



Partea I

Baze de date relaționale

Modelul relațional



- Structura modelului relațional
- Constrângeri de integritate



Structura modelului relațional

- Conceptele ce stau la baza modelului relațional:
 - relație
 - tabel } diferă ca noțiuni dar sunt strâns legate
- **Relație** - noțiune formală - domeniul teoriei mulțimilor
- **Tabel** - noțiune simplă și intuitivă

Definiții

Fie date mulțimile D_1 și D_2 .

Produsul cartezian dintre D_1 și D_2 (notat $D_1 \times D_2$) se definește ca fiind mulțimea perechilor ordonate (v_1, v_2) , $v_1 \in D_1$, $v_2 \in D_2$.

O **relație** pe mulțimile D_1 și D_2 (numite *domeniile relației*) este o submulțime a produsului cartezian $D_1 \times D_2$.

Exemplu

CURSURI

01	Fizica
03	Chimie
04	Biologie

D1 = ?

D2 = ?

Definiții

Fie date mulțimile D_1 și D_2 .

Produsul cartezian dintre D_1 și D_2 (notat $D_1 \times D_2$) se definește ca fiind mulțimea perechilor ordonate (v_1, v_2) , $v_1 \in D_1$, $v_2 \in D_2$.

O **relație** pe mulțimile D_1 și D_2 (numite *domeniile relației*) este o submulțime a produsului cartezian $D_1 \times D_2$

■ Mulțimile D_1 și D_2 sunt finite sau nu ?

- Este de dorit ca domeniile să aibă dimensiuni infinite, a.î. să putem presupune existența unei valori care nu e prezentă în baza de date.
- În practică relațiile trebuie să fie finite deoarece bazele de date trebuie stocate în sisteme computerizate de dimensiuni finite.



Vom presupune că **bazele de date sunt alcătuite din relații finite, definite pe domenii infinite**

Definițiile anterioare pot fi generalizate

Fie $n > 0$ mulțimi D_1, D_2, \dots, D_n , nu neapărat distincte

STUDENTI

276545	Ionescu	Maria	25/11/1980
485745	Popescu	Ana	23/04/1981
200768	Georgescu	Paul	12/02/1981
587614	Luca	Radu	10/10/1980
937653	Maftai	Lucian	01/12/1980

Definițiile anterioare pot fi generalizate

Fie $n > 0$ mulțimi D_1, D_2, \dots, D_n , nu neapărat distincte

Produsul cartezian $D_1 \times D_2 \times \dots \times D_n$ este dat de mulțimea n -tuplurilor v_1, v_2, \dots, v_n , unde $v_i \in D_i, i = \overline{1, n}$

O **relație matematică** pe domeniile D_1, D_2, \dots, D_n este o submulțime a produsului cartezian $D_1 \times D_2 \times \dots \times D_n$

- n - gradul produsului cartezian și al relației (dat de numărul componentelor produsului cartezian)
- cardinalitatea relației este dată de numărul n -tuplurilor
- analogia relație – mulțime implică:
 - între n -tupluri nu există o ordine predefinită
 - n -tuplurile unei relații sunt distincte unele de celelalte \Rightarrow un tabel reprezintă o relație doar dacă liniile sale sunt diferite între ele

Notația non-pozițională – se asociază nume (atribute) domeniilor unei relații

Atributele descriu rolurile jucate de domenii

EchipaGazda	EchipaOaspete	GoluriGazde	GoluriOaspeti
Real Madrid	Liverpool	3	1
Liverpool	Milan	2	0
Real Madrid	Roma	1	2
Roma	Milan	0	1

Fig. 2.2 Relație cu atribute

Pentru formalizarea conceptelor introduse anterior, definim:

D - mulțimea domeniilor

X - mulțimea atributelor

funcția $DOM: X \rightarrow D$ care asociază fiecărui atribut un domeniu

Definiție. Un **tuplu** definit pe o mulțime de atribute este o **funcție** t care asociază fiecărui atribut $A \in X$ o valoare din domeniul $DOM(A)$

Definiție. O **relație** pe o mulțime de atribute X este o mulțime de tupluri definite pe X

Notăție. Fie t un tuplu definit pe o mulțime de atribute X și $A \in X$ un atribut. Prin $t[A]$ se notează valoarea tuplului t în domeniul $DOM(A)$

Definiție. Un **tuplu** definit pe o mulțime de atribute este o *funcție* t care asociază fiecărui atribut $A \in X$ o valoare din domeniul $DOM(A)$

Definiție. O **relație** pe o mulțime de atribute X este o mulțime de tupluri definite pe X

Exemplu

Fie t primul tuplu din relația prezentată în figura 2.2

$t[\text{EchipaOaspete}] = \text{Liverpool}$

$t[\text{EchipaOaspete}, \text{GoluriOaspeti}] = \text{Liverpool}, 1$

- BD sunt formate din mai multe relații
 - tuplurile acestor relații conțin valori comune atunci când acest lucru este necesar pentru stabilirea unor corespondențe

STUDENTI

NrInreg	Nume	Prenume	DataNastere
276545	Ionescu	Maria	25/11/1980
485745	Popescu	Ana	23/04/1981
200768	Georgescu	Paul	12/02/1981
587614	Luca	Radu	10/10/1980
937653	Maftai	Lucian	01/12/1980

EXAMENE

Student	Nota	Curs
276545	8	01
276545	9	04
937653	9	01
200768	9	04

CURSURI

Cod	Denumire	Titular
01	Fizica	Melinte
03	Chimie	Mardare
04	Chimie	Dascalu

Fig. 2.3 Exemplu de bază de date relațională

Schema unei relații este formată din numele relației R și o mulțime de attribute $X = \{A_1, A_2, \dots, A_n\}$ și se notează $R(X)$; fiecărui atribut i se asociază un domeniu

Schema bazei de date este formată dintr-o mulțime de scheme de relații:

$$\mathbf{R} = \{R_1(X_1), R_2(X_2), \dots, R_n(X_n)\}$$

Instanța unei relații (pe scurt – *relația*) având schema $R(X)$ este dată de mulțimea r a tuplurilor definite pe mulțimea de attribute X

Instanța bazei de date (pe scurt – *baza de date*) având schema

$\mathbf{R} = \{R_1(X_1), R_2(X_2), \dots, R_n(X_n)\}$ este mulțimea $r = \{r_1, r_2, \dots, r_n\}$ de relații în care fiecare $r_i, i = \overline{1, n}$ este o relație definită pe schema $R_i(X_i)$

Exemplu

Schema bazei de date din figura 2.3 este

**R = {STUDENTI(NrInreg, Nume, Prenume, DataNastere),
EXAMENE(Student, Nota, Curs),
CURSURI(Cod, Denumire, Titular)}**

STUDENTI

NrInreg	Nume	Prenume	DataNastere
276545	Ionescu	Maria	25/11/1980
485745	Popescu	Ana	23/04/1981
200768	Georgescu	Paul	12/02/1981
587614	Luca	Radu	10/10/1980
937653	Maftai	Lucian	01/12/1980

EXAMENE

Student	Nota	Curs
276545	8	01
276545	9	04
937653	9	01
200768	9	04

CURSURI

Cod	Denumire	Titular
01	Fizica	Melinte
03	Chimie	Mardare
04	Chimie	Dascalu

Structura modelului relațional

- Informațiile trebuie reprezentate sub forma unor **tupluri omogene de date**
- În particular, putem reprezenta în cadrul unei relații doar tupluri ce corespund schemei relației
- În practică există cazuri în care datele disponibile nu corespund cu exactitate formatului ales

Exemplu

Se consideră schema **PERSOANĂ (Nume, Prenume, Adresă, TelefonFix)**

- Problemă: Pot exista tupluri pentru care valoarea atributului **TelefonFix** nu este disponibilă
- Rezolvarea problemei indisponibilității valorilor: s-a inclus posibilitatea ca fiecare atribut al unui tuplu să poată lua fie valori din domeniul asociat, fie o valoare specială, denumită valoare **null**
- Valoarea **null** indică absența informației
 - este o valoare suplimentară, ce nu este conținută de domeniul;
 - la definirea unei relații se pot specifica acele attribute care acceptă valori **null**

Constrângeri de integritate

STUDENTI

NrInreg	Nume	Prenume	DataNastere
200768	George	Paul	12/02/1981
937653	Maftai	Lucian	10/10/1980
937653	Luca	Radu	01/12/1980

EXAMENE

Student	Nota	Promovat	Curs
200768	11	DA	05
937653	4	DA	01
937653	6	DA	04
276545	7	DA	01

CURSURI

Cod	Denumire	Titular
01	Fizica	Melinte
03	Chimie	Mardare
04	Chimie	Dascalu



Constrângeri de integritate

- Există situații când nu orice mulțime de tupluri în cadrul unei scheme reprezintă informații corecte pentru aplicație
 - Pentru înlăturarea situațiilor de **incorectitudine a informațiilor** a fost introdus conceptul de *constrângere de integritate*
- **Constrângerea de integritate** - proprietate ce trebuie satisfăcută de toate instanțele corecte ale bazei de date
- O constrângere poate fi privită ca un *predicat* ce asociază valoarea *adevărat* sau *fals* fiecărei instanțe
- Se pot defini mai multe constrângeri pentru o bază de date
- Vom considera *corecte* (sau *legale*) acele instanțe care satisfac toate constrângerile impuse

Clasificare

- **constrângeri intra-relaționale** - sunt definite pe o singură relație a bazei de date și pot fi de două tipuri:
 - constrângeri la nivel de tuplu – pot fi evaluate pe fiecare tuplu, independent de celelalte tupluri
 - constrângeri la nivel de domeniu (la nivel de valoare) – impun restricții asupra domeniului unui atribut

- **constrângeri inter-relaționale** – implică mai multe relații

STUDENTI

NrInreg	Nume	Prenume	DataNastere
200768	George	Paul	12/02/1981
937653	Maftai	Lucian	10/10/1980
937653	Luca	Radu	01/12/1980

EXAMENE

Student	Nota	Promovat	Curs
200768	11	DA	05
937653	4	DA	01
937653	6	DA	04
276545	7	DA	01

CURSURI

Cod	Denumire	Titular
01	Fizica	Melinte
03	Chimie	Mardare
04	Chimie	Dascalu

Constrângeri de integritate

Exemplu. Fie baza de date din figura următoare

STUDENTI

NrInreg	Nume	Prenume	DataNastere
200768	George	Paul	12/02/1981
937653	Maftai	Lucian	10/10/1980
937653	Luca	Radu	01/12/1980

EXAMENE

Student	Nota	Promovat	Curs
200768	11	DA	05
937653	4	DA	01
937653	6	DA	04
276545	7	DA	01

CURSURI

Cod	Denumire	Titular
01	Fizica	Melinte
03	Chimie	Mardare
04	Chimie	Dascalu

Fig.2.4 Bază de date cu informații incorecte

- constrângeri la nivel de tuplu:
 - în primul tuplu al relației EXAMENE avem ca rezultat la un examen nota 11
 - în al doilea tuplu al relației EXAMENE un student este considerat promovat la examen deși nota sa este 4
- constrângeri la nivel de domeniu
 - pentru atributul Notă din relația EXAMENE, numai valorile cuprinse între 1 și 10 sunt permise
 - ultimele două tupluri ale relației STUDENTI conțin informații pentru doi studenți diferiți dar cu același număr de înregistrare, identificarea studenților fiind astfel ambiguă
- constrângeri inter-relaționale
 - al patrulea tuplu al relației EXAMENE are, pentru atributul Student, o valoare care nu apare printre numerele de înregistrare din relația STUDENTI
 - primul tuplu al relației EXAMENE are, pentru atributul Curs, o valoare care nu apare printre codurile cursurilor din relația CURSURI

Constrângerile pe tuplu - condiții impuse valorilor fiecărui tuplu, independent de celelalte tupluri

- sintaxă - *expresii booleene* (folosind conectorii AND, OR și NOT) ale căror atomi compară valorile atributelor implicate în constrângeri sau *expresii aritmetice* folosind valorile atributelor
- singura condiție este ca aceste expresii să utilizeze valorile unui singur tuplu

Exemple

1) Expresiile ce descriu constrângerile intra-relaționale încălcate în exemplul prezentat în figura 2.4:

$(Nota \geq 1) \text{ AND } (Nota \leq 10)$

$(Promovat = 'DA') \text{ AND } (Nota \geq 5)$

2) Fie dată schema

PLATĂ(Dată, Sumă, Deduceri, Net)

Se poate defini o constrângere care impune condiția ca suma netă să fie egală cu diferența dintre suma totală și deduceri:

$Net = Sumă - Deduceri$

Constrângeri de chei

O **cheie** este o mulțime de atribute ce ajută la identificarea în mod unic a tuplurilor unei relații

Definiții

- O mulțime de atribute K este o **super-cheie** a relației r dacă r nu conține două tupluri distincte t_1 și t_2 astfel încât $t_1[K] = t_2[K]$
- O mulțime K de atribute este o **cheie** a relației r dacă mulțimea K este o super-cheie minimală (nu există altă super-cheie K' a lui r astfel încât $K' \subset K$)

Exemplu. Se consideră următoarea relație

STUDENT

NrInreg	Nume	Prenume	DataNastere	Facultate
284328	Ionescu	Maria	29/04/59	AC
296328	Ionescu	Ana	29/04/59	TCM
587614	Ionescu	Lucian	01/05/61	Textile
934856	Popescu	Lucian	01/05/61	AC
965536	Popescu	Lucian	05/03/58	TCM

O mulțime K de attribute este o **cheie** a relației r dacă mulțimea K este o super-cheie minimală (nu există altă super-cheie K' a lui r astfel încât $K' \subset K$)

Constrângeri de integritate

Exemplu. Se consideră următoarea relație

NrInreg	Nume	Prenume	DataNastere	Facultate
284328	Ionescu	Maria	29/04/59	AC
296328	Ionescu	Ana	29/04/59	TCM
587614	Ionescu	Lucian	01/05/61	Textile
934856	Popescu	Lucian	01/05/61	AC
965536	Popescu	Lucian	05/03/58	TCM

Fig.2.5 Exemplu de relație pentru evidențierea cheilor

- Mulțimea {NrInreg} este o cheie
- Mulțimea {Nume, Prenume, DataNastere} este o cheie, dar este și super-cheie!
- Mulțimea {NrInreg, Facultate} este o super-cheie, dar nu este cheie!
- Mulțimea {Nume, Facultate} este o cheie

Întrebare: *Putem afirma acest lucru în cazul general ?*

Răspuns: nu, deoarece pot exista studenți cu același nume și care au terminat aceeași facultate

Spunem că, în acest caz, mulțimea {Nume, Facultate} este, *prin șansă*, cheie a relației

- Pentru schema **STUDENTI (NrInreg, Nume, Prenume, DataNastere, Facultate)** se pot stabili două constrângeri ce impun următoarele chei:

{NrInreg} și {Nume, Prenume, DataNastere}

Relația din figura 2.5 satisface ambele constrângeri de chei

Constrângeri de integritate

- Pentru orice relație se poate stabili cel puțin o cheie \Rightarrow se permite:
 - accesul la toate datele din baza de date
 - identificarea unică a datelor
 - stabilirea unor legături între datele conținute de diverse relații

Exemplu. Se consideră baza de date din figura 2.3

STUDENTI

NrInreg	Nume	Prenume	DataNastere
276545	Ionescu	Maria	25/11/1980
485745	Popescu	Ana	23/04/1981
200768	Georgescu	Paul	12/02/1981
587614	Luca	Radu	10/10/1980
937653	Maftai	Lucian	01/12/1980

EXAMENE

Student	Nota	Curs
276545	8	01
276545	9	04
937653	9	01
200768	9	04

CURSURI

Cod	Denumire	Titular
01	Fizica	Melinte
03	Chimie	Mardare
04	Chimie	Dascalu

- Relația EXAMENE face referire la studenții din relația STUDENTI prin NrInreg și la cursurile din relația CURSURI prin Cod
 - NrInreg este cheia relației STUDENTI
 - Cod este cheia relației CURSURI
- Valorile atributelor cheie sunt utilizate pentru referirea conținutului altor relații

Cheia primară - cheia prin intermediul căreia se realizează referințe între relații

- Constrângere: este interzisă utilizarea valorilor **null** pentru cheile primare,
 - Valorile **null** sunt permise pentru celelalte chei
- Pot apărea situații în care nu există attribute ale căror valori să fie disponibile pentru o cheie primară ⇒ se va introduce un atribut suplimentar care va fi generat și asociat fiecărui tuplu în momentul inserării în relația corespunzătoare

Constrângeri de referință

O constrângere de referință (sau cheie străină) între o mulțime de atribute X ale relației R_1 și altă relație R_2 este satisfăcută dacă valorile fiecărui tuplu din R_1 corespunzătoare mulțimii X se regăsesc printre valorile cheii primare a relației R_2 . Apar două situații:

a) Cheia relației R_2 este unică și conține un singur atribut B

Constrângerea de referință între mulțimea X (formată din atributul A) și relația R_2 este satisfăcută dacă $\forall t_1 \in R_1$ cu $t_1[A] \neq NULL$, $\exists t_2 \in R_2$ astfel încât $t_1[A] = t_2[B]$

b) Cheia relației R_2 este unică și coincide cu o mulțime K de atribute

Pentru stabilirea unei constrângeri între X și R_2 trebuie specificată o anumită ordine pentru ambele mulțimi (X și K). Indicând atributele în ordine, $X = A_1 A_2 \dots A_p$, $K = B_1 B_2 \dots B_p$, constrângerea este satisfăcută dacă $\forall t_1 \in R_1$ cu $t_1[A_i] \neq NULL$, $i = \overline{1, p}$, $\exists t_2 \in R_2$ astfel încât $t_1[A_i] = t_2[B_i]$, $i = \overline{1, p}$

Constrângeri de integritate

Exemplu. Se consideră baza de date din figura următoare

CADRE

NrInreg	Nume	Prenume
567	Dascalescu	Eugen
456	Ionescu	George
638	Popescu	Dan

AUTOVEHICUL

NrInmat	Judet	Proprietar	Adresa
03 BBB	IS	Maftai Eduard	Nicolina 30
01 CCC	IS	Maftai Eduard	Nicolina 30
02 AAA	IS	Luca Marian	Primaverii 4
03 BBB	VS	Melinte Dan	Primaverii 17

CONTRAVENȚIE

Cod	Data	Cadru	Judet	NrInmat
143256	25/10/92	567	IS	02 AAA
987554	26/10/92	456	IS	02 AAA
987557	26/10/92	456	IS	03 BBB
630876	15/10/92	456	VS	03 BBB
539856	12/10/92	567	VS	03 BBB

Fig.2.6 Bază de date cu constrângeri de referință

Se definesc următoarele constrângeri de referință:

- între atributul **Cadru** al relației **CONTRAVENȚIE** și relația **CADRE**
- între attributele **NrInmat** și **Judet** ale relației **CONTRAVENȚIE** și relația **AUTOVEHICUL**, ordinea atributelor în cheia relației **AUTOVEHICUL** fiind **NrInmat**, **Judet**

Baza de date din figura 2.6 respectă ambele constrângeri

Probleme propuse

1. Se consideră baza de date din figura următoare

PACIENT

Cod	Nume	Prenume
A102	Popescu	Maria
B372	Popa	Mihai
B543	Ionescu	Ioana
B444	Ionescu	Radu
S555	Vasile	Ion

INTERNARE

Pacient	Intrare	Iesire	Sectie
A102	2/05/2004	9/05/2004	A
A102	2/12/2004	2/01/2005	A
S555	5/10/2004	3/12/2004	B
B444	1/12/2004	1/01/2005	B
S555	5/10/2004	1/11/2004	A

MEDIC

Numar	Nume	Prenume	Sectie
203	Dumitrescu	Mihai	A
574	Georgescu	Stefan	B
461	Barbu	Petre	B
530	Petrescu	George	C
405	Filipescu	George	A
501	Barbulescu	Stefan	A

SECTIE

Cod	Nume	Specialist
A	Chirurgie	203
B	Pediatric	574
C	Medicala	530

Puneți în evidență cheile și constrângerile de referință din baza de date prezentată. Precizați dacă respectivele constrângeri sunt satisfăcute de toate bazele de date care au aceeași schemă cu baza de date din figură. Precizați care atribute pot admite valori **null**.




2. Se consideră următoarele informații referitoare la managementul împrumuturilor dintr-o bibliotecă personală:

- proprietarul împrumută cărți prietenilor, care se înregistrează prin intermediul numelor (astfel încât să se evite repetițiile)
- o carte este referită prin titlul său (nu există două cărți cu același titlu)
- când clientul împrumută o carte, se înregistrează data împrumutului și proprietarul fixează data returnării

Definiți o schema relațională pentru reprezentarea informațiilor de mai sus, precizând domeniile adecvate pentru atribute și o instanță sub formă tabelară

Precizați cheia sau cheile relației

- 
3. Reprezentați prin intermediul unei relații sau a mai multora informațiile conținute în mersul trenurilor dintr-o stație: numărul trenului, ora sosirii, ora plecării, punctul de plecare, destinația finală, tipul de tren și opririle de pe parcurs

 4. Definiți o schemă a unei baze de date în care se organizează informațiile referitoare la o companie care are angajați (fiecare cu codul numeric personal, nume, prenume și data nașterii) și subsidiari (fiecare cu cod, ramură și director, care este angajat). Fiecare angajat lucrează pentru un subsidiar. Indicați cheile și constrângerile de referință ale schemei. Arătați o instanță a bazei de date și verificați dacă se respectă constrângerile.