ΗΥ 360 – Αρχεία και Βάσεις Δεδομένων
Χειμερινό Εξάμηνο 2013
Διδάσκουσα: Ειρήνη Φουντουλάκη

Ενδεικτικές λύσεις 3ης Σειράς Ασκήσεων

Άσκηση 1 [25]Για τα παρακάτω σχεσιακά σχήματα και τα αντίστοιχα σύνολα συναρτησιακών εξαρτήσεων, βρείτε: (α) τις μη-τετριμμένες συναρτησιακές εξαρτήσεις που προκύπτουν και οι οποίες έχουν μόνο ένα γνώρισμα το δεξί μέλος, (β) τα κλειδιά των σχέσεων.

I. R(A, B, C, D, E), $\{AB \rightarrow C, C \rightarrow D, D \rightarrow E, E \rightarrow A\}$
II. S(A, B, C, D), $\{A \rightarrow B, A \rightarrow C, D \rightarrow A\}$
III. T(A, B, C, D), $\{AB \rightarrow C, BC \rightarrow D, D \rightarrow A, AD \rightarrow B\}$
IV. U(A, B, C, D), $\{A \rightarrow B, B \rightarrow C, C \rightarrow D, CD \rightarrow A\}$

__Λύση__

AB → D, (από *AB → C, C → D με μεταβατικότητα*)

C → A, (από *C → D, D → E, E → A με μεταβατικότητα*)

D → A, (από *D → E, E → A με μεταβατικότητα*)

C → E, (από *C → D, D → E με μεταβατικότητα*)

Και οι εξαρτήσεις:

ABC → D(από *AB → D επαύξηση ABC→ DC και αποσύνθεση ABC→ D)*

ABD → C(από *AB → C επαύξηση ABD→ CD και αποσύνθεση ABD→ C)*

AC → D, (από *C → D επαύξηση AC→ AD και αποσύνθεση AC→ D)*

BC → D, (από *C → D επαύξηση BC→ BD και αποσύνθεση BC→ D)*

BC → E, (από *C → E επαύξηση BC→ BE και αποσύνθεση BC→ E)*

AC → E, (από *C → E επαύξηση AC→ AE και αποσύνθεση AC→ E)*

BC → A (από *C → A επαύξηση BC→ BA και αποσύνθεση BC→ A)*

BD → A (από *D → A επαύξηση BD→ BA και αποσύνθεση BD→ A)*

BD → C (από *D → A ψευδομεταβατικότητα AB→ C)*

BCD → A (από *BC → A επαύξηση BCD→ AD και αποσύνθεση BCD→A)*

***AD → E** (από D → E επαύξηση AD → AE και αποσύνθεση AD → E)*

***BD → E** (από D → E επαύξηση BD → BE και αποσύνθεση BD → E)*

***CD → E** (από D → E επαύξηση DE → AD και αποσύνθεση CD → E)*

***BE* → *A* (από Ε → Α επαύξηση ΒΕ → ΑΒ και αποσύνθεση ΒΕ → Α)**

***CE* → *A* (από Ε → Α επαύξηση CE → AC και αποσύνθεση CE → A)**

***DE* → *A* (από Ε → Α επαύξηση DE → AD και αποσύνθεση DE → A)**

***BE* → *C* (από Ε → Α ψευδομεταβατικότητα AB → C)**

CDE* → *A *(από Ε → Α επαύξηση DE → DA, αποσύνθεση DE → A, επαύξηση CDE → AD και αποσύνθεση CDE → A)*

BCDE* → *A *(από ΒΕ → Α επαύξηση BCE → A, αποσύνθεση BCE → A, επαύξηση BCDE → AD και αποσύνθεση BCDE → A)*

Αντίστοιχα, καθώς επίσης και με τον υπολογισμό της κλειστότητας κάθε γνωρίσματος (ή συνόλου γνωρισμάτων) υπολογίζονται και οι παρακάτω μη-τετριμμένες συναρτησιακές εξαρτήσεις:

AB → E, ABC → E, ABD → E, ABE → C, ABE → D, ACD → E, ACE → D, BCD → E, BCE → A, BCE → D, BDE → A, BDE → C, ABCD → E, ABCE → D, ABDE → C.

Άσκηση 2 [25]Έστω η σχέση R(A,B,C,D,E,F,G,H) με τις συναρτησιακές εξαρτήσεις$\{C \to D, B \to H, FD \to G, E \to AG, H \to A\}$.

(α) Να βρείτε το κλειδί της σχέσης

(β) Είναι σε Τρίτη κανονική μορφή; Αν όχι να κάνετε την αναγκαία αποσύνθεση, εξηγώντας το κάθε βήμα.

(α)

Στα δεξιά μέλη των Σ.Ε. δεν υπάρχουν τα B,C,E,F.

Οπότε θα βρούμε το **(BCEF)**+

(BCEF)	B→H
BCEFH	H→A
BCEFHA	C→D
BCEFHAD	FD→G
BCEFHADG ,	δηλαδή ABCDEFGH

Το **(BCEF)+** περιέχει όλα τα γνωρίσματα της αρχικής σχέσης R, άρα το **BCEF** αποτελεί κλειδί.

(β)

Για να είναι η σχέση **<u>R (A,B,C,D,E,F,G,H)</u>** σε **<u>3η κανονική μορφή</u>** πρέπει να είναι **<u>σε δεύτερη κανονική μορφή (2NF)</u>** και να **<u>μην υπάρχουν μεταβατικές συναρτησιακές εξαρτήσεις.</u>**

Από το προηγούμενο ερώτημα ξέρουμε ότι το κλειδί της σχέσης είναι το **BCEF**.
Άρα:
B,C,E,F: πρωτεύοντα γνωρίσματα
A,D,G,H : **μη** πρωτεύοντα γνωρίσματα

Για να είναι η σχέση R σε **3NF**, πρέπει να είναι σε **2NF** και να **μην** υπάρχουν **μεταβατικές συναρτησιακές εξαρτήσεις**.
Οπότε πρώτα ελέγχουμε εάν είναι σε **2NF**.

Η σχέση R (A,B,C,D,E,F,G,H) είναι σε 2NF?

Όχι, γιατί υπάρχουν οι παρακάτω Σ.Ε. που παραβιάζουν την δεύτερη κανονική μορφή 2NF, γιατί τα δεξιά μέλη των Σ.Ε. δεν εξαρτώνται από ολόκληρο το κλειδί αλλά από ένα μέρος αυτού.

- **B→H**
- **C→D**
- **E→AG**
- **FD→G**

Οπότε μετατρέπουμε το αρχικό σχήμα σε **2NF**, κάνοντας τις παρακάτω αποσυνθέσεις.

R1 = (B,C,E,F)	με κλειδί το BCEF
R2 = (B,H,A)	με κλειδί το B
R3 = (C,D)	με κλειδί το C
R4 = (E,A,G)	με κλειδί το E
R5 = (F,D,G)	με κλειδί το FD

Οι παραπάνω σχέσεις R1,R2,R3,R4,R5 που προέκυψαν είναι σε 3NF?

Η R2 δεν είναι σε 3NF γιατί υπάρχουν οι παρακάτω μεταβατικές εξαρτήσεις.

- B→ **H**
- **H**→ A

Άρα την αποσυνθέτουμε σε R_{21}=(B,H) και R_{22}=(H,A).
Οπότε, προέκυψαν οι παρακάτω σχέσεις:

- R1 = (B,C,E,F) με κλειδί το BCEF
- R_{21}=(B,H) με κλειδί το B
- R_{22}=(H,A) με κλειδί το H
- R3 = (C,D) με κλειδί το C
- R4 = (E,A,G) με κλειδί το E
- R5 = (F,D,G) με κλειδί το FD

Άσκηση 3[25] Θεωρείστε τη σχέση R(A,B,C,D,E) για την οποία ισχύουν οι ακόλουθες εξαρτήσεις$F1 = \{A \to B, B \to E, E \to C, C \to B, A \to E, B \to AD\}$.Βρείτε την ελάχιστη κάλυψη αυτού του συνόλου των εξαρτήσεων. Είναι μοναδική; Αν όχι, βρείτε μια διαφορετική ελάχιστη κάλυψη.

- Αποσυνθέτουμε συναρτησιακές εξαρτήσεις που έχουν παραπάνω από ένα γνωρίσμα στο δεξί μέλος :

  Η $B \rightarrow AD$ γίνεται $B \rightarrow A$, $B \rightarrow D$, Οπότε το αρχικό σύνολο γίνεται:
  $G = \{A \rightarrow B, B \rightarrow E, E \rightarrow C, C \rightarrow B, A \rightarrow E, B \rightarrow A, B \rightarrow D\}$.

- Αφαιρούμε τις συναρτησιακές εξαρτήσεις που δεν επηρεάζουν το σύνολο G.
  - Για την $A \rightarrow B$ στο $\{B \rightarrow E, E \rightarrow C, C \rightarrow B, A \rightarrow E, B \rightarrow A, B \rightarrow D\}$ βρίσκω $A^+=ABCDE$ που περιέχει το B και την αφαιρώ
    Άρα το σύνολο $G = \{B \rightarrow E, E \rightarrow C, C \rightarrow B, A \rightarrow E, B \rightarrow A, B \rightarrow D\}$
  - Για την $B \rightarrow E$ στο $\{E \rightarrow C, C \rightarrow B, A \rightarrow E, B \rightarrow A, B \rightarrow D\}$
    βρίσκω $B^+ = ABCDE$ που περιέχει E και την αφαιρώ
    Άρα το σύνολο $G = \{E \rightarrow C, C \rightarrow B, A \rightarrow E, B \rightarrow A, B \rightarrow D\}$
  - Για την $E \rightarrow C$ στο $\{C \rightarrow B, A \rightarrow E, B \rightarrow A, B \rightarrow D\}$
    βρίσκω $E^+ = E$ που δεν περιέχει το C άρα το G παραμένει ως έχει .

Ομοίως ελέγχω και τις υπόλοιπες Σ.Ε. στο σύνολο G και διαπιστώνω ότι δεν μπορεί να αφαιρεθεί καμία άλλη Σ.Ε. από το σύνολο G.

Εφαρμόζοντας τον κανόνα της ένωσης η ελάχιστη κάλυψη που προκύπτει είναι:
$G = \{E \rightarrow C, C \rightarrow B, A \rightarrow E, B \rightarrow AD\}$ που είναι και μοναδική.

Εάν κάποιος εφαρμόσει τον αλγόριθμο από την αρχή και αλλάξει την σειρά των Σ.Ε. μπορεί να καταλήξει και σε άλλη ελάχιστη καλύψη όπως:

$$G = \{A \rightarrow B, B \rightarrow E, E \rightarrow C, C \rightarrow B, B \rightarrow A, B \rightarrow D\}$$

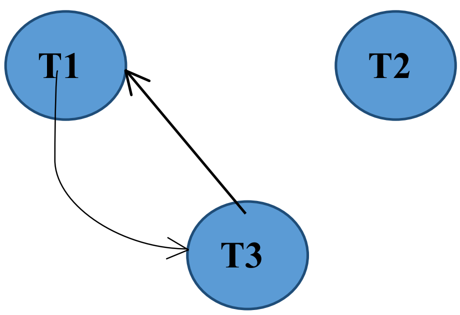
Άσκηση 4[25] Για κάθε μια από τις παρακάτω εκτελέσεις

$S1 = r_2[x]\ w_3[y]\ r_1[z]\ r_3[z]\ w_3[z]\ r_1[y]\ w_2[x]$
$S2 = w_2[y]\ r_3[x]\ r_1[z]\ r_3[z]\ r_2[y]\ w_3[x]$

να εντοπίσετε τις συγκρουόμενες ενέργειες και να σχεδιάσετε το γράφο σειριακοποιησιμότητας. Να εξηγήσετε τη σημασία του κυκλικού σε σχέση με τον άκυκλο γράφο σειριακοποιησιμότητας. Στη συνέχεια να εφαρμόσετε το πρωτόκολλο 2PLκαι να δείξετε προγράμματα σύγχρονης εκτέλεσης των ίδιων δοσοληψιών και τα ισοδύναμα σειριακά.

<u>Λύση</u>

Συγκρουόμενες ενέργειες: $w_3[y] \ldots r_1[y]$, $r_1[z] \ldots w_3[z]$.
Άρα ο γράφος σειριακοποιησιμότητας είναι:



Ο γράφος είναι κυκλικός. Αυτό σημαίνει ότι υπάρχει η απαίτηση δύο δοσοληψίες να εκτελεστούν με δύο διαφορετικές ιεραρχίες (T1T3) και (T3T1). Για παράδειγμα, το ζεύγος ενεργειών w_3[y]…r_1[y] απαιτεί να εκτελεστεί πρώτα η T3 και μετά η T1και ταυτόχρονα το ζεύγος ενεργειών r_1[z]… w_3[z] απαιτεί να εκτελεστεί πρώτα η T1 και μετά η T3. Συνεπώς λόγω αυτού του κύκλου, η εκτέλεσηS1 δεν είναι σειριακοποιήσιμη (δε μπορεί δηλαδή να αντιστοιχιστεί ως έχει σε μια μοναδική ακολουθία εκτέλεσης των T1, T2 και T3).

Για να επιτύχουμε σειριακοποιησιμότητα, χρησιμοποιούμε το πρωτόκολλο 2PL όπως φαίνεται παρακάτω (θεωρούμε ότι τα readlocks είναι shared και tawritelocks είναι exclusive. Επίσης με το commit εννοούνται και όλα τα unlock για τα locks που έχει πάρει η αντίστοιχη διεργασία):

T1					RL1(x)	R1(z)		
T2	RL2(x)	R2(x)						
T3			WL3(y)	W3(y)			RL3(z)	R3(z)

T1			RL1(y)	wait		R1(y)	C1
T2			**DEADLOCK**				
T3	WL3(z)	wait			abort		

T1					
T2	WL2(x)	W2(x)	**C2**		
T3				restart	C3

Ισοδύναμο σειριακό πρόγραμμα με αυτή την εκτέλεση είναι το T3T2T1. Επίσης δεδομένου ότι η T2 δε συγκρούεται με καμία από τις άλλες δύο δοσοληψίες, ισοδύναμα προγράμματα είναι και τα T2T3T1 και T3T1T2 (προϋπόθεση δηλαδή για ισοδυναμία είναι μόνο να εκτελεστεί η T3 πριν την T1).

Άλλη εφαρμογή του 2PLείναι η παρακάτω:

T1					RL1(x)	R1(z)		
T2	RL2(x)	R2(x)						
T3			WL3(y)	W3(y)			RL3(z)	R3(z)

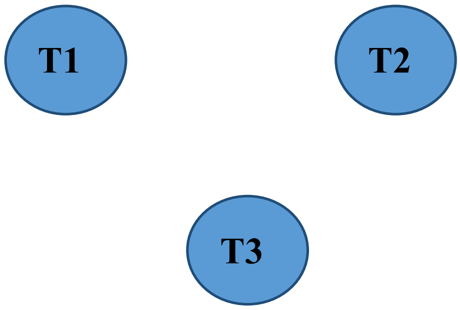
T1			RL1(y)	wait	abort		
T2			**DEADLOCK**				
T3	WL3(z)	wait				W3(z)	C3

T1				restart	C1
T2	WL2(x)	W2(x)	**C2**		
T3					

Ισοδύναμο σειριακό πρόγραμμα με αυτή την εκτέλεση είναι το T1T2T3. Επίσης δεδομένου ότι η T2 δε συγκρούεται με καμία από τις άλλες δύο δοσοληψίες, ισοδύναμα προγράμματα είναι και τα T2T1T3 και T1T3T2 (προϋπόθεση δηλαδή για ισοδυναμία είναι μόνο να εκτελεστεί η T1 πριν την T3).

S2= $\mathbf{w_2[y]\ r_3[x]\ r_1[z]\ r_3[z]\ r_2[y]\ w_3[x]}$

Συγκρουόμενες ενέργειες: δεν υπάρχουν.
Άρα ο γράφος σειριακοποιησιμότητας είναι:



Ο γράφος είναι άκυκλος (για την ακρίβεια στο συγκεκριμένο παράδειγμα ο γράφος δεν έχει καν ακμές). Αυτό σημαίνει δεν υπάρχουν διαφορετικές απαιτήσεις για τη σειρά εκτέλεσης των T1, T2 και T3 και άρα η εκτέλεση S2 είναι σειριακοποιήσιμη.
Το πρωτόκολλο 2PL δεν είναι απαραίτητο να χρησιμοποιηθεί δεδομένου ότι επιτυγχάνεται σειριακοποιησιμότητα και χωρίς τη χρήση locks. Σε κάθε περίπτωση, μια πιθανή εφαρμογή του πρωτοκόλλου είναι:

T1					RL1(z)	R1(z)	C1
T2	WL2(y)	W2(y)					
T3			RL3(x)	R3(x)			

T1									
T2			RL2(y)	R2(y)	C2				
T3	RL3(z)	R3(z)					WL3(x)	W3(x)	C3

Ισοδύναμο σειριακό πρόγραμμα με αυτή την εκτέλεση είναι το T1T2T3. Δεδομένου όμως ότι δεν υπάρχει καμία σύγκρουση μεταξύ των ενεργειών των τριών δοσοληψιών, όλες οι δυνατές σειριακές εκτελέσεις είναι ισοδύναμες, δηλαδή και οι T2T3T1, T1T3T2, T2T1T3, T3T1T2, T3T2T1.