



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Εισαγωγή στις Βάσεις Δεδομένων II

Ενότητα: 1. Εισαγωγή στις Βάσεις Δεδομένων I - Δεδομένα

Διδάσκων: Πηγουνάκης Κωστής
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΣΧΟΛΗ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται στην άδεια χρήσης **Creative Commons** και ειδικότερα ***Αναφορά – Μη εμπορική Χρήση – Όχι Παράγωγο Έργο 3.0 Ελλάδα*** (***Attribution – Non Commercial – Non-derivatives 3.0 Greece***)



[ή επιλογή ενός άλλου από τους έξι συνδυασμούς]

[και αντικατάσταση λογότυπου άδειας όπου αυτό έχει μπει (σελ. 1, σελ. 2 και τελευταία)]

- Εξαιρείται από την ως άνω άδεια υλικό που περιλαμβάνεται στις διαφάνειες του μαθήματος, και υπόκειται σε άλλου τύπου άδεια χρήσης. Η άδεια χρήσης στην οποία υπόκειται το υλικό αυτό αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης





Εισαγωγή στις Βάσεις Δεδομένων

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΕΒΔΟ100)

Κωστής Πηγουνάκης

Τελευταία ενημέρωση : 12/2/2013

Η Πληροφορία ως έννοια

Ορισμός εργασίας : η επικοινωνία είναι η ανταλλαγή πληροφορίας. Η διαδικασία της επικοινωνίας προϋποθέτει την αποστολή και λήψη ενός μηνύματος μεταξύ ενός πομπού και ενός δέκτη, όπου και οι δύο χρησιμοποιούν ένα **κοινό κώδικα**.

Οι πληροφορίες μπορεί να είναι οποιασδήποτε μορφής και **μεταφέρουν κάποιο περιεχόμενο**.

Οι πληροφορίες μπορεί να παραμένουν **στατικές** δηλ. σταθερές στην πάροδο του χρόνου ή να είναι **δυναμικές**, δηλ. να αλλάζουν με την πάροδο του χρόνου.

Η πληροφορία ορίζεται ως το **στοιχείο γνώσης** που μετουσιώνει και δίνει αξία στα πράγματα/γεγονότα, τα οποία αποκτούν σημασία με αυτόν τον τρόπο. Πρέπει πάντως να γίνει ο διαχωρισμός μεταξύ πληροφορίας και γνώσης: η πληροφορία είναι το μέγεθος που, καθώς αυξάνει και επεξεργάζεται, προσδίδει σημασία στα πράγματα και παράγει γνώση.

Πληροφορίες και δεδομένα

Τα δεδομένα δεν είναι πληροφορία, όταν δεν έχουν λάβει από τη νόηση συγκεκριμένη σημασία, ενώ η πληροφορία είναι δεδομένα με σημασία, με «ουσιαστικό» περιεχόμενο.

Τα δεδομένα είναι **ποιοτικά** ή **ποσοτικά** χαρακτηριστικά μιας μεταβλητής ή ενός συνόλου μεταβλητών.

Σε αντίθεση με τα δεδομένα η **πληροφορία** έχει ορισμένο νόημα και είναι οργανωμένη για συγκεκριμένο σκοπό. Με την προσθήκη αξιών στα δεδομένα με διαμόρφωση, οργάνωση, μαθηματική ή στατιστική ανάλυση, διόρθωση λαθών ή συμπύεση παράγεται η πληροφορία. Τη διαδικασία αυτή ονομάζουμε **επεξεργασία δεδομένων**.

Δεδομένα στην Πληροφορική

Στην Πληροφορική συχνά δεν γίνεται διάκριση ανάμεσα στα δεδομένα και την πληροφορία.

Δεδομένα : τρόποι αναπαράστασης εννοιών και γεγονότων που δύνανται να υποστούν **διαχείριση** και **επεξεργασία** από τον άνθρωπο, ή από ένα αυτοματοποιημένο υπολογιστικό σύστημα.

Αυτοί οι τρόποι αναπαράστασης σχετίζονται με τη **φύση της πληροφορίας** που μεταφέρουν.

Μοντελοποίηση δεδομένων : η επιλογή του κατάλληλου είδους αναπαράστασης ώστε να μπορεί να αποθηκευτεί η επιθυμητή πληροφορία χωρίς να υπάρξουν απώλειες και χωρίς να γίνει σπατάλη πόρων

Αδόμητα Δεδομένα

Αδόμητα ή ακατέργαστα δεδομένα (raw data) : Συνήθως πρόκειται για δεδομένα που συλλέγονται από κάποια πηγή χωρίς επεξεργασία και αναφέρονται και ως **πρωτεύοντα δεδομένα (primary data)**.

Τα αδόμητα δεδομένα μπορεί να προέρχονται, π.χ., από χειρονακτική εισαγωγή στοιχείων σε ένα πρόγραμμα.

Έτσι, τα αδόμητα δεδομένα εν γένει περιέχουν σφάλματα και διαφορετική ή καθόλου **μορφοποίηση (format)**: μια εισαγωγή δεδομένων ημερομηνίας μπορεί να περιέχει τις παρακάτω μορφές: "**31st January 1999**", "**31/01/1999**", "**31/1/99**", "**31 Jan**".

Άπαξ και αντληθούν τα αδόμητα δεδομένα, πρέπει να υποστούν **επεξεργασία** για να αποκτήσουν ενιαία μορφή, απαλλαγμένη από σφάλματα, ώστε να παρέχουν πραγματική πληροφορία.

Τύποι Δεδομένων

Στον προγραμματισμό τα δεδομένα χαρακτηρίζονται από τον **τύπο τους (data type)**. Ο τύπος δεδομένων είναι μια ταξινόμηση που καθορίζει :

- τις πιθανές **τιμές** που μπορεί να έχουν τα δεδομένα μας (π.χ. ακέραιος αριθμός ή πραγματικός, συμβολοσειρά κλπ)
- τις **πράξεις** που μπορούν να γίνουν με τις τιμές αυτού του τύπου
- το **νόημα** των δεδομένων, και
- τον **τρόπο αποθήκευσης** των δεδομένων.

Ο τύπος των δεδομένων είναι μια τυπική περιγραφή των τιμών που μπορεί να αποθηκευτούν σε μια σταθερά ή μεταβλητή καθώς και των πράξεων που μπορεί να εκτελεστούν με τις τιμές αυτές.

Συστήματα Τύπων Δεδομένων

Οι τύποι δεδομένων χρησιμοποιούνται εντός **συστημάτων τύπων**, τα οποία προσφέρουν διαφορετικούς τρόπους για τον ορισμό, την ανάπτυξη και τη χρήση τους και διαφοροποιούνται ως προς τους βαθμούς ασφάλειας των τύπων.

Ο κάθε τύπος δεδομένων αναπαριστά και έναν **περιορισμό** για το σύστημα τύπων δεδομένων, περιγράφοντας μια αναπαράσταση που μεταφράζεται σε μια δομή τιμών ή αντικειμένων τα οποία αποθηκεύονται στη μνήμη του υπολογιστή.

Το σύστημα τύπων δεδομένων χρησιμοποιεί την πληροφορία των τύπων για να **ελέγξει την ορθότητα** των προγραμμάτων που προσπελούν ή χειρίζονται τα δεδομένα.

Τύποι Μηχανής

Όλα τα δεδομένα στους υπολογιστές που βασίζονται στην ψηφιακή ηλεκτρονική αναπαρίστανται στο κατώτερο επίπεδο ως **bits**, δηλ. εναλλακτικές καταστάσεις (0 ή 1).

Η μικρότερη μονάδα μνήμης δεδομένων είναι συνήθως μια ομάδα από bits που ονομάζονται **byte** και συνήθως πρόκειται για μια οκτάδα από bits (1 byte = 8 bits).

Η μονάδα που μπορεί να επεξεργαστεί ένας υπολογιστής σε επίπεδο κώδικα μηχανής λέγεται **λέξη (word)** και σήμερα (2013) είναι μήκους 32 ή 64bits.

Λογικός Τύπος (Boolean)

Τα δεδομένα αληθείας αποτυπώνονται δύο καταστάσεις :

- **ΑΛΗΘΕΣ** - **TRUE** και
- **ΨΕΥΔΕΣ** - **FLASE**.

Τα δεδομένα αυτά αποτελούν το **λογικό τύπο ή τύπο Boole (Boolean type)**.

Παρότι μόνο δύο τιμές είναι πιθανές, σπάνια αναπαρίστανται στους υπολογιστές με ένα μόνο δυαδικό ψηφίο για λόγους αποδοτικότητας.

Πολλές γλώσσες προγραμματισμού δεν είναι εφοδιασμένες με ρητό τύπο αληθείας. Αντί αυτού ερμηνεύουν το 0 (για παράδειγμα) ως ψευδές και τις υπόλοιπες τιμές ως αληθείς.

Αριθμητικοί τύποι

ΟΝΟΜΑΣΙΑ ΤΥΠΟΥ	ΘΕΣΕΙΣ ΑΠΟΘΗΚΕΥΣΗΣ	ΕΥΡΟΣ ΤΙΜΩΝ
TINYINT	1 byte	-128 ΕΩΣ 127 [0 ΕΩΣ 255 ΧΩΡΙΣ ΠΡΟΣΗΜΟ - UNSIGNED]
SMALLINT	2 bytes	-32,768 ΕΩΣ 32,767 [0 ΕΩΣ 65,535]
MEDIUMINT	3 bytes	-8,388,608 ΕΩΣ 8,388,607 [0 ΕΩΣ 16,777,215]
INT	4 bytes	-/+2.147E+9 [0 ΕΩΣ 4.294E+9]
BIGINT	8 bytes	-/+9.223E+18 [0 ΕΩΣ 18.45E+18]
FLOAT	4 bytes	Min=+/-1.15E-38 Max=+/-3.403E+38
DOUBLE	8 bytes	Min=+/-2.225E-308 Max=+/-1.798E+308
DECIMAL (M , [D]) <i>αποθηκεύονται ως συμβολοσειρά</i>	M+2	Max εύρος = το εύρος DOUBLE

Τύποι Κειμένου/Συμβολοσειρών

CHAR (size)	Αποθηκεύει ένα σταθερό μήκος συμβολοσειράς Το μέγεθος (μήκος) ορίζεται σε παρένθεση. Μπορεί να αποθηκεύσει μέχρι 255 χαρακτήρες.
VARCAR (size)	Αποθηκεύει ένα μεταβλητό μήκος συμβολοσειράς Το μέγιστο μέγεθος (μήκος) ορίζεται σε παρένθεση. Μπορεί να αποθηκεύσει μέχρι 255 χαρακτήρες. Σημείωση: Αν το μήκος της συμβολοσειράς είναι μεγαλύτερο από 255 χαρακτήρες, μετατρέπεται σε τύπο TEXT
TINYTEXT	Αποθηκεύει το πολύ 255 χαρακτήρες
TEXT	Αποθηκεύει το πολύ 65,535 χαρακτήρες
BLOB	Binary Large Objects. Αποθηκεύει το πολύ 65,535 bytes δεδομένων
MEDIUMTEXT	Αποθηκεύει το πολύ 16,777,215 χαρακτήρες
MEDIUMBLOB	Binary Large Objects. Αποθηκεύει το πολύ 16,777,215 bytes δεδομένων
LONGTEXT	Αποθηκεύει το πολύ 4,294,967,295 χαρακτήρες
LOBLOB	Binary Large Objects. Αποθηκεύει το πολύ 4,294,967,295 bytes δεδομένων
ENUM (x, y, z, etc.)	Επιτρέπει την εισαγωγή λίστας πιθανών τιμών (μέχρι 65535 τιμές). Αν η τιμή που εισάγεται δεν είναι μεταξύ των πιθανών τιμών, αποθηκεύεται η κενή τιμή.
SET	Παρόμοιος τύπος με το ENUM εκτός του ότι το SET μπορεί να περιέχει μέχρι 64 τιμές και μπορεί να αποθηκεύσει περισσότερες από μία

Τύποι Χρόνου

DATE ()	Ημερομηνία με μορφή οργάνωσης: EEEE-MM-ΗΗ. Υποστηρίζονται ημερομηνίες από '1000-01-01' έως '9999-12-31'
DATETIME ()	Ημερομηνία και ώρα με μορφή οργάνωσης: EEEE-MM-ΗΗ ΩΩ:ΛΛ:ΔΔ. Υποστηρίζονται ημερομηνίες από '1000-01-01 00:00:00' έως '9999-12-31 23:59:59'
TIMESTAMP ()	Τρόπος αποτύπωσης της τρέχουσας ώρας που προέρχεται από το λειτουργικό σύστημα UNIX με μορφή οργάνωσης: EEEE-MM-ΗΗ ΩΩ:ΛΛ:ΔΔ. Υποστηρίζονται χρόνοι από '1970-01-01 00:00:01' έως '2038-01-09 03:14:07'
TIME ()	Χρόνος με μορφή οργάνωσης: ΩΩ:ΛΛ:ΔΔ. Υποστηρίζονται χρόνοι από '-838:59:59' έως '838:59:59'
YEAR ()	Έτος με 2 ή 4 ψηφία. Για 4 ψηφία υποστηρίζονται τα έτη : 1901 έως 2155 ενώ για τα 2 ψηφία τα έτη από 70 έως 69, αναπαριστώντας τα έτη από 1970 έως 2069.

Τύπος Απαρίθμησης

Ο τύπος της **απαρίθμησης (enumeration)** αφορά σε μεταβλητές που παίρνουν δεδομένες τιμές που διαφέρουν μεταξύ τους, οι οποίες μπορούν να συγκριθούν, αλλά δεν έχουν απαραίτητα κάποια συγκεκριμένη αναπαράσταση στην μνήμη του υπολογιστή.

Οι μεταγλωττιστές και οι διερμηνευτές τις αναπαριστούν αυθαίρετα. Στη MySQL αποτελούν μέρος του τύπου κειμένου.

Για παράδειγμα, τα τέσσερα χαρτιά σε μία τράπουλα θα μπορούσαν να είναι τέσσερις απαριθμητές με ονόματα ΣΠΑΘΙ, ΚΑΡΩ, ΜΠΑΣΤΟΥΝΙ, ΚΟΥΠΑ, που θα ανήκουν σε μία απαρίθμηση με όνομα τράπουλα. Αν η μεταβλητή V δηλωθεί να έχει τράπουλα για τύπο, μπορούμε να αναθέσουμε οποιαδήποτε από αυτές τις τιμές σε αυτή την μεταβλητή.

Σύνθετοι τύποι

Η ομαδοποίηση απλών τύπων σε μια πιο σύνθετη δομή αποθήκευσης δημιουργεί ένα **σύνθετο τύπο δεδομένων**.

ΟΝΟΜΑ
ΕΠΩΝΥΜΟ
ΗΛΙΚΙΑ
ΚΩΔΙΚΟΣ

EMPLOYEE
ΟΝΟΜΑ
ΕΠΩΝΥΜΟ
ΗΛΙΚΙΑ
ΚΩΔΙΚΟΣ

Στην τεχνολογία των βάσεων δεδομένων, οι δομές αυτές ονομάζονται **εγγραφές (records)**.

Δομές Δεδομένων

Δομή δεδομένων (data structure): διαφορετικοί τρόποι οργάνωσης και αποθήκευσης δεδομένων, ώστε αυτά να μπορούν να χρησιμοποιηθούν αποδοτικά στον προγραμματισμό.

Σε συγκεκριμένες εφαρμογές χρησιμοποιούνται συγκεκριμένες δομές δεδομένων, δηλαδή αυτές που είναι οι πιο αποδοτικές αλγοριθμικά για το κάθε είδος εφαρμογής.

Χρησιμοποιούνται σχεδόν σε κάθε πρόγραμμα ή σύστημα λογισμικού. Παρέχουν έναν τρόπο αποδοτικής διαχείρισης τεράστιου όγκου δεδομένων, όπως μεγάλες βάσεις δεδομένων και υπηρεσίες ευρετηρίου στο διαδίκτυο.

Οι δομές δεδομένων διακρίνονται επίσης σε **στατικές** και σε **δυναμικές**:

- Στις στατικές δομές το πλήθος των στοιχείων είναι σταθερό και καθορίζεται στον ορισμό του τύπου στο τμήμα δηλώσεων πχ ο πίνακας.
- Στις δυναμικές δομές το πλήθος των στοιχείων της δομής καθορίζεται κατά τη διάρκεια εκτέλεσης του προγράμματος.

Οι Βάσεις Δεδομένων

- Βάση δεδομένων : ένα σύνολο δεδομένων που χαρακτηρίζονται από κάποια λογική οργάνωση και ομαδοποίηση έτσι ώστε να είναι εύκολη και αποτελεσματική η διαχείρισή τους.
- Οι βάσεις δεδομένων κατέχουν κεντρική θέση στην επιστήμη της Πληροφορικής, διότι επιτυγχάνουν τη διαχείριση της πληροφορίας.
- Τα τελευταία χρόνια παρατηρείται μια ραγδαία ανάπτυξη της τεχνολογίας των βάσεων δεδομένων τόσο σε θεωρητικό επίπεδο όσο και σε επίπεδο εφαρμογών:
 - Επιχειρήσεις - Τραπεζικός τομέας
 - Δημόσια Διοίκηση
 - Εκπαίδευση
 - Δίκτυα Μεταφορών
 - Δίκτυα Επικοινωνιών κλπ

Συστήματα διαχείρισης ΒΔ

- Η διαχείριση μιας βάσης δεδομένων πραγματοποιείται με τη βοήθεια ειδικών προγραμμάτων, που ονομάζονται **συστήματα διαχείρισης βάσεων δεδομένων (ΣΔΒΔ) - Database Management Systems (DBMS)**.
- Ένα ΣΔΒΔ είναι **λογισμικό** που χρησιμοποιείται για:
 - τη δημιουργία,
 - τη συντήρηση,
 - την επεξεργασία
 - τη συνολική διαχείριση μιας ΒΔ.

Παραδείγματα ΣΔΒΔ

- Microsoft Access,
- MS SQL Server
- Oracle,
- Sybase
- MySQL
- Postgress,

Microsoft Access : σε προσωπικούς υπολογιστές και σε επιχειρήσεις μικρού μεσαίου /μεγέθους.

MS SQL Server, Oracle, Sybase : σε μεγάλες επιχειρήσεις, υποστηρίζουν πολλούς χρήστες, πολλές ταυτόχρονες προσπελάσεις πάνω στα ίδια δεδομένα, καθώς και απομακρυσμένη πρόσβαση μέσω δικτύου.

MySQL και Postgress : ανοικτού λογισμικού, κυρίως σε εφαρμογές διαδικτύου



PostgreSQL



ORACLE

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

