Σχεδίαση Β.Δ. (Database Design)

- Η σχεδίαση ενός σχήματος μιας Β.Δ. βασίζεται σε μεγάλο βαθμό στη διαίσθηση του σχεδιαστή σχετικά με τον κόσμο που θέλει να αναπαραστήσει.
- Η εννοιολογική σχεδίαση υπαρκτών κόσμων παράγει σχήματα υψηλού επιπέδου, π.χ. διαγράμματα Οντοτήτων – Σχέσεων.
- Ένα εννοιολογικό μοντέλο πρέπει να μετατραπεί σε ένα λογικό μοντέλο, π.χ. σχεσιακό.
- Προκειμένου να κρίνουμε αν η σχεδίαση είναι ορθή πρέπει να χρησιμοποιήσουμε κάποια τυπικά κριτήρια.

Θεωρία Σχεδίασης Σχεσιακών Β.Δ.

- Η σχεδίαση του σχήματος μιας σχεσιακής Β.Δ. μπορεί να τυποποιηθεί με χρήση της θεωρίας κανονικοποίησης (normalization theory)
- Με τον όρο «κανονικοποίηση» εννοούμε την εφαρμογή κανόνων σχεδίασης που αποκαλούνται κανονικές μορφές (normal forms) και οι οποίοι περιορίζουν τις δυνατές μορφές σχεσιακών σχημάτων.
- Αν ακολουθούνται οι κανόνες αυτοί αποφεύγεται ανώμαλη ή λανθασμένη συμπεριφορά του συστήματος.
- Επιβεβαιώνεται επίσης ότι το σχήμα έχει διάφορες επιθυμητές ιδιότητες.

Κανονικές Μορφές

- 1η Κανονική Μορφή (1NF)
  - Μια σχέση είναι σε 1η κανονική μορφή αν δεν έχει πλειότιμα γνωρίσματα
- 2η Κανονική Μορφή
- 3η Κανονική Μορφή
- Κανονική Μορφή Boyce-Codd

- Υπόθεση: όλες οι σχέσεις είναι σε 1NF

Κανονικές Μορφές

- Παράδειγμα: Employee Database
  - για κάθε υπάλληλο αποθηκεύεται η ακόλουθη πληροφορία: **emp_id, emp_name, emp_phone, dept_name, dept_phone, dept_mgrname, skill_id, skill_name, skill_date, skill_lvl**
  - Τα γνωρίσματα **emp_id, dept_name** και **skill_id** προσδιορίζουν μοναδικά υπαλλήλους, τμήματα και δεξιότητες αντίστοιχα.
  - Μια καθολική σχέση (universal relation) είναι μια σχέση η οποία περιέχει όλα τα γνωρίσματα που αντιστοιχούν σε πληροφορία σχετικά με τους υπαλλήλους.

Κανονικές Μορφές

- Προβλήματα με τη χρήση καθολικής σχέσης:
  - Επανάληψη πληροφορίας: για έναν υπάλληλο με πολλές δεξιότητες εμφανίζονται ισάριθμες πλειάδες στη σχέση
  - Αν ένας υπάλληλος αλλάξει τμήμα ή αριθμό τηλεφώνου, όλες οι πλειάδες του υπαλλήλου στη σχέση πρέπει να αλλαχθούν.
  - Αντίστοιχα, αν ένα τμήμα αποκτήσει νέο διευθυντή.
  - Η επαναλαμβανόμενη πληροφορία δεν είναι μόνο περιττή, αλλά πρέπει επίσης να κρατείται ενημερωμένη
  - Αν ένας υπάλληλος ο οποίος έχει μια και μόνο δεξιότητα, την χάσει, τότε όλη η πληροφορία σχετικά με τον υπάλληλο πρέπει να χαθεί επίσης.

Προβλήματα Κακής Σχεδίασης

- Πρόβλημα ενημέρωσης:
  - Μια σχέση R πάσχει από πρόβλημα ενημέρωσης αν, όποτε αλλάζει η τιμή ενός γνωρίσματος για ένα στιγμιότυπο της οντότητας ή της σχέσης την οποία αναπαριστά η R, είναι απαραίτητη η ενημέρωση πολλαπλών πλειάδων της R.
- Πρόβλημα διαγραφής:
  - Μια σχέση R πάσχει από πρόβλημα διαγραφής αν η διαγραφή μιας πλειάδας της σχέσης έχει ως αποτέλεσμα την απώλεια πληροφορίας σχετικά με μια σχετιζόμενη οντότητα ή σχέση.
- Πρόβλημα εισαγωγής:
  - Μια σχέση R πάσχει από πρόβλημα εισαγωγής αν πληροφορία δεν μπορεί να αναπαρασταθεί παρά μόνο αν περιληφθεί πληροφορία σχετική με κάποια άλλη οντότητα ή σχέση (η οποία μπορεί να μην υπάρχει).

Προβλήματα Κακής Σχεδίασης

- Στο προηγούμενο παράδειγμα, κάποια από τα προβλήματα αυτά μπορούν να λυθούν αν η καθολική σχέση χωριστεί σε δύο πίνακες:

```
Employees (emp_id, emp_phone, dept_name,
           dept_phone, dept_mgrname)
```

```
Skills (emp_id, skill_id, skill_name,
        skill_date, skill_lvl)
```

- Ο πίνακας **`Employees`** περιέχει μια μοναδική πλειάδα για κάθε υπάλληλο.
- Ο πίνακας **`Skills`** περιέχει μια μοναδική πλειάδα για κάθε ζεύγος **`emp_id, skill_id`**.
- Η συνένωση των δύο σχέσεων δίνει την αρχική σχέση.
- Άρα έχουμε την ίδια πληροφορία χωρίς περιττή επανάληψη.

Συναρτησιακές Εξαρτήσεις

- Μια συναρτησιακή εξάρτηση (Functional Dependency) είναι ένας περιορισμός μεταξύ δύο συνόλων γνωρισμάτων.
  - Έστω $A_1, A_2, \ldots, A_n$ όλα τα γνωρίσματα μιας σχέσης R. Αν X και Y είναι υποσύνολα του $\{A_1, A_2, \ldots, A_n\}$, τότε η συναρτησιακή εξάρτηση $X \rightarrow Y$ καθορίζει ότι για οποιεσδήποτε δύο πλειάδες t_1, t_2 της R, αν $t_1[X] = t_2[X]$, τότε πρέπει επίσης να ισχύει ότι $t_1[Y] = t_2[Y]$.

- Δεδομένης μιας συναρτησιακής εξάρτησης $X \rightarrow Y$, λέμε ότι το σύνολο X προσδιορίζει συναρτησιακά το σύνολο Y ή ότι το σύνολο Y εξαρτάται συναρτησιακά από το σύνολο X

Συναρτησιακές Εξαρτήσεις

- Αν η συναρτησιακή εξάρτηση $X \rightarrow Y$ ισχύει σε μια σχέση R, τότε δεν είναι δυνατόν να υπάρχουν πλειάδες οι οποίες συμφωνούν στις τιμές όλων των γνωρισμάτων στο X και συγχρόνως δε συμφωνούν στην τιμή κάποιου από τα γνωρίσματα του Y.

- Αν το X είναι κλειδί της R, τότε η εξάρτηση $X \rightarrow Y$ ισχύει για κάθε υποσύνολο Y των γνωρισμάτων της R.

- Η εξάρτηση $X \rightarrow Y$ δεν συνεπάγεται την εξάρτηση $Y \rightarrow X$

Συναρτησιακές Εξαρτήσεις

- Παράδειγμα: Έστω ότι οι παρακάτω σχέσεις έχουν το περιεχόμενο που φαίνεται στους αντίστοιχους πίνακες. Προσδιορίστε τις συναρτησιακές εξαρτήσεις οι οποίες ισχύουν.

T1

A	B
x_1	y_1
x_2	y_2
x_3	y_1
x_4	y_1
x_5	y_2
x_6	y_2

T2

A	B
x_1	y_1
x_2	y_4
x_1	y_1
x_3	y_2
x_2	y_4
x_4	y_3

T3

A	B
x_1	y_1
x_2	y_4
x_1	y_1
x_3	y_2
x_2	y_4
x_4	y_4

T1: $A \rightarrow B$

$B \nrightarrow A$

T2: $A \rightarrow B$

$B \rightarrow A$

T3: $A \rightarrow B$

$B \nrightarrow A$

Συναρτησιακές Εξαρτήσεις

- Παράδειγμα: Έστω το σχήμα

$$EMPLOYEE\ (emp\_id, emp\_name, emp\_phone, dept\_name)$$
$$DEPARTMENT\ (dept\_id, dept\_name, dept\_phone, dept\_mgrname)$$
$$SKILL\ (skill\_id, skill\_name)$$
$$EMP\_HAS\_SKILL\ (emp\_id, skill\_id, skill\_date, skill\_level)$$

- Οι συναρτησιακές εξαρτήσεις που ισχύουν είναι:

1. $emp\_id \rightarrow emp\_name, emp\_phone, dept\_name$
2. $dept\_name \rightarrow dept\_phone, dept\_mgrname$
3. $skill\_id \rightarrow skill\_name$
4. $emp\_id, skill\_id \rightarrow skill\_date, skill\_level$

Λογικές Συνέπειες Συναρτησιακών Εξαρτήσεων

- **Κανόνας Εγκλεισμού** (Inclusion Rule): Αν X, Y είναι σύνολα γνωρισμάτων από το σχήμα της σχέσης R και $Y \subseteq X$, τότε $X \rightarrow Y$.

- Μια συναρτησιακή εξάρτηση $X \rightarrow Y$ λέγεται **τετριμμένη** αν ισχύει για κάθε σχέση R της οποίας το σχήμα περιέχει X και Y

- Τετριμμένες εξαρτήσεις εμφανίζονται σαν αποτέλεσμα της εφαρμογής του κανόνα εγκλεισμού

- **Θεώρημα:** Αν $X \rightarrow Y$ είναι τετριμμένη συναρτησιακή εξάρτηση, πρέπει να ισχύει ότι $Y \subseteq X$

Λογικές Συνέπειες Συναρτησιακών Εξαρτήσεων

- **Θεώρημα:** Αν $X \rightarrow Y$ είναι τετριμμένη συναρτησιακή εξάρτηση, πρέπει να ισχύει ότι $Y \subseteq X$
- **Απόδειξη:** Υποθέστε ότι $Y \supset X$. Δημιουργήστε μια σχέση με όλα τα γνωρίσματα των X και Y και θεωρήστε ένα γνώρισμα A του $Y - X$. Εφόσον $A \in Y$ και $A \notin X$, είναι δυνατόν να κατασκευάσουμε δύο πλειάδες u και v, οι οποίες έχουν κοινές τιμές σε όλα τα γνωρίσματα στο X αλλά έχουν διαφορετικές τιμές στο A. Τότε όμως η τετριμμένη εξάρτηση δεν ισχύει. Άρα, δε μπορεί να υπάρχει τέτοιο γνώρισμα A στο $Y - X$. Επομένως, $Y \subseteq X$.

- Από ένα μικρό αριθμό κανόνων συνεπαγωγής και ένα αρχικό σύνολο συναρτησιακών εξαρτήσεων μπορεί να εξαχθεί ένας αριθμός πρόσθετων συναρτησιακών εξαρτήσεων.

Αξιώματα Armstrong

- Έστω ότι τα σύνολα γνωρισμάτων X, Y, Z περιέχονται στο σχήμα της σχέσης R. Τότε ισχύουν οι παρακάτω κανόνες:

1. **Κανόνας Εγκλεισμού (Inclusion Rule)**

   Αν $Y \subseteq X$ τότε $X \rightarrow Y$

2. **Κανόνας Μεταβατικότητας (Transitivity Rule)**

   Αν $X \rightarrow Y$ και $Y \rightarrow Z$ τότε $X \rightarrow Z$

3. **Κανόνας Επαύξησης (Augmentation Rule)**

   Αν $X \rightarrow Y$ τότε $XZ \rightarrow YZ$

Συνέπειες των Αξιωμάτων

- Θεώρημα: Αν W, X, Y, Z, B περιέχονται στο σχήμα της R, τότε:

1. **Κανόνας Ένωσης (Union Rule)**
   Αν $X \to Y$ και $X \to Z$ τότε $X \to YZ$
2. **Κανόνας Αποσύνθεσης (Decomposition Rule)**
   Αν $X \to YZ$ τότε $X \to Y$ και $X \to Z$
3. **Κανόνας Ψευδομεταβατικότητας (Pseudotransitivity Rule)**
   Αν $X \to Y$ και $WY \to Z$ τότε $XW \to Z$
4. **Κανόνας Συσσώρευσης (Accumulation Rule)**
   Αν $X \to YZ$ και $Z \to B$ τότε $X \to YZB$

Αξιώματα Armstrong

- **Παράδειγμα:** Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

1. Σ.Ε. με ένα γνώρισμα στο αριστερό μέλος
   - Οι τετριμμένες εξαρτήσεις $A \rightarrow A, B \rightarrow B, C \rightarrow C, D \rightarrow D$ δεν περιλαμβάνονται στο ελάχιστο σύνολο.
   - Οι εξαρτήσεις $A \rightarrow B, C \rightarrow B, D \rightarrow B$ προκύπτουν από τον πίνακα καθώς όλες οι τιμές του B είναι ίδιες.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

1. Σ.Ε. με ένα γνώρισμα στο αριστερό μέλος
   - Τα γνωρίσματα A, C, D έχουν τουλάχιστον δύο διακεκριμένες τιμές. Άρα, $B \nrightarrow A, B \nrightarrow C, B \nrightarrow D$.
   - Όλες οι τιμές του D είναι διαφορετικές Άρα, $D \rightarrow A, D \rightarrow B, D \rightarrow C$.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

1. Σ.Ε. με ένα γνώρισμα στο αριστερό μέλος
   - Τα γνωρίσματα A, και C έχουν τουλάχιστον δύο επαναλαμβανόμενες τιμές. Άρα, $A \nrightarrow D$, C $\nrightarrow D$.
   - $A \nrightarrow C$ και $C \nrightarrow A$, εξαιτίας των πλειάδων 1,2 και 1,3 αντίστοιχα.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

1. Σ.Ε. με ένα γνώρισμα στο αριστερό μέλος
   - Άρα ισχύουν οι ακόλουθες Σ.Ε.:
     $A \rightarrow B, C \rightarrow B, D \rightarrow A, D \rightarrow B, D \rightarrow C$
   - Από τον κανόνα ένωσης: $A \rightarrow B, C \rightarrow B, D \rightarrow ABC$.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

2. Σ.Ε. με ζεύγος γνωρισμάτων στο αριστερό μέλος
   - Εξαιτίας της $D \rightarrow ABC$, κάθε ζεύγος που περιέχει το D προσδιορίζει συναρτησιακά όλα τα υπόλοιπα γνωρίσματα (κανόνας επαύξησης). Αυτές οι εξαρτήσεις είναι συνεπαγωγές εξαρτήσεων που ήδη ανήκουν στο ζητούμενο σύνολο.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

2. Σ.Ε. με ζεύγος γνωρισμάτων στο αριστερό μέλος
    - Ζεύγη γνωρισμάτων που περιλαμβάνουν το B στο αριστερό μέλος δίνουν είτε τετριμμένες εξαρτήσεις, είτε συνεπαγωγές εξαρτήσεων.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

2. Σ.Ε. με ζεύγος γνωρισμάτων στο αριστερό μέλος
   - $AC \rightarrow ABCD$ καθώς το ζεύγος AC έχει διακεκριμένες τιμές σε κάθε πλειάδα. Η μόνη νέα εξάρτηση είναι η $AC \rightarrow D$. Οι εξαρτήσεις $AC \rightarrow A, AC \rightarrow C, AC \rightarrow B$ ήδη εξάγονται από άλλες εξαρτήσεις.

Αξιώματα Armstrong

- Παράδειγμα: Βρείτε ένα ελάχιστο σύνολο συναρτησιακών εξαρτήσεων οι οποίες ικανοποιούνται στον ακόλουθο πίνακα:

T

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_3
a_2	b_1	c_3	d_4

3. Δεν υπάρχουν εξαρτήσεις με 3 ή 4 γνωρίσματα στο αριστερό μέλος οι οποίες ανήκουν σε αυτό το σύνολο.

Το ελάχιστο σύνολο εξαρτήσεων είναι

$$\{A \rightarrow B, C \rightarrow B, D \rightarrow ABC, AC \rightarrow D\}$$

Κλείσιμο (Closure)

- Τα αξιώματα Armstrong και οι συνέπειές τους παράγουν ένα σύνολο το οποίο είναι πολύ μεγαλύτερο από το αρχικό σύνολο των ΣΕ.
- Δεδομένου ενός συνόλου F από ΣΕ, το κλείσιμο (closure) F^+ του F ορίζεται σαν το σύνολο των ΣΕ οι οποίες συνεπάγονται από το F.

- Παράδειγμα: Θεωρείστε το σύνολο των ΣΕ:

$$F = \{A \to B, B \to C, C \to D, D \to E, E \to F, F \to G, G \to H\}$$

✓Από $A \to B$ και $B \to C$, συνεπάγεται λόγω μεταβατικότητας η $A \to C$.

✓Επίσης: $A \to D, A \to E, A \to F, A \to G, A \to BC, A \to EF$ κ.ο.κ.

✓Παρόμοια για εξαρτήσεις με B στο αριστερό μέλος κ.ο.κ.

Κάλυψη (Cover)

- Το μέγεθος του κλεισίματος ενός συνόλου ΣΕ μεγαλώνει εκθετικά με αυτό του αρχικού συνόλου.
- Χρειαζόμαστε έναν τρόπο για να αναφερόμαστε στο σύνολο των εξαρτήσεων που συνεπάγονται από ένα αρχικό σύνολο χωρίς να πρέπει να υπολογίσουμε το κλείσιμο του συνόλου αυτού.
- Για κάθε σύνολο ΣΕ μπορούμε να βρούμε ένα «ισοδύναμο» σύνολο το οποίο είναι ελάχιστο.

- Ένα σύνολο F από ΣΕ για μια σχέση R καλύπτει ένα άλλο σύνολο G από ΣΕ για την R, αν το σύνολο G μπορεί να εξαχθεί με την εφαρμογή των κανόνων συνεπαγωγής στις ΣΕ του F, δηλαδή αν $G \subseteq F^+$.

Κάλυψη (Cover)

- Αν το F καλύπτει το G και το G καλύπτει το F τότε τα F και G λέγονται ισοδύναμα ($F \equiv G$).
- Αν $F \equiv G$, τότε $F^+ = G^+$
- Παράδειγμα: Θεωρείστε τα σύνολα ΣΕ

  $F = \{B \rightarrow CD, AD \rightarrow E, B \rightarrow A\}$

  $G = \{B \rightarrow CDE, B \rightarrow ABC, AD \rightarrow E\}$

  Δείξτε ότι το F καλύπτει το G.

Πρέπει να δείξουμε ότι κάθε ΣΕ του G μπορεί να εξαχθεί από το F με χρήση των κανόνων συνεπαγωγής.

- Η ΣΕ $AD \rightarrow E$ είναι ήδη στο F.
- Από την $B \rightarrow CD$ και την $B \rightarrow A$, εξάγουμε την $B \rightarrow ACD$ (κανόνας ένωσης)

Κάλυψη (Cover)

- Παράδειγμα (συνέχεια): Θεωρείστε τα σύνολα ΣΕ

  $F = \{B \rightarrow CD, AD \rightarrow E, B \rightarrow A\}$

  $G = \{B \rightarrow CDE, B \rightarrow ABC, AD \rightarrow E\}$

  Δείξτε ότι το F καλύπτει το G.

Πρέπει να δείξουμε ότι κάθε ΣΕ του G μπορεί να εξαχθεί από το F με χρήση των κανόνων συνεπαγωγής.

- Από $B \rightarrow ACD$ και $B \rightarrow B$, εξάγουμε την $B \rightarrow ABCD$ (κανόνας ένωσης)
- Από $B \rightarrow ABCD$ εξάγουμε την $B \rightarrow AD$ (κανόνας αποσύνθεσης)
- Από $B \rightarrow AD$ και $AD \rightarrow E$ εξάγουμε την $B \rightarrow E$ (κανόνας μεταβατικότητας)

Κάλυψη (Cover)

- **Παράδειγμα:** Θεωρείστε τα σύνολα ΣΕ

  $F = \{B \to CD, AD \to E, B \to A\}$

  $G = \{B \to CDE, B \to ABC, AD \to E\}$

  Δείξτε ότι το F καλύπτει το G.

Πρέπει να δείξουμε ότι κάθε ΣΕ του G μπορεί να εξαχθεί από το F με χρήση των κανόνων συνεπαγωγής.

- Από $B \to ABCD$ και $B \to E$ εξάγουμε την $B \to ABCDE$ (κανόνας ένωσης)
- Από $B \to ABCDE$ εξάγουμε την $B \to CDE$ και την $B \to ABC$ (κανόνας αποσύνθεσης)