ένα Σ.Δ.Β.Δ. είναι οι παρακάτω:

- **Διαχειριστής Β.Δ.**
 - εξουσιοδοτεί προσπέλαση στη Β.Δ.
 - είναι υπεύθυνος για προβλήματα ασφάλειας, απόδοσης του συστήματος
 - είναι υπεύθυνος για τη διαχείριση πόρων
- **Σχεδιαστές Β.Δ.**
 - είναι υπεύθυνοι για τον προσδιορισμό δεδομένων
 - αναπτύσσουν τις διαφορετικές όψεις της βάσης που καλύπτουν τις απαιτήσεις διαφορετικών χρηστών
- **Τελικοί χρήστες**
 - άτομα που η εργασία τους απαιτεί προσπέλαση δεδομένων
- **Αναλυτές Συστημάτων και Προγραμματιστές**
 - αναπτύσσουν προδιαγραφές και υλοποιούν προκαθορισμένες δοσοληψίες βάσει των απαιτήσεων των τελικών χρηστών

3.6.5 Γλώσσα SQL

Η Δομημένη Γλώσσα Ερωτημάτων ή Structured Query Language (*SQL*) είναι μια γλώσσα για τη διαχείριση σχεσιακών βάσεων δεδομένων. Περιλαμβάνει εκφράσεις με τις οποίες ο χρήστης μπορεί να καθορίσει τη δομή/σχήμα της βάσης αλλά και εκφράσεις με τις οποίες μπορεί να προσθέσει, να τροποποιήσει και να διαγράψει τα δεδομένα. Επιπλέον, η βασική της λειτουργία είναι η δημιουργία ερωτημάτων (*queries*) για τη λήψη συγκεκριμένων δεδομένων από τις αποθηκευμένες πληροφορίες.

Οι πιο γνωστές εντολές της γλώσσας SQL είναι οι παρακάτω:

- **Select:** Επιλέγει συγκεκριμένες τιμές από τους πίνακες με βάση διάφορα κριτήρια.
- **Insert:** Εισάγει νέες εγγραφές στους διάφορους πίνακες

- **Update:** Τροποποιεί εγγραφές με νέα δεδομένα
- **Delete:** Διαγράφει εγγραφές

3.6.6 Παράδειγμα βάσης δεδομένων

Στο παράδειγμα που ακολουθεί παρουσιάζεται μια απλή βάση δεδομένων για τη λειτουργία μιας βιβλιοθήκης. Η μορφή της βάσης δεδομένων είναι απλουστευμένη σε σχέση με τα σύγχρονα λογισμικά που προσφέρονται σήμερα. Το σχήμα της βάσης δεδομένων ακολουθεί το σχεσιακό μοντέλο, όπου τα δεδομένα οργανώνονται σε πίνακες, οι οποίοι περιέχουν εγγραφές (*records*). Η κάθε εγγραφή είναι μια συλλογή από συσχετιζόμενα πεδία (*fields*) τα οποία περιέχουν μια τιμή.

Ο πίνακας των βιβλίων χρησιμεύει στην εισαγωγή των βιβλιογραφικών πληροφοριών σχετικά με τα βιβλία που διαθέτει η βιβλιοθήκη. Ως πεδία (*fields*) του συγκεκριμένου πίνακα, ορίζονται τα εξής:

- Κωδικός (πεδίο κλειδί)
- Τίτλος
- Υπότιτλος
- Συγγραφέας
- Έτος Έκδοσης
- Εκδοτικός Οίκος
- ISBN

Πίνακας 3.9. Πίνακας Βιβλίων

Κωδικός	Τίτλος	Υπότιτλος	Συγγραφέας	Έτος Έκδ.	Εκδ/κός Οίκος	ISBN
001	Η Φόνισσα		Αλέξανδρος Παπαδιαμάντης	1998	Νεφέλη	
002	Οδηγός της Java 2		Steven Haines	2001	Γκιούρδας	960-387-0412
003	Σύγχρονη βιολογία	εισαγωγή στη μελέτη των οργανισμών	Νικολάου Σοφ. Χριστοδουλάκη	1994	Πατάκης	960-360-0482
....						

Ο πίνακας των αντιτύπων περιέχει τις πληροφορίες σχετικά με τα αντίτυπα από κάθε βιβλίο που διαθέτει η βιβλιοθήκη. Τα αντίτυπα είναι οι φυσικές οντότητες οι οποίες δανείζονται στον τελικό χρήστη. Ως πεδία του συγκεκριμένου πίνακα, ορίζονται τα εξής:

- Κωδικός (πεδίο κλειδί)
- Βιβλίο (πεδίο διασύνδεσης με τον πίνακα των βιβλίων, ένα βιβλίο μπορεί να έχει ένα ή περισσότερα αντίτυπα $I \times N$, FK: Foreign Key)
- Τοποθεσία
- Περιορισμός Δανεισμού
- Ημέρες Δανεισμού
- Κράτηση
- Ειδικοί Περιορισμοί

Πίνακας 3.10. Πίνακας Αντιτύπων

Κωδικός	Βιβλίο	Τοποθεσία	Περιορισμός Δανεισμού	Ημέρες Δανεισμού	Κράτηση	Ειδικοί Περιορισμοί
012	001	Όροφος Β – Ράφι 5	Δανείζεται	5	-	-
013	001	Όροφος Β – Ράφι 5	Δεν δανείζεται	-	-	-
014	002	Όροφος Α – Ράφι 2	Δανείζεται	5	-	Διαθέτει CD
015	003	Όροφος Α – Ράφι 2	Προς συντήρηση	5	-	-
....						

Στη συνέχεια ο πίνακας των χρηστών περιέχει τις πληροφορίες σχετικά με τους εγγεγραμμένους χρήστες στη βιβλιοθήκη. Ο δανεισμός των αντιτύπων πραγματοποιείται με βάση τα στοιχεία του χρήστη, τα οποία είναι αποθηκευμένα σ' αυτόν τον πίνακα. Ως πεδία του συγκεκριμένου πίνακα, ορίζονται τα εξής:

- Κωδικός (πεδίο κλειδί)
- Όνομα
- Επώνυμο
- Διεύθυνση
- Τηλέφωνο
- e-mail
- Ημερομηνία Λήξης Συνδρομής

Πίνακας 3.11. Πίνακας Χρηστών

Κωδικός	Όνομα	Επώνυμο	Διεύθυνση	Τηλέφωνο	e-mail	Ημερομηνία
---------	-------	---------	-----------	----------	--------	------------

						Λήξης Συνδρομής
0100	Δημήτριος	Στεργίου	Γ. Αιτωλού 14	344553		15/1/2012
0101	Ιωάννης	Παπαδόπουλος	Αχαΐας 14	344566		1/1/2011
....						

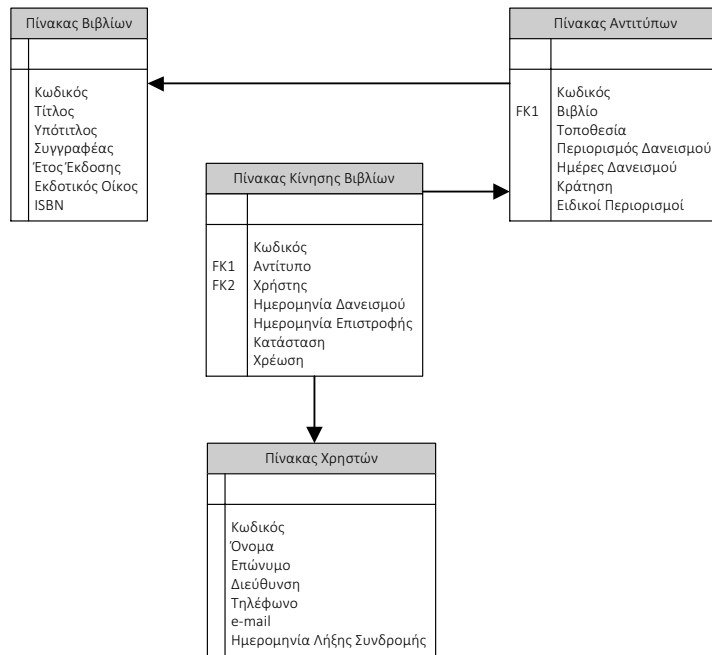
Τέλος, ο πίνακας κίνησης βιβλίων καταγράφει τους δανεισμούς των βιβλίων από τους χρήστες και δίνει πληροφορίες σχετικά με την κατάσταση δανεισμού του κάθε βιβλίου. Ως πεδία του συγκεκριμένου πίνακα, ορίζονται τα εξής:

- Κωδικός (πεδίο κλειδί)
- Αντίτυπο (πεδίο διασύνδεσης με τον πίνακα των αντιτύπων, ένα αντίτυπο μπορεί να δανειστεί πολλές φορές από τον ίδιο ή διαφορετικούς χρήστες $M \times N$)
- Χρήστης (πεδίο διασύνδεσης με τον πίνακα των χρηστών, ένας χρήστης μπορεί να δανειστεί ένα ή περισσότερα αντίτυπα βιβλίων $M \times N$)
- Ημερομηνία Δανεισμού
- Ημερομηνία Επιστροφής
- Κατάσταση
- Χρέωση

Πίνακας 3.12. Πίνακας Κίνησης Βιβλίων

Κωδικός	Αντίτυπο	Χρήστης	Ημερομηνία Δανεισμού	Ημερομηνία Επιστροφής	Κατάσταση	Χρέωση
1045	012	0101	10/7/2010	16/7/1010	Επεστράφη	0 ευρώ
1044	012	0100	10/8/2010	17/8/1010	Σε καθυστέρηση	5 ευρώ
1045	014	0101	20/8/2010	30/8/2010	Κανονική	0 ευρώ
....						

Το σχήμα της βάσης που προκύπτει απεικονίζεται στην παρακάτω εικόνα.



Εικόνα 3.16. Σχεσιακό Μοντέλο Βάσης Δεδομένων βιβλιοθήκης

Ένας χρήστης μπορεί να θέσει μέσω του Σ.Δ.Β.Δ., σύνθετες ερωτήσεις (SQL queries) στη βάση, όπως ποιοι είναι οι τίτλοι των βιβλίων και τα ονοματεπώνυμα των χρηστών που έχουν δανειστεί βιβλία μετά την τάδε ημερομηνία και δεν τα έχουν επιστρέψει. Για να δοθούν αυτά τα στοιχεία πρέπει να συνδεθούν και οι τέσσερις πίνακες. Στην αρχή πρέπει να αναζητηθούν εκείνες οι κινήσεις με πεδίο «Ημερομηνία Δανεισμού» τη δοθείσα και πεδίο «Κατάσταση» να μην έχει την ένδειξη «Επεστράφη». Με βάση αυτές τις εγγραφές θα αναζητηθούν στη συνέχεια, μέσω του πεδίου σύνδεσης «Αντίτυπο» ο τίτλος του βιβλίου, και μέσω του πεδίου σύνδεσης «Χρήστης» το όνομα και το επώνυμο του χρήστη.