ΗΥ360
Αρχεία και Βάσεις Δεδομένων

Διδάσκων: Δ. Πλεξουσάκης

Συναρτησιακές Εξαρτήσεις
Αξιώματα Armstrong
Ελάχιστη κάλυψη

Φροντιστήριο

Συναρτησιακές Εξαρτήσεις

- Οι Συναρτησιακές εξαρτήσεις είναι περιορισμοί ακεραιότητας πάνω σε σύνολα γνωρισμάτων των σχέσεων ενός σχήματος.
- Σε οποιαδήποτε σχέση R,
  αν $\alpha \subseteq$ Head(R) και $\beta \subseteq$ Head(R) η συναρτησιακή εξάρτηση $\alpha \rightarrow \beta$
  καθορίζει ότι
  για οποιεσδήποτε δύο πλειάδες t1 και t2 της R:

  Αν t1[α] = t2[α], τότε t1[β] = t2[β].

#	DeptNo	SupSSN	City	#
	2	1234	N.York	
	3	5432	Paris	
	2	1234	Barcelona	
	3	5432	N.York	
	2	1234	Chicago	

Συναρτησιακές Εξαρτήσεις

- Έστω ότι ο διευθυντής τμήματος είναι μόνο σε ένα τμήμα. Και κάθε τμήμα έχει έναν διευθυντή. Ένα τμήμα είναι σε πολλές πόλεις.
- Ισχύει η συναρτησιακή εξάρτηση
  SupSSn → DeptNo
  - Το SupSSn καλείται προσδιοριστικό
- Αλλά δεν ισχύει
  SupSSn → City
- Ούτε ισχύει
  DeptNo → City

#	DeptNo	SupSSN	City	#
	2	1234	N.York	
	3	5432	Paris	
	2	1234	Barcelona	
	3	5432	Paris	
	2	1234	Chicago	

Συναρτησιακές Εξαρτήσεις

- Έστω ότι ο διευθυντής τμήματος διευθύνει μόνο ένα τμήμα σε κάποια πόλη.
- Ισχύει η συναρτησιακή εξάρτηση
  {City, DeptNo } → SupSSN
  - Το {City, DeptNo } προσδιορίζει συναρτησιακά το SupSSN.
  - Το SupSSN εξαρτάται συναρτησιακά από το {City, DeptNo } .
- Αλλά δεν ισχύει
  SupSSn → { City, DeptNo}
- Ούτε ισχύει
  {SupSSN, DeptNo } → City

#	DeptNo	SupSSN	City	#
	2	1234	N.York	
	2	1234	Moscow	
	4	1234	Chicago	
	3	5432	N.York	
	2	1234	Barcelona	

Συναρτησιακές Εξαρτήσεις

- Οι συναρτησιακές εξαρτήσεις έχουν να κάνουν με <u>όλ</u>α τα πιθανά έγκυρα στιγμιότυπα μιας βάσης.
  - δεν εξάγεται μια συναρτησιακή εξάρτηση που ικανοποιεί απλά κάποιο ή κάποια από τα πιθανά στιγμιότυπα μιας σχέσης.
- Οι συναρτησιακές εξαρτήσεις απαιτούν η τιμή ορισμένων γνωρισμάτων να προσδιορίζει μοναδικά την τιμή άλλων γνωρισμάτων.
- Οι περιορισμοί του <u>**κύριου κλειδιού**</u> αποτελεί μια ειδική περίπτωση συναρτησιακών εξαρτήσεων.
  - Οι τιμές των γνωρισμάτων του **κλειδιού** <u>**προσδιορίζουν μοναδικά ολόκληρες τις πλειάδες μιας σχέσης**</u>

Κλείσιμο (Closure)

- Σε ένα σχεσιακό σχήμα R, αν διαθέτουμε ένα σύνολο συναρτησιακών εξαρτήσεων F, τότε ίσως μπορούμε να συνεπάγουμε λογικά ότι ισχύουν και κάποιες άλλες.
- Για παράδειγμα:
  - R = (A, B, C, D)
  - Και ισχύει
    - $A \rightarrow B$
    - $A \rightarrow C$
    - $B \rightarrow D$
  - Τότε
    - $A \rightarrow D$
    - $A \rightarrow A$ (τετριμμένηΣ.Ε.)
- Αν έχουμε ένα σύνολο από συναρτησιακές εξαρτήσεις F, τότε το σύνολο F^{+} όλων των συναρτησιακών εξαρτήσεων που συνεπάγονται λογικά καλείται κλείσιμο του F.

Αξιώματα Armstrong

- Σε ένα σχεσιακό σχήμα, αν διαθέτουμε τα σύνολα γνωρισμάτων A, B, C που είναι μη κενά υποσύνολα μιας σχέσης R, τότε :
  - Ανακλαστικότητα :
    - Αν B ⊆ A, τότε A → B
      - Παράδειγμα: XYZ → XY
  - Προσαύξηση:
    - Αν A → B , τότε AC → BC.
      - Παράδειγμα: XY → Z και W, τότε XYW → ZW
  - Μεταβατικότητα ή Επαγωγή:
    - Αν A → B και B → C, τότε A → C
      - Παράδειγμα: XY → Z, Z →WV, τότε XY → WV

Αξιώματα Armstrong

- Ενώ το παραπάνω σύνολο κανόνων είναι έγκυρο και πλήρες, προκύπτουν και οι χρήσιμοι **κανόνες**:
  - Ένωση :
    - Αν A → B και A → C, τότε A → BC
  - Αποσύνθεση:
    - Αν A → BC , τότε A → B και A → C.
  - Ψευδό-μεταβατικότητα:
    - Αν A → B και BC → D, τότε AC → D

Παράδειγματα

- Έστω R = (A,B,C,G,H,I)
- Και οι συναρτησιακές εξαρτήσεις:

  $A \rightarrow B$

  $A \rightarrow C$

  $CG \rightarrow H$

  $CG \rightarrow I$

  $B \rightarrow H$
- Τότε εξάγονται και οι μη τετριμμένες (μέσω Armstrong)
  - $A \rightarrow H$, από **<u>μεταβατικότητα</u>** μέσω $A \rightarrow B$, $B \rightarrow H$
  - $A \rightarrow BC$, από **<u>ένωση</u>** μέσω $A \rightarrow B$, $A \rightarrow C$
  - $AG \rightarrow I$, από **<u>ψευδό-μεταβατικότητα</u>** μέσω $A \rightarrow C$, $CG \rightarrow I$
  - $CG \rightarrow HI$, από **<u>ένωση</u>** μέσω $CG \rightarrow I$, $CG \rightarrow H$

Κάλυψη

- Ένα σύνολο συναρτησιακών εξαρτήσεων F μιας σχέσης R, λέμε ότι καλύπτει ένα άλλο σύνολο συναρτησιακών εξαρτήσεων G της σχέσης R, αν το σύνολο G μπορεί να εξαχθεί από το F, με την εφαρμογή κανόνων συνεπαγωγής.
  - Ώστε το G να είναι υποσύνολο του F^+

Υπολογισμός του X^+

- Σε ένα σχήμα R, το σύνολο των γνωρισμάτων που εξαρτώνται από το σύνολο γνωρισμάτων X, δεδομένου ενός συνόλου συναρτησιακών εξαρτήσεων F, καλείται X^+:
- Αλγόριθμος:

  X^+ = X
  Αρχή_επανάληψης
  Παλιό_X^+ = X^+
    Για κάθε συναρτησιακή εξάρτηση Y → Z στο F
      Αν Y ⊆ X^+ τότε X^+ = X^+ ∪ Z
  μέχρι_ότου (Παλιό_X^+ == X^+)

Παράδειγμα 1

$X_+ = X$
Αρχή_επανάληψης
Παλιό_$X_+ = X_+$
Για κάθε συναρτησιακή εξάρτηση $Y \rightarrow Z$ στο F
Αν $Y \subseteq X_+$ τότε $X_+ = X_+ \cup Z$
μέχρι_ότου (Παλιό_$X_+ == X_+$)

- Έστω R = (A,B,C,G,H,I)
- Και F= {$A \rightarrow B$, $A \rightarrow C$, $CG \rightarrow H$, $CG \rightarrow I$, $B \rightarrow H$}
- Για το $(AG)^+$:
  - AG
  - ABG , από $A \rightarrow B$
  - ABCG , από $A \rightarrow C$
  - ABCGH , από $CG \rightarrow H$ (ή από $B \rightarrow H$)
  - ABCGHI , από $CG \rightarrow I$

Παράδειγμα 2

$X_+ = X$
Αρχή_επανάληψης
Παλιό_X_+ = X_+
Για κάθε συναρτησιακή εξάρτηση $Y \rightarrow Z$ στο F
Αν $Y \subseteq X_+$ τότε $X_+ = X_+ \cup Z$
μέχρι_ότου (Παλιό_X_+ == X_+)

- Έστω R = (A,B,C,D,E)
- F= {A→B, A→C, BC→D, D→E}

- $X^+ = A^+$
  - 0. A^+ = {A}
  - 1. A^+ = {A, B} , από A→B
  - 2. A^+ = {A, B, C}, από A→C
  - 3. A^+ = {A, B, C, D}, από BC→D
  - 4. A^+ = {A, B, C, D, E}, από D→E

- $X^+ = BC^+$
  - 0. BC^+ = {B, C}
  - 1. BC^+ = {B, C, D}, από BC→D
  - 2. BC^+ = {B, C, D, E}, από D→E

Παράδειγμα 3

$X_+ = X$
Αρχή_επανάληψης
Παλιό_$X_+ = X_+$
Για κάθε συναρτησιακή εξάρτηση $Y \rightarrow Z$ στο F
Αν $Y \subseteq X_+$ τότε $X_+ = X_+ \cup Z$
μέχρι_ότου (Παλιό_$X_+ == X_+$)

- Έστω R = (A,B,C,D,E) και F= {D → E, BC → D, A → C, A → D}
- Τα αποτελέσματα του αλγορίθμου για μερικές περιπτώσεις του X.
  - Για το $(A)^+$: ACDE
  - Για το $(B)^+$: B
  - Για το $(C)^+$: C
  - Για το $(D)^+$: DE
  - Για το $(E)^+$: E
  - Για το $(BC)^+$: BCDE
  - Για το $(AB)^+$: ABCDE

Ελάχιστη Κάλυψη

- Δοθέντος ενός συνόλου F από συναρτησιακές εξαρτήσεις (ΣΕ), το ελάχιστο σύνολο ισοδύναμων συναρτησιακών εξαρτήσεων καλείται **Fc** (**ελάχιστο κάλυμμα ή ελαχιστοποιημένη κάλυψη).**
  - Η ελαχιστοποιημένη κάλυψη περιέχει εξαρτήσεις με όσο το δυνατόν πιο οικονομική σύνταξη.
  - Η ελαχιστοποιημένη κάλυψη περιέχει όλες τις εξαρτήσεις ώστε το κλείσιμο της να ισούται με το κλείσιμο του αρχικού συνόλου F^+

- Αλγόριθμος υπολογισμού της ελάχιστης κάλυψης ενός συνόλου ΣΕ
  - Δημιουργούμε ένα νέο σύνολο G ισοδύναμο του F όπου φροντίζουμε να έχουμε ΣΕ, με μόνο ένα γνώρισμα στο δεξιό μέλος της συνάρτησης. (Αποσύνθεση)
  - Αντικαθιστούμε τις ΣΕ με άλλες, που έχουν λιγότερα γνωρίσματα στο αριστερό μέλος εφόσον δεν επηρεάζουν την κλειστότητα του.
  - Αφαιρούμε από το G όλες τις ΣΕ που δεν επηρεάζουν την κλειστότητα του G αν αφαιρεθούν, (πλεονάζουσες)
  - Συγχωνεύουμε τις ΣΕ που έχουν το ίδιο αριστερό μέλος

Παράδειγμα 1

- Έστω R = (A,B,C,D)
- F= {A→BC, B→C, A→B, AB→C}
  - Αποσυνθέτουμε συναρτησιακές εξαρτήσεις
    - Η A→BC γίνεται, A→B, A→C
    - G = {A→B, A→C , B→C, A→B, AB→C}
  - Αφαιρώ τις πλεονάζουσες
    - Δεν επηρεάζει την κλειστότητα η αφαίρεση μιας A→B (υπάρχει δύο φορές)
    - G = {A→B, A→C , B→C, AB→C}
  - Αφαιρώ την AB→C αφού δεν επηρεάζει την κλειστότητα (συνεπάγεται από A→C και B→C)
    - G = {A→B, A→C , B→C}
  - Αφαιρώ την A→C αφού δεν επηρεάζει την κλειστότητα (συνεπάγεται από A→B και B→C)
    - G = {A→B, B→C}
- Η ελάχιστη κάλυψη είναι{A→B, B→C}

Παράδειγμα 2

- Έστω R = (A,B,C,D,E,F,G,H)
- F= {A→B, ABCD→E, EF→G, EF→H, ACDF→EG}
  - Αποσυνθέτουμε συναρτησιακές εξαρτήσεις
    - Η ACDF→EG γίνεται, ACDF→E, ACDF→G
    - G = {A→B, ABCD→E, EF→G, EF→H, ACDF→E, ACDF→G}
  - Αντικαθιστώ την ABCD→E με ACD→E, αφού υπάρχει η A→B
    - G = {A→B, ACD→E, EF→G, EF→H, ACDF→E, ACDF→G}
  - Αφαιρώ την ACDF→G αφού δεν επηρεάζει την κλειστότητα (συνεπάγεται από ACD→E και EF→G)
    - G = {A→B, ACD→E, EF→G, EF→H, ACDF→E}
  - Αφαιρώ την ACDF→E πού αποτελεί πλεονασμό αφού υπάρχει η ACD→E.
    - G = {A→B, ACD→E, EF→G, EF→H}
- Η ελάχιστη κάλυψη είναι

  {A→B, ACD→E, EF→G, EF→H}

Ελάχιστη Κάλυψη

- Ένας αναλυτικός αλλα χρονοβόρος αλγόριθμος για την εύρεση της ελάχιστης κάλυψης G ενός συνόλου συναρτησιακών εξαρτήσεων F είναι:

G = F

!2 Αποσυνθέτω για να φτιάξω ΣΕ με δεξί μέλος να περιέχει ένα γνώρισμα
Αποσύνθεσε όλες τις ΣΕ με πολλαπλό δεξί μέλος.

!3 Αφαιρώ τις πλεονάζουσες
Για κάθε ΣΕ X → A στο G επανάλαβε
    Υπολόγισε το X^+ με αναφορά στο σύνολο {G - (X → A)}
    Αν το X^+ περιέχει το A τότε
        G = {G - (X → A)}, δηλαδή αφαιρώ αυτήν την ΣΕ
    Τέλος_Αν
Τέλος_Επανάληψης

!4 Ελαχιστοποιώ τα αριστερά μέλη
Για κάθε ΣΕ X → A στο G με πολλαπλό X επανάλαβε
    Για κάθε γνώρισμα B στο X επανάλαβε
        Υπολόγισε το $(X\text{-}B)^+$ με αναφορά στο σύνολο {G - (X → A)} ∪ {(X-B) → A}
        Αν το $(X\text{-}B)^+$ περιέχει το B τότε
            G = {G -(X → A)} ∪ {(X-B) → A}, δηλαδή αφαιρώ αυτό το γνώρισμα από το αριστερό μέλος της ΣΕ
        Τέλος_Αν
    Τέλος_Επανάληψης
Τέλος_Επανάληψης

Παράδειγμα 3

- Έστω R = (A,B,C,D,E,F) με ΣΕ:
- F= {AB→D, B→C, AE→B, A→D, D→EF}

  - Εφαρμόζω τον αλγόριθμο:

  Βήμα 1:

  G = {AB→D, B→C, AE→B, A→D, D→EF}

  Βήμα 2:

  Αποσυνθέτουμε συναρτησιακές εξαρτήσεις

  - G = {AB→D, B→C, AE→B, A→D, D→E, D→F}

Παράδειγμα 3

Βήμα 3:

Αφαιρώ τις πλεονάζουσες

- Για την $AB \to D$, στο $\{B \to C, AE \to B, A \to D, D \to E, D \to F\}$ βρίσκω $AB_+ = ABCDEF$, που περιέχει το D, και την αφαιρώ
  - $G = \{B \to C, AE \to B, A \to D, D \to E, D \to F\}$
- Για την $B \to C$, στο $\{AE \to B, A \to D, D \to E, D \to F\}$ βρίσκω $B_+ = B$, δεν περιέχει το C το G παραμένει
  - $G = \{B \to C, AE \to B, A \to D, D \to E, D \to F\}$
- Για την $AE \to B$, στο $\{B \to C, A \to D, D \to E, D \to F\}$ βρίσκω $AE_+ = AEDF$, δεν περιέχει το B το G παραμένει
  - $G = \{B \to C, AE \to B, A \to D, D \to E, D \to F\}$
- Για την $A \to D$, στο $\{B \to C, AE \to B, D \to E, D \to F\}$ βρίσκω $A_+ = A$, δεν περιέχει το D το G παραμένει
  - $G = \{B \to C, AE \to B, A \to D, D \to E, D \to F\}$
- Για την $D \to E$, στο $\{B \to C, AE \to B, A \to D, D \to F\}$ βρίσκω $D_+ = DF$, δεν περιέχει το E το G παραμένει
  - $G = \{B \to C, AE \to B, A \to D, D \to E, D \to F\}$
- Για την $D \to F$, στο $\{B \to C, AE \to B, A \to D, D \to E\}$ βρίσκω $D_+ = DE$, δεν περιέχει το F το G παραμένει
  - $G = \{B \to C, AE \to B, A \to D, D \to E, D \to F\}$

Παράδειγμα 3

Βήμα 4:

Ελαχιστοποιώ τα αριστερά μέλη στην

G = {B→C, AE→B, A→D, D→E, D→F}

- Για την AE→B, που είναι η μόνη με πολλαπλό αριστερό μέλος
  - **Για το A: στο {B→C, <u>E→B</u>, A→D, D→E, D→F}**
    - » **E+ = EBC που δεν περιέχει το A η** AE→B παραμένει στο G
    - » **G = {B→C, <u>AE→B</u>, A→D, D→E, D→F}**
  - **Για το E: στο {B→C, <u>A→B</u>, A→D, D→E, D→F}**
    - » **A+ =ABCDEF που περιέχει το E, άρα αντικαθιστώ την** AE→B, με το A→B
    - » **G = {B→C, <u>A→B</u>, A→D, D→E, D→F}**

Η ελάχιστη κάλυψη είναι:

G = {B→C, A→B, A→D, D→E, D→F}

Εύρεση κλειδιών

- Το κλείσιμο του συνόλου των γνωρισμάτων που αποτελεί κλειδί, είναι το σύνολο των γνωρισμάτων της σχέσης.
- Για να βρω <u>τα κλειδιά μιας σχέσης</u>
  - Ξεκινάω να ελέγξω το κλείσιμο <u>ατομικών γνωρισμάτων</u>.
    - Αν το κλείσιμο του μου δίνει το σύνολο των γνωρισμάτων της σχέσης αποτελεί κλειδί.
  - Αν κανένα ατομικό γνώρισμα δεν είναι κλειδί δοκιμάζω με το κλείσιμο ενός συνόλου με δύο γνωρίσματα. Δοκιμάζω όλα τα διμελή σύνολα μέχρι να βρω κλειδί.
  - Αν δεν έχω βρει κλειδί η διαδικασία συνεχίζεται με τριμελή σύνολα κ.ο.κ.
- Μια <u>βελτιστοποίηση</u> που μπορούμε να κάνουμε είναι:
  - Εντοπίζω τα γνωρίσματα που <u>δεν βρίσκονται σε κανένα δεξί μέλος των εξαρτήσεων.</u>
    - Το σύνολο τους X θα είναι απαραίτητα μέρος του κλειδιού.
    - Αν το σύνολο αυτό δεν είναι κενό, ξεκινάω για την εύρεση κλειδιού με το κλείσιμο αυτού του συνόλου και όχι με ατομικά γνωρίσματα.
    - Αν το κλείσιμο του δεν είναι το σύνολο των γνωρισμάτων της σχέσης άρα δεν είναι κλειδί, τότε προσθέτω ένα ακόμα γνώρισμα στο σύνολο και συνεχίζω.

Παράδειγμα

- Έστω η σχέση R(A,B,C,D,E) και το σύνολο ΣΕ: F={ A ->B, BC ->E, ED->A}

  - Στα δεξιά μέλη δεν έχω D ούτε C. Ξεκινάω με το CD.

  - CD^+ = CD, δεν είναι το σύνολο των γνωρισμάτων της σχέσης άρα συνεχίζω με

  - Προσθέτω το A:
    - ACD^+ = ACDBE, Το ACD είναι κλειδί

  - Προσθέτω το B:
    - BCD^+ = BCDEA, Το BCD είναι κλειδί

  - Προσθέτω το E:
    - CDE^+ = ACDBE, Το CDE είναι κλειδί

- Κλειδιά είναι τα ACD, BCD, CDE

Κλειδιά και Γνωρίσματα

- Τα γνωρίσματα που είναι μέρη κλειδιών ονομάζονται πρωτεύοντα (prime).
  - Στο προηγούμενο παράδειγμα με κλειδιά τα ACD, BCD, CDE, πρωτεύοντα είναι τα γνωρίσματα A, B, C, D, E, και δεν υπάρχουν μη-πρωτεύοντα γνωρίσματα.
  - Στην σχέση R(A, B, C, D, E), αν κλειδιά είναι τα AB και BC τότε:
    - Πρωτεύοντα(prime): A, B, C
    - Μη πρωτεύοντα(non-prime): D, E