BAZE DE DATE

Bibliografie

1. Abiteboul S., R.Hull, V.Vianu – *“Foundations of Databases”*, Addison-Wesley, Reading, MA, 1995.
2. Chandra, A. K., D. Harel – *“Structure and complexity of relational queries”*, J. Computer and Systems Sciences 25:1, pp.99-128.
3. Date, C.J. – *“Baze de date”*, traducere de Simona Preda si Titi Preda, editura PLUS, 2005.
4. Date, C.J., H. Darwen – *“A Guide to the SQL standard”*, Addison-Wesley, Reading, MA, 1997.

5. Date, C.J., Hugh Darwen – *“A Guide to the SQL Standard”*, Reading Mass., Addison-Wesley, 1997.
6. Date, C.J. – *“A Critique of the SQL Database Language”*, ACM SIGMOD Record 14, N.3, 1984.
7. Dubois, Paul – *“MySQL”* , traducere în limba română, Editura Teora, 2001.
8. England K. – *“Microsoft SQL Server 200 Performance Optimization and Tuning Handbook”*, Digital Press, 2001.
9. Fagin, R. – *“On an authorization mechanism”*, ACM Transactions on Database systems, 3:3, pp.310-319, 1978.
10. Fehily, C. – *“SQL visual quickstart guide”*, traducere în limba română, Editura B.I.C. ALL, 2004.

11. V.Felea: Baze de date relationale. Dependente. Editura Universitatii “Al.I.Cuza” Iasi, 1996
12. V.Felea, C.Matei si M.Balta: Interogarea bazelor de date, Aplicatii in Oracle si SQL Server, Editura MATRIXROM Bucuresti 2005.
13. V.Felea: Elemente ale implementarii modelului relational in Sisteme de Gestiune de Baze de Date”, Editura MATRIX-ROM, Bucuresti 2007.
14. Fotache Marin, Catalin Strimbei, Liviu Cretu – *“ORACLE 9i2, Ghidul dezvoltarii aplicatiilor profesionale”*, Editura Polirom, 2003.
15. Garcia-Molina Hector, Jeffrey D.Ullman, Jennifer Widom – *“Database Sistems: The complete book”*, Pearson Education International, 2002.

16. Gulutzan, P., T.Pelzer – *“SQL-99 Complete, Really, R&D Books”*, Lawrence, KA, 1999.
17. Gulutzan P., Peltzer T. – *“SQL Performance Tuning”*, Addison Wesley, 2002
18. Lorie, R.A., J.F. Nilsson – *“An Access Specification Language for a Relational Data Base System”*, IMB J. R&D, 23, 3, 1979.
19. Lorie, R.A., J.J. Daudenarde – *“SQL and its Applications”*, Englewoods Cliffs, N.J., Prentice_Hall,1991.
20. Melton, J., A.R.Simon – *“SQL-1999-Understanding Relational Components”*, San Francisco, Calif., Morgan-Kaufmann, 2002.

21. Melton, J., A.R.Simon – *“Understanding the New SQL: A Complete Guide”*, Morgan-Kaufmann, San Francisco, 1993.
22. Palinski, J.A. – *“Oracle SQL and PL/SQL Handbook”*, Addison-Wesley, 2003.
23. Perkins, J. şi Bryan Morgan – *“SQL fără profesor în 14 zile”*, traducere în limba română, Editura Teora, 1997.
24. Popescu, Ileana, *“Oracle8. – “Prelucrarea avansata a informatiei”*, Editura Tehnica, Bucuresti, 1999.
25. Riordan R. – *“Designing Relational Database Systems”*, Microsoft Press, 1999.

26. Rozenshtein, D.A., Abramovich and E.Birger – *“Optimizing Transact-SQL: Advanced Programming Techniques”*, Fremont, Calif., SQL Forum Press, 1995.
27. Sheldon R., Wilansky E. – *“MCSE Training Kit : Microsoft SQL Server 2000 Database Design and Implemention”*, Microsoft Press, 2001
28. Smith, J.M., P.Y.Chang – *“Optimizing the performance of a relational algebra database interface”*, Comm. ACM 18:10 (1975), pp.568-579.
29. Stonebraker M., J.M. Hellerstein (eds) – ”Reading in Databases Systems”, Morgan-Kaufmann, Sab Francisco, 1998.

30. Ullman J.D. – *“Principles of Database and Knowledge-Base Systems”*, Volume I, Computer Science Press, New York, 1988.
31. Ullman J.D. – *“Principles of Database and Knowledge-Base Systems”*, Volume II, Computer Science Press, New York, 1989.
32. Urman, S. – *“Oracle 9i PL/SQL Programming”, Osborne/McGraw-Hill, 2002*.
33. Welling, L., Laura Thomson – “Dezvoltarea de aplicații WEB cu PHP si MySQL”, traducere în limba română, Editura Teora, 2004.

CAPITOLUL I
ELEMENTE ALE MODELULUI RELAȚIONAL

Considerăm o mulțime nevidă, finită de simboluri, notată cu **U**. Numim elementele lui **U** atribute.

Fie **U** = $\{A_1, A_2, \ldots, A_n\}$. Fiecărui atribut A_i îi vom asocia o mulțime nevidă de valori notată dom(A_i), care va fi numită domeniul valorilor atributului A_i. Această mulțime dom(A_i) este numită şi mulțime valorilor posibile pentru A_i, $1 \leq i \leq n$.

Vom numi uplu peste **U** o aplicație, notată φ, φ : **U** $\rightarrow \bigcup_{1 \leq i \leq n}$ dom(A_i), astfel încât $\varphi(A_i) \in$ dom(A_i), $1 \leq i \leq n$.

Pentru fiecare uplu φ peste **U** putem să-i asociem mulțimea $\{A_1:V_1, A_2 : V_2, \ldots, A_n : V_n\}$, unde $V_i = \varphi(A_i)$, $1 \leq i \leq n$.

Invers, dacă este dată mulțimea $\{A_1:V_1, A_2 : V_2, \ldots, A_n : V_n\}$, cu $V_i \in dom(A_i)$, $1 \leq i \leq n$, atunci putem defini uplul φ peste **U** prin $\varphi(A_i) = V_i$, $1 \leq i \leq n$. Deci există o corespondență biunivocă între mulțimea de uple peste **U** şi mulțimea mulțimilor de forma considerată. Mai mult, dacă ordonăm mulțimea de atribute ale lui **U** sub forma: **U** $= (A_1, A_2, \ldots, A_n)$, atunci în locul mulțimii $\{A_1:V_1, A_2 : V_2, \ldots, A_n : V_n\}$ se poate considera numai vectorul $(V_1, V_2, \ldots,V_n)$.

În sisteme de gestiune de baze de date (FOX, DBASE, ORACLE, etc.) se lucrează cu astfel de ordonări. Vom privi uplele atât ca aplicații, cât şi ca vectori, considerând **U** ordonată.

(i) Relatia ca set.

O relație peste **U**, notată r, este o mulțime de uple peste **U**. În cazul în care mulțimea de uple este vidă, spunem că relația este vidă. În cazul în care mulțimea de uple este finită, spunem că relația este finită.

Rezultă că într-o astfel de relație nu putem avea două uple identice.

Ca operații de actualizare a unei relații putem avea adăugarea de noi uple, ştergerea unor uple, modificarea de uple.

În prelucrarea unei relații, mulțimea de uple variază în timp, ceea ce este constant în timp este structura sa, adică numele relației, împreună cu mulțimea de atribute. Să notăm această structură prin R(**U**) sau R[**U**], sau R(A_1, A_2, …, A_n) sau R[A_1, A_2, …, A_n], unde R este un simbol numit numele relației, iar **U** sau {A_1, A_2, …, A_n} este mulțimea de atribute corespunzătoare.

R(**U**) poartă denumirea şi de schemă de relație.

În cazul în care r este o relație finită peste **U** şi **U** se ordonează sub forma (A_1, A_2, …, A_n), atunci rezultă o reprezentare a lui r sub forma unui tablou:

	A_1	A_2	…	A_n
	V_1	V_{12}	…	V_{1n}
	1			
r	……….			
	V_{h1}	V_{h2}	…	V_{hn}

unde $(V_{i1}, V_{i2}, \dots, V_{in})$ constituie un uplu din r, $1 \le i \le h$. Avem $V_{ij} \in dom(A_j)$, $1 \le j \le n$, $1 \le i \le h$.

Din acest motiv, în majoritatea sistemelor de gestiune de baze de date, structurile de date ce memorează aceste relații se numesc tabele, iar uplele se numesc linii.

O schemă de baze de date, notată D, este o mulțime finită de scheme de relație: $D = \{R_1[U_1], \ldots, R_p[U_p]\}$, unde $R_j[U_j]$ este o schemă de relație, $1 \leq j \leq p$.

O bază de date peste schema de baze de date D, este o aplicație ce asociază fiecărei scheme de relație $R_j[U_j]$ o relație r_j peste U_j, $1 \leq j \leq p$. Dacă D se consideră ordonată sub forma $(R_1[U_1], \ldots, R_p[U_p])$, atunci baza de date este un vector de relații, notat $\bar{r} = (r_1, r_2, .., r_p)$.

Structura unei relații se numeşte partea sa intensională, iar uplele relației poartă denumirea de partea sa extensională.

Operații referitoare la relații:

a) Proiecția unei relații.

Fie t un uplu peste R[U] şi X o submulțime a lui U. Proiecția lui t relativ la X, notată prin t[X] este restricția lui t (ca uplu) la submulțimea X. Aici t este considerat ca aplicație. Dacă X este mulțimea vidă, atunci t[X] il vom considera uplul vid. Dacă U este ordonată sub forma $(A_1,A_2,\ldots,A_n)$, iar în această ordonare $X=(A_{i1},A_{i2},\ldots,A_{ik})$, cu $i_1<i_2<\ldots<i_k$ şi $X=(V_1,V_2,\ldots,V_n)$, atunci $t[X] =(V_{i1},V_{i2},\ldots,V_{ik})$. Dacă r este o relație peste U, atunci proiecția lui r relativă la X, notată r[X] va fi definita prin: $r[X]=\{t[X] \ / \ t\in r\}$.

b)Reuniunea a două relații.

Fie r_1 şi r_2 relații peste R[U]. Atunci $r_1\cup r_2$ notează reuniunea celor două şi se defineşte ca:

$r_1 \cup r_2 = \{t \,/\, t \text{ uplu}, t \in r_1 \text{ sau } t \in r_2\}$

c) Diferența a două relații.

Fie r_1 şi r_2 relații peste R[U]. Diferența lor, notată r_1-r_2 se defineşte astfel:

$r_1 - r_2 = \{t \,/\, t \text{ uplu}, t \in r_1 \text{ şi } t \notin r_2\}$

d) Unire (join) oarecare.

Fie relațiile r_i definite peste $R_i[U_i]$, i=1,2 $U_1 \cap U_2 = \emptyset$.
Fie $A_{\alpha 1}$, $A_{\alpha 2}, \ldots, A_{\alpha q} \in U_1$, $B_{\beta 1}$, $B_{\beta 2}, \ldots, B_{\beta q} \in U_2$ şi Θ_i un operator de comparare între $dom(A_{\alpha i})$ şi $dom(B_{\beta i})$, $1 \leq i \leq q$. Sintactic Θ_i este unul din operatorii =, <, >, <=, >=, <>. Deci Θ_i este o relație binară pe $dom(A_{\alpha i}) \times dom(B_{\beta i})$.

Rezultă că cele două domenii conțin valori comparabile prin Θ_i. Considerăm $\Theta=(A_{\alpha 1}\Theta_1 B_{\beta 1})\wedge\ldots\wedge(A_{\alpha q}\Theta_q B_{\beta q})$, unde semnul "$\wedge$" reprezintă conjuncția.

Join-ul oarecare între r_1 şi r_2 prin expresia Θ se notează prin $r_1 \underset{\Theta}{\overset{*}{—}} r_2$ şi se defineşte prin:

$r_1 \underset{\Theta}{\overset{*}{—}} r_2=\{t \ / \ t$ uplu peste $U_1\cup U_2$, $t[U_i]\in r_i$, $i=1,2$ şi $t[A_{\alpha j}]\Theta_j t[B_{\beta j}]$, $j=1,q\}$

Se pot defini expresii de tip Θ mai generale decit cea de sus, folosind parantezele şi operatorii de conjuncție şi disjuncție ($\wedge$, respectiv $\vee$) astfel:

- o expresie elementară este de forma $A\Theta B$, unde A şi B sunt atribute, iar Θ este operator de comparare.
- o expresie join se defineşte astfel:

1. Dacă e este o expresie elementară, atunci e şi (e) sunt expresii join.
2. Dacă e_1 şi e_2 sunt expresii join, atunci $e_1 \wedge e_2$, $e_1 \vee e_2$, $(e_1 \wedge e_2)$, $(e_1 \vee e_2)$ sunt expresii join.
3. Orice expresie join se obține numai prin regulile 1 şi 2.

Dacă Θ este o expresie join, atunci se defineşte faptul că (t_1,t_2) satisface Θ în mod recursiv, $t_1 \in r_1$, $t_2 \in r_2$.

1. (t_1,t_2) satisface $A\Theta B$ dacă $t_1[A]\Theta t_2[B]$.
2. (t_1,t_2) satisface $e_1 \wedge e_2$ şi $(e_1 \wedge e_2)$ dacă (t_1,t_2) satisface e_1 şi e_2.
3. (t_1,t_2) satisface $e_1 \vee e_2$ şi $(e_1 \vee e_2)$, dacă (t_1,t_2) satisface e_1 sau e_2.

În acest caz joinul oarecare se defineşte prin:
$r_1 \underset{\Theta}{\overset{*}{-}} r_2=\{(t_1,t_2)$ / (t_1,t_2) uplu peste $U_1 \cup U_2$, $t_i \in r_i$, $i=1,2$ şi (t_1,t_2) satisface $\Theta\}$

e) Produsul cartezian.
Fie r_i relații definite peste $R_i[U_i]$, $i=1,2$ şi $U_1 \cap U_2=\varnothing$.
Produsul cartezian al relațiilor r_1 şi r_2 se noteaza prin $r_1 \times r_2$ şi se defineşte prin: $r_1 \times r_2=\{t$ / t uplu peste $U_1 \cup U_2$, $t[U_i] \in r_i$, $i=1,2\}$

Obs. Produsul cartezian este un join oarecare cu expresia join Θ=TRUE.

f) Unirea naturală (join natural)

Fie r_i relații peste $R_i[U_i]$, $i=1,2$. Se numeşte join natural sau unire a celor două relații, notat r_1*r_2 o relație peste $U_1 \cup U_2$ definită prin:

$r_1*r_2=\{t \,/\, t \text{ uplu peste } U_1 \cup U_2, t[U_i] \in r_i, i=1,2\}$

Fie R un nume pentru relația peste $U_1 \cup U_2$. Deci relația r_1*r_2 va fi considerată peste schema $R[U_1 \cup U_2]$. Considerăm un exemplu pentru operația de join natural. Fie $R_1[ABCD]$, $R_2[CDE]$ şi relațiile r_1 şi r_2 date astfel:

r_1

	A	B	C	D	
	1	0	0	0	t_1
	1	1	0	0	t_2
	0	1	0	1	t_3
	0	0	0	1	t_4
	1	1	1	1	t_5

r_2

	C	D	E	
	0	0	0	V_1
	1	1	1	V_2
	1	1	0	V_3
	1	0	0	V_4
	1	0	1	V_5

Joinul natural va fi:

$r_1 * r_2 =$

A	B	C	D	E
1	0	0	0	0
1	1	0	0	0
1	1	1	1	1
1	1	1	1	0

Se observă că uplele t_3 şi t_4 din r_1 nu contribuie la construcția joinului, deoarece nu există uple $V_j \in r_2$, astfel încît $t_3[CD]=V_j[CD]$, şi nu există uple $V_j \in r_2$, astfel încît $t_4[CD]=V_j[CD]$. La fel uplele V_4 şi V_5 din r_2 nu contribuie la formarea joinului, deoarece $V_4[CD]=V_5[CD] \notin r_1[CD]$.
Să remarcăm faptul că perechea (t_i, V_j), $t_i \in r_1$, $V_j \in r_2$ contribuie la formarea joinului, dacă $t_i[CD]=V_j[CD]$.
În acest caz, va rezulta un uplu în joinul natural, notat w, definit prin: $w[U_1]=t_i[U_1]$, $w[U_2-U_1]=V_j[U_2-U_1]$.

Se observă uşor următoarele relații:

1. $(r_1 * r_2)[U_1] \subseteq r_1$, $(r_1 * r_2)[U_2] \subseteq r_2$.
2. Dacă $r_1' = \{t_1 \ / \ t_1 \in r_1, \exists t_2 \in r_2, \text{ a.î., } t_1[U_1 \cap U_2]=t_2[U_1 \cap U_2]\}$
   şi $r_2' = \{t_2 \ / \ t_2 \in r_2, \exists t_1 \in r_1, \text{ a.î., } t_1[U_1 \cap U_2]=t_2[U_1 \cap U_2]\}$

și $r_1'' = r_1 - r_1'$, $r_2'' = r_2 - r_2'$, atunci

$$r_1 * r_2 = r_1' * r_2',\ (r_1 * r_2)[U_1] = r_1',\ (r_1 * r_2)[U_2] = r_2'.$$

3. Dacă $\overline{r_1} \subseteq r_1$ și $\overline{r_2} \subseteq r_2$ satisfac relația $\overline{r_1} * \overline{r_2} = r_1 * r_2$, atunci $r_1' \subseteq \overline{r_1}$ și $r_1' \subseteq \overline{r_2}$, adică relațiile r_1', r_2' sunt minimale în r_1 respectiv r_2 cu proprietatea $r_1' * r_2' = r_1 * r_2$. În cazul în care $U_1 \cap U_2 = \varnothing$, atunci este clar că joinul natural este tocmai produsul cartezian $r_1 \times r_2$.

Operația de join natural se poate extinde la mai multe relații.

Fie r_i peste $R_i[U_i]$, i=1,h. Joinul natural al acestora se notează prin

$r_1 * r_2 * \ldots * r_h$ sau $*\langle r_i, i=1,h \rangle$, sau $* \langle r_i, i=1,h \rangle$ și se definește prin:

$r_1 * r_2 * \ldots * r_h = \{t \,/\, t$ uplu peste $U_1 \cup \ldots \cup U_h$, a.î. $t[U_i] \in r_i$, $i=1,h\}$

Este clar că operația de join natural nu depinde de ordinea considerării relațiilor r_i, $i=1,h$.

Dacă notăm cu $*_2$ operația de join natural între două relații şi cu $*_h$ operația de join natural între h relații, apare următoarea problemă:

Problema 1. Operația $*_h$ se poate exprima cu ajutorul operației $*_2$?

Răspunsul la problema 1 este NU. Demonstrați.

g) Selecția.

Fie r o relație peste R[U].

Definim întâi expresia elementară de selecție prin:

> $A\Theta B$ sau $A\Theta c$ sau $c\Theta B$, unde $A,B \in U$, iar c este o constantă, comparabilă cu elementele din dom(A) şi dom(B).

Dacă e_1 şi e_2 sunt expresii elementare de selecție, atunci (e_1), $e_1 \wedge e_2$, $e_1 \vee e_2$, $(e_1 \wedge e_2)$, $(e_1 \vee e_2)$ este o expresie de selecție. Orice expresie de selecție se obține numai prin regulile de mai sus. Fie E o expresie de selecție.

Precizăm cazul în care uplul t satisface E.

Dacă $E \equiv A\Theta B$, atunci t satisface E, dacă $t[A]\Theta t[B]$.

Dacă $E \equiv A\Theta c$, atunci t satisface E, dacă $t[A]\Theta c$.

Dacă $E \equiv c\Theta B$, atunci t satisface E, dacă t satisface $c\Theta t[B]$.

Dacă $E \equiv e_1 \wedge e_2$, atunci t satisface E, dacă t satisface e_1 şi e_2.
Dacă $E \equiv e_1 \vee e_2$, atunci t satisface E, dacă t satisface e_1 sau t satisface e_2.
Fie F o expresie de selecție.
Selecția se notează prin $\sigma_F(\underline{r})$ şi se defineşte prin:
$\sigma_F(\underline{r})=\{t$ / t uplu peste R[U], t satisface F$\}$.

h) Intersecția relațiilor.
Fie r_i relații peste $R_i[U_i]$, $i=1,2$ şi $U_1=U_2$.
Atunci intersecția se defineşte prin:

$$r_1 \cap r_2=\{t \text{ / t uplu peste } U_1, \text{ a.î. } t \in r_1 \text{ şi } t \in r_2\}$$

Se arată că o parte din operațiile cu relații se exprimă cu ajutorul altor operații. Mulțimea de operații: reuniune, diferență, produs cartezian, proiecție şi selecție se arată a fi minimală cu proprietatea că celelalte operații se exprimă cu ajutorul acestora.

(ii) Relatia ca multiset

Fie schema R(U). O relatie ca multiset peste R(U) este un multiset de uple peste U.
$r=\{u_1:m_1,\ldots, u_h:m_h\}$.
Orice relatie de tip set este multiset. Operatii:

a) Proiectia. $r=\{u_1:m_1,\ldots, u_h:m_h\}$ peste R(U) si X <=U, r[X]
$v_i=u_i[X]$, consideram $v_1:m_1,\ldots, v_h:m_h$, v_i pot fie egale. Se grupeaza toate egale si se aduna

multiplicitatile pentru cele egale.

A	B	C	D	m_i
0	0	0	0	3
0	0	1	0	2
0	0	0	1	4
1	1	0	1	2
1	1	1	2	1

X=AB,

$v_1[X]=(0\ \ 0)\ :\ 3$

$v_2[X]=(0\ \ 0)\ :\ 2$

$v_3[X]=(0\ \ 0)\ :\ 4$ $\quad$ $r[X]=\{(0\ \ 0):9,\ (1\ \ 1):3\}$.

$v_4[X]=(1 \quad 1) \ : \ 2$

$v_5[X]=(1 \quad 1) \ : \ 1$

----

b) Reuniunea a doua relatii:

$r_1=\{u_1:m_1, \ldots, u_h:m_h\}$, $r_2=\{v_1:n_1,\ldots, v_p:n_p\}$, peste R(U).

$r_1 \cup r_2 =\{w_1 : q_1,\ldots, w_l: q_l \}$, $u_1,\ldots, u_h$, $v_1,\ldots, v_p$,

For i=1 to h

1) If exista v_j astfel incat $u_i= v_j$, atunci $w_\alpha= u_i$, $q_\alpha=m_i+n_j$,
2) altfel $w_\alpha= u_i$, $q_\alpha=m_i$

endfor

For j=1 to p

3) If $v_j \neq u_i$, for all i=1,h, atunci $w_\alpha = v_j$, $q_\alpha = n_j$

Endfor

Exemplu: $r_1 = \{(0\ \ 0){:}2, (0\ \ 1){:}3, (1\ \ 0){:}2\}$,

$r_2 = \{(0\ \ 0){:}1, (1\ \ 1){:}3\}$,

$r_1 \cup r_2 = \{(0\ \ 0){:}3, (0\ \ 1){:}3, (1\ \ 0){:}2, (1\ \ 1){:}3\}$

c) Diferenta a 2 relatii multiset.

$r_1 = \{u_1{:}m_1, \ldots, u_h{:}m_h\}$, $r_2 = \{v_1{:}n_1, \ldots, v_p{:}n_p\}$, peste R(U).

$r_1 - r_2 = \{w_1 : q_1, \ldots, w_l : q_l\}$,

For i=1 to h

4) If exista v_j astfel incat $u_i = v_j$, si $m_i > n_j$, atunci $w_\alpha = u_i$, $q_\alpha = m_i - n_j$, altfel

5) If exista v_j astfel incat $u_i = v_j$, si $m_i <= n_j$, atunci continue,

6) altfel $w_\alpha = u_i$, $q_\alpha = m_i$

endfor

d) Joinul oarecare.

$r_1=\{u_1:m_1, \ldots, u_h:m_h\}$, peste $R_1(U_1)$, $r_2=\{v_1:n_1,\ldots, v_p:n_p\}$, peste $R_2(U_2)$, $U_1 \cap U_2 = \Phi$,
$r_1 \,\theta\, r_2 = \{w : q \mid w$ peste $U_1 \cup U_2$, exista i si j, $w[U_1] = u_i$, $w[U_2] = v_j$, w satisface $\theta\}$.

e) Produsul cartezian.

$r_1=\{u_1:m_1, \ldots, u_h:m_h\}$, peste $R_1(U_1)$, $r_2=\{v_1:n_1,\ldots, v_p:n_p\}$, peste $R_2(U_2)$, $U_1 \cap U_2 = \Phi$,

$r_1 \times r_2 = \{w : q \mid w$ peste $U_1 \cup U_2$, exista i si j,
$w[U_1] = u_i$, $w[U_2] = v_j\}$.

f) Join natural.

(C1) $u_i [U_1 \cap U_2] = v_j[U_1 \cap U_2]$,
(C2) $w[U_1] = u_i [U_1]$, $w[U_2 - U_1] = v_j[U_2 - U_1]$, $q = m_i * n_j$.

r= Φ
For i=1,h
For j=1,p
If perechea (u_i, v_j) satisface (C1), then
calculeaza w:q conform lui (C2) si adauga w:q la r,
endif
endfor

endfor

Exemplu:

A	B	C	D	m_i	
0	1	0	0	3	u_1
0	0	1	0	2	u_2
0	1	1	1	2	u_3

C	D	E	n_j	
0	0	1	2	v_1
1	0	0	3	v_2
0	1	1	1	v_3
0	0	0	3	v_4

	A	B	C	D	E	q_l
(u_1,v_1)->	0	1	0	0	1	6
$(u_1,v4)$->	0	1	0	0	0	9
(u_2,v_2)->	0	0	1	0	0	6

Sunt valabile proprietatile joinului natural pentru cazul set?

Dependente functionale

1 Definitie. Proprietati. Sisteme de reguli de inferenta.

Fie $X, Y \subseteq U$. Vom nota sintactic o dependenta functionala prin $X \rightarrow Y$.

Semantic, vom spune ca o relatie r peste U satisface dependenta functionala

$X \rightarrow Y$ daca: $(\forall t1, t2)(t1, t2 \in r)[t1[X] = t2[X] \Rightarrow t1[Y] = t2[Y]]$.

Daca $X = \emptyset$,atunci spunem ca r satisface $\emptyset \to Y$ daca

$(\forall t1, t2)(t1, t2 \in r)[t1[Y] = t2[Y]]$, altfel

spus, r[Y] consta dintr-un singur element.

Daca $Y = \emptyset$, atunci consideram ca

orice relatie r peste U satisface dependenta functionala $X \to \emptyset$. Daca r satisface dependenta functionala $X \to Y$, atunci exista o functie : $r[X] \to r[Y]$ definita

prin: $F(t) = t'[Y]$, unde $t' \in r$, si $t'[X] = t \in r[X]$.

Daca r satisface $X \rightarrow Y$ se mai spune ca X determina functional pe Y ın r.

Proprietati ale dependentelor functionale:

FD1. (Reflexivitate) Daca $Y \subseteq X$, atunci r satisface $X \rightarrow Y$ pentru orice relatie r peste U.

FD2 (Extensie) Daca r satisface $X \rightarrow Y$, si $Z \subseteq W$, atunci r satisface $XW \rightarrow YZ$.

FD3 (Tranzitivitate) Daca r satisface $X \rightarrow Y$, si $Y \rightarrow Z$, atunci r satisface $X \rightarrow Z$.

FD4 (Pseudotranzitivitate) Daca r satisface $X \to Y$, si $YW \to Z$, atunci r satisface $XW \to Z$.

FD5 (Uniune) Daca r satisface $X \to Y$ si $X \to Z$, atunci r satisface $X \to Y Z$.

FD6 (Descompunere) Daca r satisface $X \to Y Z$, atunci r satisface $X \to Y$ si $X \to Z$.

FD7 (Proiectabilitate) Daca r peste U satisface $X \to Y$ si $X \subset Z \subseteq U$, atunci r[Z] satisface $X \to Y \cap Z$.

FD8 (Proiectabilitate inversa) Daca $X \to Y$ este satisfacuta de o proiectie a lui r, atunci $X \to Y$ este satisfacuta de r.

Definitia 2.

Fie Σ o multime de dependente functionale peste U. Spunem ca $X \to Y$ este consecinta din Σ, daca orice relatie ce satisface toate dependentele lui Σ va

satisface si $X \to Y$.

Vom nota aceasta situatie prin $\Sigma \mid\!- X \to Y$.

Deci Σ |- $X \rightarrow Y$ daca pentru $(\forall r)[\forall \alpha \in \Sigma$, r satisface $\alpha \rightsquigarrow$ r satisface $X \rightarrow Y]$.

Fie $\Sigma^* = \{X \rightarrow Y \mid \Sigma \vdash X \rightarrow Y\}$. Fie Σ_1 o multime de dependente functionale.

Spunem ca Σ_1 constituie o acoperire pentru Σ^*, daca $\Sigma^*_1 = \Sigma^*$.

Propozitia 1.1 Pentru orice multime Σ de dependente functionale exista o acoperire Σ_1 pentru Σ^*, astfel ıncat toate dependentele din Σ_1 sunt de forma

$X \rightarrow A$, A fiind un atribut din U.

Demonstratie. Pentru fiecare $X \rightarrow Y \in \Sigma$, cu $Y = B1B2 \ldots Bh$, consideram

$X \rightarrow Bj$ incluse ın Σ_1, $j = 1, h$. Dupa proprietatile FD5 si FD6 avem: r satisface $X \rightarrow Y$ daca si numai daca r satisface $X \rightarrow Bj$, $j = 1, h$. Dupa aceleasi proprietati si din modul de definire a lui $\Sigma 1$ avem: r satisface α, $\forall \alpha \in \Sigma$ daca si numai daca r satisface $\alpha 1$, $\forall \alpha 1 \in \Sigma_1$. Aceasta din urma conduce la $\Sigma \vdash X \rightarrow Y$ daca si numai daca $\Sigma_1 \vdash X \rightarrow Y$, ceea ce ınseamna $\Sigma^*_1 = \Sigma^*$.

Propozitia 1.2 $\Sigma \vdash X \rightarrow Y$ daca si numai daca $\Sigma \vdash X \rightarrow Bj$ pentru j = 1, h, unde Y = B1 . . .Bh.

Justificarea propozitiei rezulta imediat din FD5 si FD6. Din cele doua propozitii de mai sus, rezulta ca studiul dependentelor functionale, din punct de vedere al relatiei de “consecinta”, se reduce la studiul acestei “consecinte” pe multimea dependentelor functionale ın care membrul doi are un singur atribut.

Reguli de inferenta.

In continuare vom considera reguli formale de deducere a noi dependente functionale, pornind de la o multime data Σ.

Fie $\mathscr{R}$ o multime de reguli formale de deducere pentru dependente functionale si Σ o multime de dependente functionale. Spunem ca $X \rightarrow Y$ are o demonstratie ın Σ utilizand regulile $\mathscr{R}$, si vom nota $\Sigma|_{\mathscr{R}} X \rightarrow Y$, daca exista sirul $\alpha_1, \alpha_2, \ldots, \alpha_n$, astfel ıncat:

a) $\alpha_n = X \rightarrow Y$, si

b) pentru orice $i = 1, n$, $\alpha_i \in \Sigma$ sau exista o regula din $\mathscr{R}$ de forma

$$\frac{\alpha_{j1}, \alpha_{j2}, \ldots, \alpha_{jk}}{\alpha_i},$$

unde j1, j2, . . . , jk < i (adica α_i se obtine utilizand o regula din $\mathscr{R}$, cu premizele existente ın sir ınainte de α_i). Corespunzator proprietatilor FD1–FD6 se pot defini reguli formale de deducere a dependentelor functionale:

$$\text{FD1f}: \frac{Y \subseteq X}{X \rightarrow Y}$$

$$\text{FD2f}: \frac{X \rightarrow Y,\ Z \subseteq W}{}$$

$$XW \to Y\,Z$$

FD3f: $X \to Y,\ Y \to Z$

$$X \to Z$$

FD4f : $X \to Y,\ YW \to Z$

$$XW \to Z$$

FD5f : $X \rightarrow Y, X \rightarrow Z$

$X \rightarrow Y Z$

FD6f : $X \rightarrow Y Z$

$X \rightarrow Y$

Armstrong a definit urmatoarele reguli de inferenta (numite axiomele lui Armstrong):

(A1) : $$\frac{\qquad}{A1 \;.\;.\;.Am \rightarrow Ai}\;,\; i = 1,m$$

(A21) : $$\frac{A1 \;.\;.\;.Am \rightarrow B1 \;.\;.\;.Br}{A1 \;.\;.\;.Am \rightarrow Bj}\;,\; j = 1,\ r$$

(A22) : $$\frac{A1 \;.\;.\;.Am \rightarrow Bj\;,\; j = 1,\ r}{A1 \;.\;.\;.Am \rightarrow B1 \;.\;.\;.Br}$$

(A3) : $A1 \ldots Am \rightarrow B1 \ldots Br,\ B1 \ldots Br \rightarrow C1 \ldots Cp$

$$\frac{A1 \ldots Am \rightarrow B1 \ldots Br,\quad B1 \ldots Br \rightarrow C1 \ldots Cp}{A1 \ldots Am \rightarrow C1 \ldots Cp}$$

unde Ai,Bj ,Ck sunt atribute.

Observatia 1.1 Regula A3 este ın fond FD3f (tranzitivitatea).

Definitie. Regula

$$\frac{\alpha_{j1}, \ldots, \alpha_{jh}}{\alpha_i}$$

se exprima cu ajutorul regulilor sistemului $\mathscr{R}$, daca:

$\{ \alpha_{j1}, \ldots, \alpha_{jh}\} \mid\!-_{\mathscr{R}} \alpha_i$

Propozitia 1.3 Regulile FD4f, FD5f, FD6f se exprima cu ajutorul regulilor FD1f, FD2f, FD3f.

In adevar, fie $X \to Y$ si $YW \to Z$ date. Aplicam FD2f pentru $X \to Y$ si $W \subseteq W$, si obtinem $XW \to YW$. Aceasta din urma si $YW \to Z$ si FD3f conduc la $XW \to Z$.

Fie date $X \to Y$ si $X \to Z$. Aplicand FD2f pentru $X \to Y$ si $X \subseteq X$ obtinem $X \to XY$; de asemenea aplicam FD2f pentru $X \to Z$ si $Y \subseteq Y$ si obtinem $XY \to Y Z$. Prin tranzitivitate din $X \to XY$ si $XY \to Y Z$ obtinem

$X \to Y Z$.

Fie $X \to Y Z$. Dupa FD1f obtinem $Y Z \to Y$ si $Y Z \to Z$, aplicand FD3f se obtine $X \to Y$ si $X \to Z$.

Sa notam prin $\Sigma^{+}_{\mathscr{R}} = \{X \to Y \mid \Sigma \vdash_{\mathscr{R}} X \to Y\}$.

Fie $\mathscr{R}_1$ = {FD1f, FD2f, FD3f}, $\mathscr{R}_2 = \mathscr{R}_1 \cup$ {FD4f, FD5f, FD6f}, $\mathscr{R}_A$ = {A1,A21, A22, A3}.

Observatia 1.2. $\Sigma^{+}{}_{\mathscr{R}1} = \Sigma^{+}{}_{\mathscr{R}2}$ avand ın vedere propozitia 1.3.

Propozitia 1.4. Regulile din $\mathscr{R}_1$ se exprima prin cele din $\mathscr{R}_A$ si invers.

Demonstratie. Fie X = A1 . . .Am , si Y = $A_{i1} \ldots A_{ik}$, {i1, . . . , ik} ⊆ {1, 2, . . . ,m}. Aplicand regula A1 obtinem $X \rightarrow A_{i1}, \ldots, X \rightarrow A_{ik}$, apoi considerand A22 obtinem $X \rightarrow A_{i1} \ldots A_{ik}$, adica $X \rightarrow Y$. Deci FD1f se exprima prin regulile din $\mathscr{R}_A$.

Fie $X \rightarrow Y$ si $Z \subseteq W$ date. Evidentiem atributele din fiecare:

$X = A1 \ldots Am$, $Y = B1 \ldots Bp$, $W = C1 \ldots Cq$,

$Z = C_{i1} \ldots C_{ik}$ cu $\{i1, \ldots, ik\} \subseteq \{1, 2, \ldots, q\}$.

Din $X \rightarrow Y$ prin A21 obtinem $X \rightarrow Bj$, $j = 1, p$.

Prin A1 , si A22 obtinem $XW \rightarrow X$. Din $X \rightarrow Bj$, si $XW \rightarrow X$ obtinem $XW \rightarrow Bj$, $j = 1, p$ utilizand A3.

Prin A1 obtinem $XW \rightarrow C_{il}$, $l = 1, k$.

Aplicand A22 pentru $XW \to Bj$, $j = 1, p$ si $XW \to C_{il}$, $l = 1, k$ obtinem $XW \to Y Z$.

Regula FD3f este exact A3.

Regula A1 se exprima numai prin FD1f. Regula A21 se exprima aplicand regula FD6f de r−1 ori, iar FD6f se exprima cu ajutorul celor din $\mathscr{R}_1$,

dupa propozitia 1.3.

Daca $\alpha_{j1}, \ldots, \alpha_{jh}$

$$\overline{\qquad\qquad\qquad\qquad} \in \mathscr{R}_1$$

$$\alpha_i$$

si se exprima cu ajutorul regulilor lui R2, notam

prin $trans_{\mathscr{R}2}$ (α_{j1} , . . . , α_{jh}, α_i) sirul de dependente ce se obtine pornind de la premizele α_{j1} , . . . , α_{jh} si aplicand regulile lui R2 pentru obtinerea lui α_i.

Propozitia 1.5 Fie $\mathscr{R}'_1$, si $\mathscr{R}'_2$ doua multimi de reguli, astfel ıncat $\mathscr{R}'_1$ se exprima prin $\mathscr{R}'_2$, si invers. Atunci

$\Sigma^{+}_{\mathscr{R}'1} = \Sigma^{+}_{\mathscr{R}'2}$ pentru orice multime Σ de dependente functionale.

Demonstratie. Fie $X \rightarrow Y \in \Sigma^{+}_{\mathscr{R}'1}$. Sa aratam ca $X \rightarrow Y \in \Sigma^{+}_{\mathscr{R}'2}$ (*)

Exista sirul $\alpha_1, \alpha_2, \ldots, \alpha_n = X \rightarrow Y$, astfel ıncat pentru orice i, $i = 1, n$ avem:

a) $\alpha_i \in \Sigma$ sau

b) exista $\sigma = \alpha_{j1}, \ldots, \alpha_{jh}$

---------------------- $\in \mathscr{R}'_1$

α_i

cu j1, j2, . . . , jh < i

Demonstratia o realizam prin inductie dupa n.

Daca n = 1, atunci putem avea:

c) $\alpha_1 \in \Sigma$, si deci $\alpha_1 \in \Sigma^+_{\mathscr{R}'2}$ sau

d) -------------- $\in \mathscr{R}'_1$

α_1

In cazul d) sirul $trans_{\mathscr{R}'2}(, \alpha_1)$ constituie o demonstratie pentru α_1 ın Σ, utilizand $\mathscr{R}'_2$, adica $\alpha_1 \in \Sigma^+_{\mathscr{R}'2}$.

Presupunem afirmatia (*) valabila ın cazul ın care $X \to Y$ are demonstratii ın Σ utilizand reguli din $\mathscr{R}'_1$ si lungimea demonstratiei este mai mica sau egala cu n.

Fie acum $X \to Y \in \Sigma^+_{\mathscr{R}'1}$ cu lungimea demonstratiei egala cu n + 1. Exista sirul $\alpha_1, \alpha_2, . . . , \alpha_n, \alpha_{n+1} = X \to Y$, astfel ıncat pentru orice i, $1 \leq i \leq n + 1$, avem:

a1) $\alpha_i \in \Sigma$ sau

b1) exista β, instantierea unei reguli din $\mathscr{R}'_1$

$$\beta = \frac{\alpha_{j1}, \ldots \alpha_{jh}}{\alpha_i} \in \mathscr{R}'_1$$

cu $j1, \ldots, jh < i$

Daca $\alpha_{n+1} \in \Sigma$, atunci $\alpha_{n+1} = X \rightarrow Y \in \Sigma^{+}_{\mathscr{R}'2}$.

Daca α_{n+1} se obtine prin b1), atunci exista

$\alpha_{j1}, \ldots \alpha_{jh}$

$$\frac{\qquad\qquad\qquad}{\alpha_{n+1}} \in \mathscr{R}'_1$$

cu $j1, \ldots, jh < n + 1$.

Dupa ipoteza inductiei avem: $\alpha_{j1}, \ldots, \alpha_{jh} \in \Sigma^+_{\mathscr{R}'2}$.

Fie $dem_{\mathscr{R}'2}(\alpha_{ji})$ o demonstratie pentru α_{ji} ın Σ, utilizand regulile lui $\mathscr{R}'_2$, $i = 1, h$.

Atunci sirul: $dem_{\mathscr{R}'2}(\alpha_{j1}), \ldots, dem_{\mathscr{R}'2}(\alpha_{jh}), trans_{\mathscr{R}'2}(\alpha_{j1}, \ldots, \alpha_{jh}, \alpha_{n+1})$

constituie o demonstratie pentru α_{n+1} ın Σ, utilizand regulile din $\mathscr{R}'_2$. Deci $\alpha_{n+1} \in \Sigma^+_{\mathscr{R}'2}$.

Relatia (*) ınseamna $\Sigma^+_{\mathscr{R}'1} \subseteq \Sigma^+_{\mathscr{R}'2}$. Rationamentul fiind simetric, rezulta si incluziunea inversa, deci egalitatea dorita.

Consecinta 1.1 $\Sigma^+_{\mathscr{R}1} = \Sigma^+_{\mathscr{R}A} = \Sigma^+_{\mathscr{R}2}$.

Relatii Armstrong

Definitie. Fie Σ o multime de restrictii functionale peste schema R(U). Numim relatie Armstrong pentru Σ o relatie r_0, care satisface proprietatile:

1) r_0 satisface τ , $\forall \tau \in \Sigma^+_{\mathcal{R}1}$, si

2) r_0 nu satisface γ, $\forall \gamma \notin \Sigma^+_{\mathcal{R}1}$.

Obs. Proprietatea 1) este echivalenta cu 1'):

1') r_0 satisface τ , $\forall \tau \in \Sigma$

Fie $X \subseteq U$ si $\mathscr{R}$ o multime de reguli de inferenta. Sa notam prin $X^{+}_{\mathscr{R}} = \{A | \Sigma \mid\!-_{\mathscr{R}} X \rightarrow A\}$

Lema 1.1 $\Sigma \mid\!-_{\mathscr{R}1} X \rightarrow Y$ daca si numai daca $Y \subseteq X^{+}_{\mathscr{R}1}$.

Demonstratie. Fie $Y = A_1 A_2 \ldots A_m$, $A_i \in U$, $i = 1,m$. Presupunem ca $Y \subseteq X^{+}_{\mathscr{R}1}$. Atunci $A_i \in X^{+}_{\mathscr{R}1}$, $i = 1,m$, deci $\Sigma \mid\!-_{\mathscr{R}1} X \rightarrow A_i$, $i = 1,m$. Aplicand regula A22 (A22 se exprima cu ajutorul regulilor din $\mathscr{R}1$) obtinem:

$\Sigma \mid\!-_{\mathscr{R}1} X \rightarrow Y$.

Invers, daca $\Sigma|\!-_{\mathscr{R}1} X \to Y$, atunci aplicand regula A21, (A21 se exprima cu ajutorul regulilor din $\mathscr{R}1$) se obtine: $\Sigma|\!-_{\mathscr{R}1} X \to A_i$, i = 1,m, ceea ce ınseamna $A_i \in X^+_{\mathscr{R}1}$, i = 1,m, deci $Y \subseteq X^+_{\mathscr{R}1}$.

Lema 1.2 Fie Σ o multime de dependente functionale si σ: $X \to Y$ o dependenta functionala astfel ıncat $\Sigma|\!\!-\!\!\!/\,_{\mathscr{R}1} X \to Y$. Atunci exista o relatie r_σ ce satisface toate dependentele din Σ si r_σ nu satisface $X \to Y$.

Demonstratie. Avem $X^+_{\mathscr{R}1} \not\subseteq U$, caci ın caz contrar,

$X^{+}_{\mathscr{R}1} = U$ si utilizand lema1.1 si faptul ca $Y \subseteq U$, obtinem $\Sigma |-_{\mathscr{R}1} X \rightarrow Y$, deci contradictie.

Definim relatia $r_{\sigma} = \{ t1, t2 \}$, unde t1 este uplul continınd 1 pentru toate atributele din U, iar t2 este definit prin: $t2[A] = 1$ daca $A \in X^{+}_{\mathscr{R}1}$ si 0 altfel.

Deoarece $X^{+}_{\mathscr{R}1} \not\subseteq U$, rezulta $t1 \neq t2$.

1) Aratam ca r_{σ} satisface orice dependenta functionala $V \rightarrow W \in \Sigma$.

Presupunem contrariul, deci r_σ nu satisface $V \to W$.

Atunci $V \subseteq X^+_{\mathscr{R}1}$, caci daca $V \not\subseteq X^+_{\mathscr{R}1}$, exista $B \in V$, si $B \not\in X^+_{\mathscr{R}1}$, ceea ce ınseamna $t1[B] \not\doteq t2[B]$, deci

$t1[V] \not\doteq t2[V]$, de unde rezulta r_σ satisface $V \to W$ (contradictie).

Pe de alta parte, avem: $W \not\subseteq X^+_{\mathscr{R}1}$, caci daca $W \subseteq X^+_{\mathscr{R}1}$, atunci $t1[W] = t2[W]$ (dupa definitia lui r_σ), ceea ce ar ınsemna ca r_σ satisface $V \to W$ (contradictie).

Din W $\not\subseteq X^{+}_{\mathscr{R}1}$ rezulta ca exista $A \in W$, $A \not\in X^{+}_{\mathscr{R}1}$. Deci $t2[A] = 0$ si $t1[A] = 1$.

Relatia $V \subseteq X^{+}_{\mathscr{R}1}$ conduce la $\Sigma |-_{\mathscr{R}1} X \rightarrow V$, dupa lema 1.1. Dar $V \rightarrow W \in \Sigma$, aplicand tranzitivitatea rezulta

$\Sigma |-_{\mathscr{R}1} X \rightarrow W$, ceea ce ınseamna $W \subseteq X^{+}_{\mathscr{R}1}$, contradictie.

2) Aratam ca r_σ nu satisface $X \rightarrow Y$.

Presupunem ca r_σ satisface $X \rightarrow Y$. Aceasta ınseamna ca $t1[X] = t2[X]$ implica $t1[Y] = t2[Y]$.

Deoarece $X \subseteq X^+_{\mathscr{R}1}$ si $t1[A] = t2[A]$, $\forall A \in X^+_{\mathscr{R}1}$, rezulta $t1[X] = t2[X]$, deci avem $t1[Y] = t2[Y]$, ceea ce ınseamna dupa definitia lui r_σ ca $Y \subseteq X^+_{\mathscr{R}1}$ si dupa lema 1.1, obtinem $\Sigma |-_{\mathscr{R}1} X \rightarrow Y$, deci o contradictie.

Teorema 1.1 (Armstrong). Fie Σ o multime de dependente functionale.

Atunci exista o relatie r0 ce satisface exact elementele lui $\Sigma^+_{\mathscr{R}1}$, adica:

1) r0 satisface τ , $\forall \tau \in \Sigma^+_{\mathscr{R}1}$, si

2) r0 nu satisface γ, $\forall \gamma \notin \Sigma^{+}_{\mathscr{R}1}$.

Demonstratie. Deoarece U este finita, rezulta ca Σ este finita si $P(U)\times P(U) - \Sigma^{+}_{\mathscr{R}1}$ este finita. (P(U) este multimea partilor lui U; pentru orice dependenta

functionala $X \to Y$, exista o unica pereche $(X, Y) \in P(U) \times P(U)$).

Fie $\sigma_1, \sigma_2, \ldots, \sigma_k$ dependentele care nu apartin multimii $\Sigma^{+}_{\mathscr{R}1}$. Pentru o restrictie $\sigma \notin \Sigma^{+}_{\mathscr{R}1}$ in lema anterioara am construit $r_\sigma = \{t1, t2\}$, unde t1 si t2 aveau valorile 0 si 1.

Vom considera valori diferite pentru $\sigma_i \neq \sigma_j$, si anume:

In r $_{\sigma i}$ vom considera valorile 2i−2 si 2i−1, i = 1, k.

Desigur r $_{\sigma i}$ le vom construi ca ın lema precedenta:

$r_{\sigma i} = \{t^i{}_1, t^i{}_2\}$,

$t^i{}_1 [A] = 2i - 1, \forall A \in U$, iar

$t^i{}_2 [A] = 2i - 1$ pentru $A \in X^+{}_{i,\mathscr{R}1}$ si $2i - 2$ altfel.

$(\sigma_i : X_i \to Y_i)$.

Definim r0 ca fiind reuniunea relatiilor r $_{\sigma i}$, i = 1, k. In continuare pentru fiecare atribut A ce satisface

$\Sigma |\!-_{\mathscr{R}1} \emptyset \to A$ (echivalent cu $A \in \emptyset^{+}{}_{\mathscr{R}1}$) ınlocuim toate valorile din coloana atributului A ın r0 cu o aceeasi valoare, notata v_A. Vom lua pentru

astfel de atribute A si A′, $v_A \neq v_{A'}$ si $v_A \geq 2k$.

Vom arata ca r0, astfel construita, este o relatie Armstrong pentru Σ.

a) Aratam ca r0 satisface orice $X \to Y \in \Sigma$.

Fie $A \in Y$ oarecare. Este suficient sa aratam ca r0 satisface $X \rightarrow A$ (aplicand uniunea va rezulta ca r0 satisface $X \rightarrow Y$). Distingem doua cazuri:

I. $X \subseteq \varnothing^+_{\mathscr{R}1}$. Avem $\Sigma |\!-_{\mathscr{R}1} \varnothing \rightarrow X$, si ımpreuna cu $X \rightarrow Y \in \Sigma$, obtinem $\Sigma |\!-_{\mathscr{R}1} \varnothing \rightarrow Y$ (prin tranzitivitate), ceea ce ınseamna $Y \subseteq \varnothing^+_{\mathscr{R}1}$ (lema 1.1).

$A \in Y$ si $Y \subseteq \varnothing^+_{\mathscr{R}1}$ implica $A \in \varnothing^+_{\mathscr{R}1}$ si din constructia lui r0 toate valorile ın coloana A sunt v_A, ceea ce denota faptul ca r0 satisface $X \rightarrow A$.

II. $X \not\subseteq \emptyset^{+}_{\mathscr{R}1}$. Rezulta ca $\exists B \in X$, $B \not\in \emptyset^{+}_{\mathscr{R}1}$.

Fie t1, t2 $\in$ r0 cu proprietatea t1[X] = t2[X].

Trebuie sa aratam ca t1[A] = t2[A].

Deoarece $B \in X$ si $B \not\in \emptyset^{+}_{\mathscr{R}1}$, rezulta ca ın aceasta situatie coloana lui B contine 2k valori

distincte, deci t1 si t2 nu contin aceleasi valori.

Deoarece t1[X] = t2[X], rezulta ca t1 si t2 provin din aceeasi relatie $r_{\sigma i}$. (Eventual anumite coloane A din

$t1$ si $t2$ au suferit modificarea cu valori v_A). Dupa lema anterioara $t1$ si $t2$ initiale din $r_{\sigma i}$ au satisfacut Σ, deci si

$X \rightarrow A$. Noile $t1$ si $t2$ obtinute prin identificarea valorilor din anumite coloane, vor satisface, de asemenea,

$X \rightarrow A$, deci $t1[A] = t2[A]$.

b) Aratam ca $r0$ nu satisface σ, $\forall \sigma \notin \Sigma^{+}_{\mathscr{R}1}$ (echivalent cu $r0$ nu satisface σ_i, $i = 1, k$).

Dupa lema anterioara stim ca $r_{\sigma i}$ nu satisface σ_i.

$r_{\sigma i} = \{t^i_1, t^i_2\}$,

$t^i_1 [A] = 2i - 1, \forall A \in U$, iar

$t^i_2[A] = 2i - 1$ pentru $A \in X^+_{i,\mathscr{R}1}$, $2i - 2$ altfel.

$(\sigma_i : X_i \to Y_i)$.

Deoarece $X_i \subseteq X^+_{i,\mathscr{R}1}$ ınseamna ca $\exists B_i \in Y_i$, $B_i \notin X^+_{i,\mathscr{R}1}$

astfel ıncat $t^i_1 [B_i] /= t^i_2 [B_i]$.

Vom arata ca nu toate coloanele de acest tip B_i din Y sufera identificarea datorita apartenentei atributului respectiv la $\varnothing^{+}_{\mathscr{R}1}$.

Presupunem ca pentru orice $B_i \in Y_i$, $B_i \notin X^{+}_{i,\mathscr{R}1}$

avem $Bi \in \varnothing^{+}_{\mathscr{R}1}$. Atunci $\Sigma|\text{-}_{\mathscr{R}1} \varnothing \rightarrow B_i$, aplicand uniunea obtinem: $\Sigma|\text{-}_{\mathscr{R}1}\varnothing \rightarrow Yi - X^{+}_{i,\mathscr{R}1}$. Aceasta ımpreuna cu $\Sigma|\text{-}_{\mathscr{R}1}$ $Xi \rightarrow \varnothing$ si tranzitivitatea ne dau

$\Sigma|\text{-}_{\mathscr{R}1}Xi \rightarrow Yi - X^{+}_{i,\mathscr{R}1}$.

Dar avem $\Sigma|\!-_{\mathscr{R}1} X_i \to X^{+}_{i,\mathscr{R}1}$, aplicand uniunea pentru toate A din $X^{+}_{i,\mathscr{R}1} = \{A | \Sigma|\!-_{\mathscr{R}1} X_i \to A\}$. Aplicand ınca o data uniunea obtinem $\Sigma|\!-_{\mathscr{R}1} X_i \to Y_i$, adica $\Sigma|\!-_{\mathscr{R}1} \sigma_i$, contradictie. Rezulta ca ın urma identificarii $\exists B_i \in Y_i$, astfel ıncat $t^i_1 [B_i] /= t^i_2 [B_i]$. Avand $t^i_1 [X_i] = t^i_2 [X_i]$, rezulta ca r0 nu satisface σ_i.

Studiul dependentelor functionale utilizand calculul propozitional

Acest studiu a fost realizat de Fagin. Pentru fiecare atribut $A \in U$ se asociaza o variabila propozitionala notata a. Corespunzator dependentei functionale

$A1 \ldots Am \rightarrow B1 \ldots Bp$ asociem implicatia logica

$a1 \ldots am \Rightarrow b1 \ldots bp$, unde $a1 \ldots am$ ınseamna conjunctia logica a variabilelor $a1, a2, \ldots, am$;

similar b1 . . . bp. Semnul “$\Rightarrow$” reprezinta implicatia logica.

Vom nota valorile de adevar din calculul propozitional prin 1 (adevarat) , si 0 (fals).

O asignare o notam prin δ si este o functie δ: Var $\rightarrow$ {0, 1}, unde Var este multimea variabilelor propozitionale.

Functia δ se extinde la formule ın general (si la implicatii ın particular) prin

δ (p1 $\wedge$ p2) = δ (p1) $\wedge\delta$ (p2), p1 si p2 fiind formule,

δ (p1 v p2) = δ (p1) vδ (p2),

δ (¬ p) = ¬ δ (p),

unde conjunctia, disjunctia si negatia ın {0, 1} sunt definite ın modul cunoscut.

Rezulta ca δ (a1 . . . am $\Rightarrow$ b1 . . . bp) = 1 daca si numai daca:

1. $\exists$i, 1 ≤ i ≤ m, astfel ıncat δ (ai) = 0 sau

2. $\forall$i, i = 1,m, δ (ai) = 1 , si $\forall$j, j = 1, p, δ (bj) = 1.

In calculul propozitional exista notiunea de consecinta logica a unei formule g dintr-o multime de formule F. Se noteaza prin F $|=_{c.l.}$ g, daca

$(\forall\delta)[(\forall f \in F, \delta (f) = 1) \rightsquigarrow \delta (g) = 1]$;

δ noteaza o asignare a variabilelor propozitionale.

Pentru o dependenta σ vom nota prin $\bar{\sigma}$ implicatia atasata, iar pentru o multime de dependente functionale Σ, vom nota prin $\bar{\Sigma} = \{ \bar{\sigma} | \sigma \in \Sigma\}$.

Exemplul 2.1 Fie $\Sigma = \{AB \to C, AC \to D, BD \to E\}$. Atunci

$\bar{\Sigma} = \{ab \Rightarrow c, ac \Rightarrow d, bd \Rightarrow e\}$.

Vom stabili o legatura ıntre notiunea de “consecinta” definita ın domeniul dependentelor functionale si notiunea de “consecinta logica” definita ın domeniul calculului propozitional.

Fie $\sigma: AB \to E$ si Σ din exemplul anterior. Se pune ıntrebarea daca $\Sigma \models \sigma$?

Pentru aceasta, fie r o relatie ce satisface toate elementele lui Σ si fie t1, t2 $\in$ r

astfel ıncat t1[AB] = t2[AB]. Din r satisface AB $\rightarrow$ C rezulta t1[C] = t2[C], deci t1[AC] = t2[AC].

Din faptul ca r satisface AC $\rightarrow$ D rezulta t1[D] = t2[D], deci t1[BD] = t2[BD].

Din faptul ca r satisface BD $\rightarrow$ E rezulta ca t1[E] = t2[E],

deci r satisface AB $\rightarrow$ E. In concluzie $\Sigma \models$ AB $\rightarrow$ E.

Vrem sa vedem acum daca $\bar{\Sigma} \models_{c.l.} \bar{\sigma}$? Pentru aceasta vom putea proceda astfel: construim tabela cu cele 25 asignari ale variabilelor propozitionale a, b, c, d, e , si vom calcula valoarea de asignare pentru toate elementele lui $\bar{\Sigma}$ si pentru $\bar{\sigma}$, si daca ın fiecare caz, ın care toate elementele lui $\bar{\Sigma}$ sunt 1, rezulta si $\bar{\sigma}$ este 1, atunci vom avea $\bar{\Sigma} \models_{c.l.} \bar{\sigma}$.

Dar putem reduce acest calcul considerand

asignarile δ astfel:

1. Daca $\delta(a) = 0$ sau $\delta(b) = 0$, atunci $\delta(\bar{\sigma}) = 1$.

2. Daca $\delta(a) = \delta(b) = 1$, atunci deoarece δ trebuie sa satisfaca $ab \Rightarrow c$, rezulta $\delta(c) = 1$. Cum δ trebuie sa satisfaca $ac \Rightarrow d$, rezulta $\delta(d) = 1$, si la fel δ trebuie sa satisfaca $bd \Rightarrow e$, rezulta $\delta(e) = 1$, deci $\delta(ab \Rightarrow e) = 1$, adica $\delta(\bar{\sigma}) = 1$.

Vom stabili echivalenta ıntre “consecinta” din universul dependentelor functionale si “consecinta logica” din calculul propozitional.

Teorema 2.1 (de echivalenta). Fie Σ o multime de dependente functionale si σ o dependenta functionala. Fie $\bar{\Sigma}$ multimea de implicatii corespunzatoare lui Σ si $\bar{\sigma}$ implicatia asociata lui σ. Atunci avem:

σ este consecinta a lui Σ daca si numai daca

$\bar{\sigma}$ este consecinta logica a lui $\bar{\Sigma}$

$$\Sigma \models \sigma \text{ iff } \bar{\Sigma} \models_{c.l.} \bar{\sigma}.$$

Conform acestei teoreme problema deciderii daca o dependenta functionala este consecinta a unei multimi de dependente se transforma ıntr-o problema de a decide daca o implicatie este consecinta logica a unei multimi de implicatii.

Pentru aceasta ultima problema de decizie exista un algoritm eficient, datorat lui Chang.

Dupa propozitiile 1.1, 1.2, putem presupune ca Σ are dependentele cu un singur atribut ın partea dreapta , si la fel σ are ın partea dreapta un singur atribut.

Algoritmul lui Chang:

1. Se considera cuvinte formate peste $\overline{U} \cup \{\sim, *\}$ astfel:

pentru a1 . . . am $\Rightarrow$ b $\in \overline{\Sigma}$ se considera cuvantul

$\sim a1 * \sim a2 * \ldots \sim am * b$

Daca U = {A1, . . . ,An}, atunci $\overline{U}$ = {a1, . . . , an}.

Sa notam prin S multimea de cuvinte obtinute.

2. Daca $\bar{\sigma}$ are forma: c1 . . . ck ⇒ d, atunci adaugam la S urmatoarele k + 1 cuvinte:

c1

c2

...

ck

~ d

Numim o variabila ai “atom” , si ~ ai “atom negat”.

3. Algoritmul cauta un atom X, astfel ıncat X este un cuvant ın S , si exista un cuvant ce ıncepe cu ~ X. Daca exista un astfel de atom, se selecteaza

unul arbitrar si se sterge ~ X din fiecare sir ce ıncepe cu ~ X (eventual si * daca acest caracter * urmeaza ın sir).

4. Se continua pasul 3) atat timp cat exista un astfel de atom. Este clar ca algoritmul se opreste ıntotdeauna si exista 2 situatii:

(a) s-a obtinut sirul vid, notat cu λ sau

(b) nu exista nici un atom X ce satisface cele 2 conditii si nu s-a obtinut sirul vid.

Daca avem (a) atunci $\bar{\sigma}$ este consecinta logica a lui $\bar{\Sigma}$, altfel $\bar{\sigma}$ nu este consecinta logica a lui $\bar{\Sigma}$.

Exemplul 2.2 Fie $\bar{\Sigma}$ = {ab ⇒ c, ac ⇒ d, bd ⇒ e} si $\bar{\sigma}$: ab ⇒ e. Atunci:

S = {~ a∗ ~ b ∗ c, ~ a∗ ~ c ∗ d, ~ b∗ ~ d ∗ e, a, b, ~ e}.

Putem alege X = a, atunci S devine

S = {~ b ∗ c, ~ c ∗ d, ~ b∗ ~ d ∗ e, a, b, ~ e}

In continuare putem lua X = b si S devine

S = {c, ~ c ⋆ d, ~ d ⋆ e, a, b, ~ e}

Fie acum X = c, atunci S devine:

S = {c, d,~ d ⋆ e, a, b, ~ e}

Luand acum X = d obtinem:

S = {c, d, e, a, b,~ e},

Pentru X = e, rezulta:

S = {c, d, e, a, b, λ}, deci $\bar{\Sigma} \models_{c.l.} \bar{\sigma}$.

Teorema 2.2 (de completitudine a dependentelor). Fie Σ o multime de dependente functionale, σ o dependenta functionala. Atunci avem σ este o consecinta a lui Σ daca si numai daca σ are o demonstratie ın Σ utilizand regulile de inferenta A1,A2,A3 (axiomele lui Armstrong).

In notatii, teorema se exprima prin:

$\Sigma \models \sigma$ iff $\Sigma \vdash_{\mathscr{R}A} \sigma$.

Demonstratie. Regulile de inferenta A1,A2,A3 au forma:

(A1) $A1 \ldots Am \rightarrow Ai, \ 1 \leq i \leq m$

(A21) $$\frac{A1 \ldots Am \rightarrow B1 \ldots Br}{A1 \ldots Am \rightarrow Bj}, \ j = 1, r,$$

(A22) $$\frac{A1 \ldots Am \rightarrow Bj \ , \ j = 1, r}{A1 \ldots Am \rightarrow B1 \ldots Br}$$

(A3) $\dfrac{A1 \ldots Am \rightarrow B1 \ldots Br, \quad B1 \ldots Br \rightarrow C1 \ldots Cp}{A1 \ldots Am \rightarrow C1 \ldots Cp}$

Fapt 1: Daca $\mathscr{R}$ este un sistem de reguli de inferenta valide si daca $\Sigma \vdash_{\mathscr{R}} \sigma$, atunci $\Sigma \models \sigma$.

O regula $\dfrac{\alpha_1, \alpha_2, \ldots, \alpha_k}{\alpha}$

se numeste valida daca$\{ \alpha_1, \alpha_2, \ldots, \alpha_k\} \models \alpha$.

Justificarea faptului rezulta prin inductie asupra lungimii demonstratiei. Fie

$\sigma_1, \sigma_2, \ldots, \sigma_h = \sigma$ o demonstratie pentru σ ın Σ utilizand regulile $\mathscr{R}$. Daca

$h = 1$, atunci $\sigma_1 = \sigma$ poate fi ın una din situatiile:

a) $\sigma_1 = \sigma \in \Sigma$, si deci $\Sigma \models \sigma$, sau

b) exista o regula de forma $\frac{\quad}{\sigma} \in \mathscr{R}$;

regula fiind valida, rezulta σ dependent triviala

(deci satisfacuta de orice relatie), de unde $\Sigma \models \sigma$.

Presupunem Fapt 1 adevarat pentru demonstratii de lungime mai mica sau egala cu h. Fie o demonstratie a lui σ de lungime $h + 1$: $\sigma_1, \sigma_2, \ldots, \sigma_h, \sigma_{h+1} = \sigma$.

Pentru σ avem 2 situatii:

c) $\sigma \in \Sigma$, deci $\Sigma \models \sigma$ sau

d) $\exists \quad \sigma_{i1}, \ldots, \sigma_{ik}$

------------------- $\in \mathscr{R}$, i1, . . . , ik ≤ h.

σ

In cazul d) dupa ipoteza inductiei avem:

$\Sigma \models \sigma_{i1}$, . . . , $\Sigma \models \sigma_{ik}$, si deoarece

regula este valida, avem $\{\sigma_{i1}, \ldots, \sigma_{ik}\} \models \sigma$.

Rezulta atunci $\Sigma \models \sigma$.

Regulile (A1) (A2) (A3) sunt valide, din proprietatile

dependentelor functionale.

Deci putem aplica Fapt 1, ceeace implica:

Daca Σ|--$_{\mathscr{R}A}$ σ, atunci Σ |= σ.

Invers: presupunem Σ |= σ. Consideram $\Sigma^{+}_{\mathscr{R}A}$
(ınchiderea lui Σ referitoare la regulile de inferenta $\mathscr{R}$A).

Stim ca:

$$\Sigma^{+}_{\mathscr{R}A} = \{\sigma_1 | \Sigma |\text{--}_{\mathscr{R}A}\, \sigma_1\}.$$

Trebuie sa aratam ca $\sigma \in \Sigma^{+}_{\mathscr{R}A}$.

Avem: $\Sigma \subseteq \Sigma^{+}_{\mathscr{R}A}$.

Dupa teorema lui Armstrong (Teorema 1.1) exista o relatie r0 ce satisface exact elementele din $\Sigma^+_{\mathscr{R}1}$, unde $\mathscr{R}1$ = {FD1f, FD2f, FD3f}.

Dupa consecinta 1.1, avem $\Sigma^+_{\mathscr{R}1} = \Sigma^+_{\mathscr{R}A}$. Deci relatia r0 satisface exact elementele lui $\Sigma^+_{\mathscr{R}A}$.

Deoarece $\Sigma \subseteq \Sigma^+_{\mathscr{R}A}$, rezulta ca r0 satisface toate dependentele din Σ.

Avand $\Sigma \models \sigma$ obtinem ca r0 satisface σ. Deoarece r0 satisface exact elementele lui $\Sigma^+_{\mathscr{R}A}$, obtinem:

$\sigma \in \Sigma^{+}{}_{\mathscr{R}A}$, deci $\Sigma|\text{--}_{\mathscr{R}A}\ \sigma$.

Reguli de inferenta pentru formule.

(similare cu regulile A1,A2, A3).

(A1′) ------------------, $1 \le i \le m$

$a1 \ldots am \Rightarrow ai$

(A21′) $a1 \ldots am \Rightarrow b1 \ldots br$

$$\overline{}, \; j = 1, r,$$

$$a1 \ldots am \Rightarrow bj$$

(A22') $\quad a1 \ldots am \Rightarrow bj, \; j = 1, r$

$$\overline{}$$

$$a1 \ldots am \Rightarrow b1 \ldots br$$

(A3′) $\quad a1 \ldots am \Rightarrow b1 \ldots br, \; b1 \ldots br \Rightarrow c1 \ldots cp$

$$\overline{}$$

$$a1 \ldots am \Rightarrow c1 \ldots cp$$

Teorema 2.3 (de completitudine implicationala). Fie $\overline{\Sigma}$ o multime de implicatii asociate dependentelor functionale din Σ, $\bar{\sigma}$ implicatia asociata dependentei functionale σ. Atunci $\bar{\sigma}$ este consecinta logica a lui $\overline{\Sigma}$ daca si

numai daca $\bar{\sigma}$ are o demonstratie ın $\overline{\Sigma}$, utilizand regulile de inferenta (A1′),(A21′),(A22′), (A3′).

Demonstratie. Regulile (A1′), (A21′), (A22′), (A3′) considerate ca reguli ın calculul propozitional sunt valide, adica orice asignare care face adevarate premizele regulii, va face adevarata si concluzia regulii. Procedand ca ın demonstratia teoremei de completitudine a dependentelor obtinem ca daca $\Sigma \vdash_{A1', A21', A22', A3'} \sigma$, atunci $\Sigma \models_{c.l.} \sigma$.

Invers, fie $\Sigma \models_{c.l.} \sigma$.

Va trebui sa aratam ca: $\Sigma \vdash_{A1', A21', A22', A3'} \sigma$.

Fie $\bar{\sigma}$: $a_1 \dots a_m \Rightarrow d_1 \dots d_h$.

Sa notam prin:

PROV E = $\{e \mid \Sigma \mid\!\!-\!\!-_{A1',A21',A22',A3'}\ a_1 \dots a_m \Rightarrow e\}$

Aplicand regula (A1′) obtinem $a_i \in$ PROV E, $1 \leq i \leq m$.

Aratam ca $d_j \in$ PROV E, pentru orice $j = 1, h$.

Presupunem contrariul, deci $\exists j \in \{1, 2, \dots, h\}$, astfel

ıncat $d_j \notin$ PROV E, adica:

Σ|-/- $_{A1',A2',A22',A3'}$ $a_1 \ldots a_m \Rightarrow d_j$.

Consideram urmatoarea asignare:

$\delta_0(x) = 1$ daca $x \in$ PROV E si 0 altfel

Deoarece $a_i \in$ PROV E, rezulta ca $\delta_0(a_i) = 1$, orice i = 1,m.

Relatia: $d_j \notin$ PROV E implica $\delta_0(d_j) = 0$, deci $\delta_0 (\bar{\sigma}) = 0$. Aratam ca δ_0 satisface toate elementele din $\bar{\Sigma}$.

Fie $b_1 \ldots b_r \Rightarrow c_1 \ldots c_p \in \bar{\Sigma}$. Avem doua situatii:

1. $b_i \in$ PROV E pentru orice i = 1, r; aceasta ınseamna ca $\Sigma|\text{--}_{A1',A21',A21',A3'}\ a_1 \ldots a_m \Rightarrow b_i$, i = 1, r

Aplicand (A22′) obtinem:

$\Sigma|\text{--}_{A1',A21',A22',A3'}\ a_1 \ldots a_m \Rightarrow b_1 \ldots b_r$.

Aceasta relatie ımpreuna cu $b_1 \ldots b_r \Rightarrow c_1 \ldots c_p \in \overline{\Sigma}$ si (A3′) conduc la:

$\Sigma|\text{--}_{A1',A21',A22',A3'}\ a_1 \ldots a_m \Rightarrow c_1 \ldots c_p$.

De aici, aplicand (A22′) obtinem:

$\Sigma|\!-\!-_{A1',A21',A22',A3'}\ a_1 \ldots a_m \Rightarrow c_k$, $1 \le k \le p$,

ceea ce ınseamna ca $c_k \in$ PROVE, orice k, $1 \le k \le p$, si dupa definitia lui δ_0 obtinem :

$\delta_0\ (c_k) = 1$, $\forall k$, $1 \le k \le p$, deci $\delta_0\ (c_1 \ldots c_p) = 1$, ceea ce ınseamna:

$\delta_0(b_1 \ldots b_r \Rightarrow c_1 \ldots c_p) = 1.$

2. Exista $i \in \{1, 2, \ldots, r\}$, astfel ıncat $b_i \notin$ PROVE. Aceasta ınseamna ca:

δ_0 (b_i) = 0, deci δ_0 (b_1 . . . b_r) = 0, de unde

δ_0 (b_1 . . . $b_r \Rightarrow c_1$ . . . c_p) = 1.

Asignarea δ_0 construita mai sus satisface conditia

δ_0 ($\bar{\gamma}$) = 1, orice $\bar{\gamma} \in \overline{\Sigma}$, dar avem:

δ_0 ($\bar{\sigma}$) = 0. Acest lucru contrazice ipoteza $\Sigma \models_{c.l.} \sigma$.

Inseamna ca toate variabilele $d_j \in$ PROV E, j = 1, h.

De aici rezulta:

$\Sigma \vdash_{A1', A21', A22', A3'} a_1 \ldots a_m \Rightarrow d_j$, j = 1, h

Aplicand A22′ se obtine :

$\Sigma \vdash_{A1',A21',A22',A3'} a_1 \ldots a_m \Rightarrow d_1 \ldots d_h$, adica:

$\Sigma \vdash_{A1',A21',A22',A3'} \bar{\sigma}$.

Demonstrarea teoremei de echivalenta.

Teorema 2.1 (de echivalenta). Fie Σ o multime de dependente functionale si σ o dependenta functionala. Fie $\bar{\Sigma}$ multimea de implicatii corespunzatoare lui Σ si $\bar{\sigma}$ implicatia asociata lui σ. Atunci avem:

σ este consecinta a lui Σ daca si numai daca

$\bar{\sigma}$ este consecinta logica a lui $\bar{\Sigma}$

$$\Sigma \models \sigma \text{ iff } \bar{\Sigma} \models_{c.l.} \bar{\sigma}.$$

Demonstratie. Dupa teorema de completitudine a dependentelor avem:

$\Sigma \models \sigma$ iff $\Sigma \vdash_{\mathscr{R}A} \sigma$

Orice demonstratie a lui σ ın Σ, utilizand regulile $\mathscr{R}_A$ se poate transforma sintactic ıntr-o demonstratie a lui $\bar{\sigma}$ ın $\bar{\Sigma}$ utilizand regulile A′1,A21’,A22’,A3’ si ınlocuind

atributele prin variabilele propozitionale respective, semnul ” → ” prin ” ⇒ ”si invers. Deci avem:

$\Sigma \vdash_{\mathscr{R}A} \sigma$ iff $\bar{\Sigma} \vdash_{A1', A21', A22', A3'} \bar{\sigma}$

Dupa teorema de completitudine implicationala

avem:

Σ|-- $_{A1',A21',A22',A3'}$ $\bar{\sigma}$ iff $\bar{\Sigma}$|=$_{c.l.}$ $\bar{\sigma}$.

Astfel avem:

Σ |= σ iff $\bar{\Sigma}$|=$_{c.l.}$ $\bar{\sigma}$.

Dependente multivaluate

Fie X, Y $\subseteq$ U. O dependenta multivaluata este notata sintactic prin $X \to\to Y$.

Vom da doua definitii pentru satisfacerea unei dependente multivaluate de catre o relatie r peste U.

Definitia 1. Relatia r peste U satisface dependenta multivaluata $X \to\to Y$, daca pentru orice doua uple $t1, t2 \in r$ si $t1[X] = t2[X]$, exista uplele t3 si t4 din r, astfel ıncat:

(i) $t3[X] = t1[X]$, $t3[Y] = t1[Y]$ si $t3[Z] = t2[Z]$;

(ii) $t4[X] = t2[X]$, $t4[Y] = t2[Y]$ si $t4[Z] = t1[Z]$,

unde $Z = U - XY$.

In definitia 1 este suficient sa cerem existenta lui t3 sau t4, celalalt uplu rezulta considerand uplele ın ordinea t2,t1.

Pentru $t \in r$ avem $t[X] \in r[X]$. Notam prin $F_Y(t[X]) = \{t'[Y]/t' \in r, t'[X] = t[X]\}$. Aceasta se numeste multimea Y -valorilor asociate lui t[X].

Exemplul 1. Fie relatia r data astfel:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_2	c_2	d_2
a_1	b_1	c_1	d_2
a_1	b_2	c_2	d_1

a_2 b_3 c_1 d_1

a_2 b_3 c_1 d_2

Se verifica faptul ca r satisface $A \to\to BC$

conform definitiei 1.

Pentru $t \in r$, definim:

$M_Y(t[XZ]) = \{t'[Y]/t' \in r, t'[XZ] = t[XZ]\}$.

Definitia 2. Relatia r peste U satisface dependenta multivaluata $X \to\to Y$, daca pentru orice t1, t2 $\in$ r, astfel incat t1[X] = t2[X] avem:

M_Y (t1[XZ]) =M_Y (t2[XZ]).

Propozitia 1. Definitiile 1 si 2 sunt echivalente.

Demonstratie. Fie r ce satisface $X \to\to Y$ dupa definitia 1 si fie t_1, $t_2 \in r$, astfel ıncat $t_1[X] = t_2[X]$.

Fie $t'_1[Y] \in M_Y(t_1[XZ])$. Sa aratam ca

$t'_1[Y] \in M_Y(t_2[XZ])$.

Avem $t'_1 \in r$ si $t'_1[XZ] = t_1[XZ]$. Rezulta $t'_1[X] = t_1[X] = t_2[X]$. Deoarece r satisface $X \rightarrow\rightarrow Y$ dupa definitia 1 si $t'_1[X] = t_2[X]$, exista $t_3 \in r$, astfel ıncat: $t_3[X] = t'_1[X]$, $t_3[Y] = t'_1[Y]$ si $t_3[Z] = t_2[Z]$. De aici $t_3[XZ] = t_2[XZ]$, de unde $t_3[Y] \in M_Y(t_2[XZ])$. Deoarece $t'_1[Y] = t_3[Y]$, obtinem $t'_1[Y] \in M_Y(t_2[XZ])$.

Rationamentul fiind simetric rezulta si invers,

adica $t'_2[Y] \in M_Y(t_2[XZ])$ implica:

$t'_2[Y] \in M_Y(t_1[XZ])$.

Am aratat:

$M_Y(t_1[XZ]) = M_Y(t_2[XZ])$, adica r satisface $X \rightarrow\rightarrow Y$ dupa definitia 2.

Presupunem acum ca r satisface $X \rightarrow\rightarrow Y$ dupa definitia 2 si fie $t_1, t_2 \in r$, astfel ıncat $t_1[X]= t_2[X]$.

Din $M_Y(t_1[XZ]) = M_Y(t_2[XZ])$ si faptul ca $t_1[Y] \in M_Y(t_1[XZ])$ obtinem $t_1[Y] \in M_Y(t_2[XZ])$.

Deci $\exists t'_2 \in r$, astfel ıncat $t_1[Y] = t'_2[Y]$ si $t'_2[XZ] = t_2[XZ]$. Pentru $t'_2 \in r$ avem: $t'_2[X] = t_2[X] = t_1[X]$, $t'_2[Y] = t_1[Y]$ si $t'_2[Z] = t_2[Z]$.

Similar obtinem $t'_1 \in r$, astfel ıncat: $t'_1[X] = t_2[X]$, $t'_1[Y] = t_2[Y]$ si $t'_1[Z] = t_1[Z]$. Am aratat ca r satisface $X \rightarrow\rightarrow Y$ dupa definitia 1.

Observatia 1. Daca r satisface dependenta functionala $X \to Y$, atunci pentru orice $t \in r$, avem $M_Y(t[XZ]) = \{t[Y]\}$.

Observatia 2. Daca r satisface dependenta functionala $X \to Y$, atunci r satisface dependenta multivaluata $X \to\to Y$.

Observatia 3. Daca r satisface dependenta multivaluata $X \rightarrow\rightarrow Y$, atunci putem defini o functie $\psi : r[X] \rightarrow P(r[Y])$, prin $\psi(t[X]) = M_Y(t[XZ])$, $\forall t \in r$. Cand r satisface $X \rightarrow Y$, atunci $\psi : r[X] \rightarrow r[Y]$.

Proprietati ale dependentelor multivaluate:

Propozitia 2. MVD0 (Complementariere).

Fie $X, Y, Z \subseteq U$, astfel ıncat $XYZ = U$ si $Y \cap Z \subseteq X$. Daca r satisface $X \rightarrow\rightarrow Y$, atunci r satisface $X \rightarrow\rightarrow Z$.

MVD1 (Reflexivitate). Daca $Y \subseteq X$, atunci orice relatie r satisface $X \rightarrow\rightarrow Y$.

MVD2 (Extensie). Fie $Z \subseteq W$ si r satisface $X \to\to Y$. Atunci r satisface $XW \to\to YZ$.

MVD3 (Tranzitivitate). Daca r satisface $X \to\to Y$ si $Y \to\to Z$, atunci r satisface $X \to\to Z - Y$.

MVD4 (Pseudotranzitivitate). Daca r satisface $X \to\to Y$, si $YW \to\to Z$, atunci r satisface $XW \to\to Z - YW$.

MVD5 (Uniune). Daca r satisface $X \to\to Y$, si $X \to\to Z$, atunci r satisface $X \to\to YZ$.

MVD6 (Descompunere). Daca r satisface $X \to\to Y$, si $X \to\to Z$, atunci r satisface $X \to\to Y \cap Z, X \to\to Y - Z, X \to\to Z - Y$.

Deoarece vom lucra cu multimi de dependente, ce pot fi functionale sau multivaluate, vom avea nevoie de asa numitele proprietati mixte.

FD-MVD1. Daca r satisface $X \rightarrow Y$, atunci r satisface $X \rightarrow\rightarrow Y$.

FD-MVD2. Daca r satisface $X \rightarrow\rightarrow Z$ si $Y \rightarrow Z'$, cu $Z' \subseteq Z$ si $Y \cap Z = \varnothing$, atunci r satisface $X \rightarrow Z'$.

FD-MVD3. Daca r satisface $X \rightarrow\rightarrow Y$ si $XY \rightarrow Z$, atunci r satisface $X \rightarrow Z - Y$.

Demonstrarea lui MVD0:

Pentru X, Y, Z ⊆ U vom considera, ın general, urmatoarea diagrama:

U reprezinta ıntregul patrat, X = {1, 2, 3, 4}, Y = {2, 3, 5, 6}, Z = {3, 4, 6, 7}, U − XYZ = {8}.

In conditiile proprietatii MVD0 ($XYZ = U$ si $Y \cap Z \subseteq X$), avem $8 = \varnothing$ si $6 = \varnothing$.

Fie $T1 = U - XY = \{7\}$, $T2 = U - XZ = \{5\}$.

Presupunem ca r satisface $X \to\to Y$. Aceasta ınseamna ca pentru orice $t, t' \in r$ cu $t[X] = t'[X]$, exista t_3 si $t_4 \in r$, astfel ıncat $t_3[X] = t[X]$, $t_3[Y] = t[Y]$, $t_3[T1] = t'[T1]$ si $t_4[X] = t'[X]$, $t_4[Y] = t'[Y]$, $t_4[T1] = t[T1]$.

Sa notam prin t_i respectiv t'_i, proiectia lui t respective t′, pe domeniul i. Atunci pentru t avem:

$$\begin{array}{llll} & X & Y & T1 \\ t \rightarrow & ((t_1, t_2, t_3, t_4), & (t_2, t_3, t_5), & (t_7)) \\ t' \rightarrow & ((t'_1, t'_2, t'_3, t'_4), & (t'_2, t'_3, t'_5), & (t'_7)) \\ & X & Z & T2 \\ t \rightarrow & ((t_1, t_2, t_3, t_4), & (t_3, t_4, t_7), & (t_5)) \end{array}$$

$t' \rightarrow ((t'_1, t'_2, t'_3, t'_4), (t'_3, t'_4, t'_7), (t'_5))$

Din $t[X] = t'[X]$ rezulta $t_i = t'_i$, $i = 1, 4$.

Aplicam faptul ca r satisface $X \rightarrow\rightarrow Y$. Rezulta

ca exista $t'' \in r$, astfel ıncat:

X Y T1

$t'' \rightarrow ((t_1, t_2, t_3, t_4), (t_2, t_3, t_5), (t'_7))$.

Acest t'' proiectat pe X, Z, T2 da:

$t'' \rightarrow ((t_1, t_2, t_3, t_4), (t_3, t_4, t'_7), (t_5)) =$

$((t'_1, t'_2, t'_3, t'_4), (t'_3, t'_4, t'_7),(t_5))$

Avem deci satisfacuta definitia 1 pentru t' si t.

Multimile Y si Z intervin simetric ın MVD0, deci rezulta si invers: daca r satisface $X \to\to Z$, atunci r satisface $X \to\to Y$.

Demonstrarea proprietatii MVD1:

$Y = \{1\}$, $X = \{1, 2\}$, $Z = U - XY = \{3\}$.

Fie $t, t' \in r$, astfel ıncat $t[X] = t'[X]$, adica

$t_i = t'_i$, $i = 1, 2$.

$$\begin{array}{llll} & X & Y & Z \\ t \rightarrow & ((t_1, t_2), & (t_1), & (t_3)) \\ t' \rightarrow & ((t'_1, t'_2), & (t'_1), & (t'_3)) \end{array}$$

consideram $t'' = t'$. Avem:

$t'' \rightarrow ((t'_1, t'_2), (t'_1), (t'_3)) = ((t_1, t_2), (t_1), (t'_3))$

Demonstrarea proprietatii MVD2:

Avem $Z \subseteq W$ si r satisface $X \rightarrow\rightarrow Y$. Aratam ca r satisface $XW \rightarrow\rightarrow YZ$.

$X = \{1, 3, 4, 5, 6, 7\}$

$Y = \{2, 3, 4, 5, 8, 9\}$

$W = \{4, 5, 6, 7, 8, 9, 10, 11\}$

$Z = \{5, 7, 9, 10\}$

$T1 = U - XY = \{10, 11, 12\}$

$T2 = U - XWY\,Z = \{12\}$.

Fie $t, t' \in r$, astfel ıncat $t[XW] = t'[XW]$, adica $t_i =$

t'_i, $i = 1, 3, 4 - 11$.

XW YZ T2

$t \rightarrow ((t_1, t_3, t_4 - t_{11}), (t_2 - t_5, t_7 - t_{10}), (t_{12}))$

$t' \rightarrow ((t'_1, t'_3, t'_4 - t'_{11}), (t'_2 - t'_5, t'_7 - t'_{10}, (t'_{12})),$

unde $t_i - t_j$ noteaza toate componentele

ıncepand cu t_i si terminand cu t_j ,$(i < j)$.

Proiectam acum t si t′ pe tripleta (X, Y, T1):

$t \rightarrow ((t_1, t_3, t_4 - t_7), (t_2 - t_5, t_8, t_9), (t_{10} - t_{12}))$

$t' \rightarrow ((t'_1, t'_3, t'_4 - t'_7), (t'_2 - t'_5, t'_8, t'_9), (t'_{10} - t'_{12}))$.

Deoarece avem $t[X]=t'[X]$ si r satisface $X \rightarrow\rightarrow Y$,

rezulta ca exista $t'' \in r$, astfel ıncat:

$t'' \to ((t_1, t_3, t_4 - t_7), (t_2 - t_5, t_8, t_9), (t'_{10} - t'_{12}))$ pe X, Y, T1.

Proiectand acest t'' pe XW, YZ, T2, obtinem:

$t'' \to ((t_1, t_3, t_4 - t_9, t'_{10}, t'_{11}), (t_2 - t_5, t_7 - t_9, t'_{10}), (t'_{12}))$

$= ((t_1, t_3, t_4 - t_{11}), (t_2 - t_5, t_7 - t_{10}), (t'_{12}))$,

deoarece $t_i = t'_i$, $i = 1, 3, 4 - 11$.

Pentru t si t' am obtinut t'' care satisface definitia1.

Demonstrarea proprietatii MVD3:

Daca r satisface $X \rightarrow\rightarrow Y$, si $Y \rightarrow\rightarrow Z$, atunci r satisface $X \rightarrow\rightarrow Z - Y$.

$X = \{1, 2, 3, 4\}$

$Y = \{2, 3, 5, 6\}$

$Z = \{3, 4, 6, 7\}$,

$Z - Y = \{4, 7\}$

$T1 = U - XY = \{7, 8\}$

$T2 = U - Y\,Z = \{1, 8\}$

$T3 = U - X(Z - Y\,) = \{5, 6, 8\}$

Fie $t, t' \in r$, astfel ıncat $t[X] = t'[X]$, adica $t_i = t'_i$,

$i = 1, 4$

$t \to ((t_1 - t_4), (t_4, t_7), (t_5, t_6, t_8)$

$t' \to ((t'_1 - t'_4), (t'_4, t'_7), (t'_5, t'_6, t'_8))$ pe $X, Z - Y$, T3

Consideram t si t' proiectate pe X, Y, T1:

$t \to ((t_1 - t_4), (t_2, t_3, t_5, t_6), (t_7, t_8))$

$t' \to ((t'_1 - t'_4), (t'_2, t'_3, t'_5, t'_6), (t'_7, t'_8))$

Deoarece r satisface $X \to\to Y$, rezulta ca exista

$t'' \in r$, astfel ıncat:

$t'' \to ((t_1 - t_4), (t_2, t_3, t_5, t_6), (t'_7, t'_8))$ pe X, Y, T1.

Consideram acum t si t'' pe Y,Z, T2.

$t \rightarrow ((t_2, t_3, t_5, t_6), (t_2, t_4, t_6, t_7), (t_1, t_8))$

$t'' \rightarrow ((t_2, t_3, t_5, t_6), (t_2, t_4, t_6, t'_7), (t_1, t'_8))$

r satisface $Y \rightarrow\rightarrow Z$. Pentru t'' si t exista $t''' \in r$,

astfel ıncat

$t''' \rightarrow ((t_2, t_3, t_5, t_6), (t_2, t_4, t_6, t'_7),(t_1, t_8))$.

Considerand t''' proiectat pe X, Z − Y si T3

obtinem:

$t''' \rightarrow ((t_1 - t_4), (t_4, t'_7), (t_5, t_6, t_8)) =$

$((t'_1 - t'_4), (t'_4, t'_7), (t_5, t_6, t_8))$

In concluzie, pentru t' si $t \in r$ cu $t'[X] = t[X]$ am gasit t''' ce satisface definitia 1.

Am aratat FD-MVD1. Sa aratam acum FD - MVD2:

Fie r care satisface $X \rightarrow\rightarrow Z$ si $Y \rightarrow Z'$, cu $Z' \subseteq Z$ si $Y \cap Z = \emptyset$.

Sa aratam ca r satisface $X \rightarrow Z'$.

$X = \{1, 2, 3, 4\}$

$Y = \{2, 5\}$

$Z = \{3, 4, 6, 7\}$

$Z' = \{4, 6\}$

$T1 = U - XZ = \{5, 8\}$.

Fie $t, t' \in r$, astfel ıncat : $t[X] = t'[X]$, deci

$t_i = t'_i$, $i = 1, 4$. Sa aratam ca $t[Z'] = t'[Z']$, adica

$t_6 = t'_6$.

Consideram proiectiile uplelor t si t' pe X, Z,$T1$:

$t \to ((t_1 - t_4),(t_3, t_4, t_6, t_7),(t_5, t_8))$

$t' \to ((t'_1 - t'_4), (t'_3, t'_4, t'_6, t'_7), (t'_5, t'_8))$

Deoarece r satisface $X \to\to Z$, exista $t'' \in r$, astfel

ıncat proiectiile lui t'' pe X, Z, $T1$ sunt :

$t'' \to ((t_1 - t_4), (t_3, t_4, t_6, t_7), (t'_5, t'_8))$

Avem $t''_2 = t_2 = t'_2$ si $t''_5 = t'_5$, deci $t''[Y] = t'[Y]$.

Deoarece r satisface $Y \to Z'$, obtinem $t''[Z'] =$

$t'[Z']$, adica $t''_6 = t'_6$, dar $t''_6 = t_6$. Deci $t_6 = t'_6$.

Sa aratam acum FD-MVD3:

Presupunem ca r satisface $X \to\to Y$ si $XY \to Z$.

Aratam ca r satisface

$X \to Z - Y$.

$X = \{1, 2, 3, 4\}$

$Y = \{2, 3, 5, 6\}$

$Z = \{3, 4, 6, 7\}$

$Z - Y = \{4, 7\}$

$T1 = U - XY = \{7, 8\}$.

Fie $t, t' \in r$, astfel ıncat $t[X] = t'[X]$, adica

$t_i = t'_i$, $i = 1, 4$. Sa aratam ca

$t[Z - Y] = t'[Z - Y]$, adica $t_7 = t'_7$.

(Avem $t_4 = t'_4$).

Proiectam t si t' pe X, Y si $T1$:

$t \to ((t_1 - t_4), (t_2, t_3, t_5, t_6), (t_7, t_8))$

$t' \to ((t'_1 - t'_4), (t'_2, t'_3, t'_5, t'_6), (t'_7, t'_8))$

Deoarece r satisface $X \rightarrow\rightarrow Y$, exista $t'' \in r$,

astfel ıncat proiectiile lui t'' pe X, Y si T1 sunt:

$t'' \rightarrow ((t_1 - t_4), (t_2, t_3, t_5, t_6), (t'_7, t'_8))$

Avem $t''_i = t_i$, $i = 1, 6$.

Deoarece r satisface $XY \rightarrow Z$, rezulta $t''[Z] = t[Z]$,

de unde $t''_7 = t_7$. Dar $t''_7 = t'_7$, deci $t'_7 = t_7$.

Pentru fiecare proprietate a dependentelor multivaluate asociem o regula formala prin aceeasi metoda ca la dependentele functionale:

MVD0f: $XY\ Z = U$ si $Y \cap Z \subseteq X$, $X \to\to Y$

$$\frac{XYZ = U \text{ si } Y \cap Z \subseteq X,\ X \to\to Y}{X \to\to Z}$$

MVD1f: $Y \subseteq X$

$$X \to\to Y$$

MVD2f: $Z \subseteq W,\ X \to\to Y$

$$XW \to\to YZ$$

MVD3f: $X \to\to Y,\ Y \to\to Z$

$$X \to\to Z - Y$$

MVD4f: $X \to\to Y,\ YW \to\to Z$

$$XW \to\to Z - YW$$

MVD5f:

$$\frac{X \to\to Y,\ X \to\to Z}{X \to\to YZ}$$

MVD6f:

$$\frac{X \to\to Y,\ X \to\to Z}{X \to\to Y \cap Z,\ X \to\to Y - Z,\ X \to\to Z - Y}$$

FD-MVD1f: $X \rightarrow Y$

$$\frac{X \rightarrow Y}{X \rightarrow\rightarrow Y}$$

FD-MVD2f: $X \rightarrow\rightarrow Z,\ Y \rightarrow Z',\ Z' \subseteq Z,\ Y \cap Z = \emptyset$

$$\frac{X \rightarrow\rightarrow Z,\ Y \rightarrow Z',\ Z' \subseteq Z,\ Y \cap Z = \emptyset}{X \rightarrow Z'}$$

FD-MVD3f: $$\frac{X \rightarrow\rightarrow Y,\ XY \rightarrow Z}{X \rightarrow Z - Y}$$

Propozitia 3. Regulile de inferenta enuntate mai sus sunt valide.

Demonstratie. Rezulta imediat din propozitia 2.

Propozitia 4. Fie $\mathscr{R}$ o multime de reguli valide si o regula: $\sigma_1, \ldots, \sigma_k$

σ ,

astfel ıncat $\{ \sigma_1, . . . ,\sigma_k \} \vdash_{\mathscr{R}} \sigma$, atunci si regula este valida.

Afirmatia rezulta usor prin inductie dupa lungimea demonstratiei ın $\{ \sigma_1, . . . ,\sigma_k \}$ utilizand $\mathscr{R}$. Faptul ca $\{ \sigma_1, . . . ,\sigma_k \} \vdash_{\mathscr{R}} \sigma$ ıl vom numi: “regula se exprima cu ajutorul regulilor de inferenta din $\mathscr{R}$”. In continuare, vom considera ın afara de regulile de inferenta de mai sus si regulile de inferenta FD1f, FD2f, FD3f pentru dependentele functionale.

Propozitia 5 Fie $\mathscr{R}_{FM}$ = { FD1f-FD3f, MVD0f-MVD3f, FD-MVD1f- FD-MVD3f }. Avem:

FD-MVD3f se exprima prin celelalte reguli din $\mathscr{R}_{FM}$ si FD-MVD2f se exprima prin celelalte reguli din $\mathscr{R}_{FM}$.

Demonstratie. Fie σ_1: $X \to\to Y$ si $\sigma_2 : XY \to Z$.

Aplicam la prima MVD0f obtinem

σ_3: $X \to\to U - XY$. Din $XY \to Z$ si $Z \to Z - XY$

(obtinuta din FD1f) prin FD3f rezulta

$\sigma_4 : XY \to Z - XY$. Deoarece $Z - XY \subseteq U - XY$

si $XY \cap (U - XY) = \emptyset$, putem aplica FD-MVD2f

pentru σ_3 si σ_4 si obtinem: $\sigma_5 : X \to Z - XY$.

Dupa FD1f avem σ_6: $X \to X \cap Z - Y$.

Aplicand FD5f care se exprima cu ajutorul regulilor FD1f-FD3f (Propozitia 1.3 Cap. II) rezulta $\sigma_7 : X \to Z - Y$, adica FD-MVD3f se exprima prin celelalte reguli din $\mathscr{R}_{FM}$.

Fie date $\sigma_1: X \to\to Z$, $\sigma_2 : Y \to Z'$ cu conditiile $Z' \subseteq Z$ si $Y \cap Z = \emptyset$.

Aplicand MVD0f lui σ_1 obtinem:

$\sigma_3: X \to\to U - XZ$.

Deoarece $Y \subseteq X(U - XZ)$ prin FD1f obtinem:

σ_4: $X(U - XZ) \rightarrow Y$. Aplicand FD3f pentru σ_4 si σ_2 se obtine σ_5: $X(U - XZ) \rightarrow Z'$.

Putem aplica regula FD-MVD3f pentru σ_3 si σ_5, ceea ce conduce la σ_6: $X \rightarrow Z' - (U - XZ)$.

Dar $Z' - (U - XZ) = Z'$. Deci σ_6: $X \rightarrow Z'$.

In concluzie, regula FD-MVD2f se exprima prin celelalte reguli din $\mathscr{R}_{FM}$.

Propozitia 6. Regulile MVD4f-MVD6f se exprima cu ajutorul regulilor MVD0f-MVD3f.

Demonstratie. Fie σ_1: $X \rightarrow\rightarrow Y$ si σ_2: $YW \rightarrow\rightarrow Z$.

Aplicam pentru σ_1 si $W \subseteq W$ regula MVD2f si obtinem: σ_3: $XW \rightarrow\rightarrow YW$.

Pentru σ_3 si σ_2 aplicam MVD3f si se obtine:

σ_4: $XW \rightarrow\rightarrow Z - YW$.

Deci $\{\sigma_1, \sigma_2\} |\text{--}_{\{MVD2f,\ MVD3f\}} \sigma_4$

(regula pentru pseudotranzitivitate).

Consideram acum MVD5f. Fie σ_1: $X \rightarrow\rightarrow Y$ si σ_2 : $X \rightarrow\rightarrow Z$.

Din σ_1, $Z \subseteq Z$ si MVD2f se obtine

$\sigma_3 : XZ \rightarrow\rightarrow YZ$. Aplicand lui σ_3 regula MVD0f se obtine

σ_4: $XZ \to\to (U-XYZ)$. Din σ_2, $X \subseteq X$ si MVD2f

rezulta σ_5: $X \to\to XZ$. Din σ_5, σ_4 si regula MVD3f

rezulta σ_6: $X \to\to U - XYZ$.

Aplicand pentru σ_6 regula MVD0f rezulta

σ_7: $X \to\to YZ$.

Fie σ_1: $X \to\to Y$ si σ_2 : $X \to\to Z$.

Din $X \to\to X$ (reflexivitate)

si σ_1 prin MVD3f obtinem σ_3 : $X \rightarrow\rightarrow Y - X$.

Aplicand lui σ_3 regula MVD0f obtinem

σ_4: $X \rightarrow\rightarrow U - XY$.

Din $X \rightarrow\rightarrow X$ si $X \rightarrow\rightarrow Z$ prin MVD3f obtinem:

σ_5: $X \rightarrow\rightarrow Z - X$. Aplicand acesteia MVD0f

obtinem:

σ_6: $X \rightarrow\rightarrow U - XZ$. Aplicand acum MVD4f

(reuniunea) pentru σ_4 si σ_6 obtinem:

σ_7: $X \to\to (U-XY)(U-XZ)$. Aplicam acum MVD0f pentru σ_7, vom avea:

σ_8: $X \to\to Y \cap Z - X$. Prin MVD1f avem

σ_9: $X \to\to X \cap Y \cap Z$.

Prin reuniune (MVD4f) din σ_8 si σ_9 se obtine

σ_{10}: $X \to\to Y \cap Z$.

Astfel, pornind de la σ_1 si σ_2 si aplicand regulile MVD0f-MVD3f se obtine σ_{10}.

Sa aratam acum cea de a doua parte a lui MVD6f. Pornim de la σ_1: $X \rightarrow\rightarrow Y$ si σ_2 : $X \rightarrow\rightarrow Z$. Aplicand MVD5f se obtine $X \rightarrow\rightarrow YZ$.

De aici, prin MVD0f, se obtine:

σ_3: $X \rightarrow\rightarrow U - XYZ$. Aplicam MVD5f pentru σ_2 si σ_3 ceea ce produce:

σ_4: $X \rightarrow\rightarrow Z(U - XYZ)$. Pentru σ_4 aplicam MVD0f, ceea ce conduce la

σ_5: $X \twoheadrightarrow Y - XZ$. Prin reflexivitate avem σ_6 : $X \twoheadrightarrow X \cap Y - Z$. Prin MVD5f din σ_5 si σ_6 se obtine: σ_7 : $X \twoheadrightarrow Y - Z$. In mod similar, se obtine $X \twoheadrightarrow Z - Y$ (schimband ın fond Y cu Z peste tot).

Teorema 1. Fie Σ o multime de dependente functionale sau multivaluate si X o submultime de atribute. Atunci exista o partitie a lui $U - X$ notate prin $\{Y_1, \ldots, Y_k\}$, astfel ıncat pentru $Z \subseteq U - X$ avem $\Sigma \vdash_{\mathscr{R}_{FM}} X \twoheadrightarrow Z$ iff Z este reuniunea unui numar de multimi din partitia $\{Y_1, \ldots, Y_k\}$.

Demonstratie. Construim partitia notata P, astfel: initial consideram ın P numai $U - X$. Fie P obtinuta la un moment dat avand elementele $W_1, \ldots, W_n$.

Presupunem ca $\Sigma|\text{--}_{\mathscr{R}_{FM}}\ X \rightarrow\rightarrow W_i$, pentru orice $i = 1, n$ (initial $\Sigma \vdash X \rightarrow\rightarrow U - X$ dupa MVD0f si MVD1f).

Fie $\Sigma|\text{--}_{\mathscr{R}_{FM}}$ $X \to\to Z$ si $Z \subseteq U - X$ si Z nu este reuniune de multimi W_i.

Deoarece P este partitie pentru $U - X$, rezulta ca exista W_i din P, astfel ıncat $W_i \cap Z \neq \emptyset$ si $W_i - Z \neq \emptyset$. Pentru fiecare astfel de W_i din P ınlocuim ın P pe W_i cu $W_i \cap Z$ si $Wi - Z$. Deoarece $\Sigma|\text{--}_{\mathscr{R}_{FM}}$ $X \to\to Z$ si dupa ipoteza inductiei $\Sigma|\!\!-\!\!-_{\mathscr{R}_{FM}} X \rightarrow\rightarrow W_i$, aplicand MVD6f se obtine $\Sigma|\!\!-\!\!-_{\mathscr{R}_{FM}} X \rightarrow\rightarrow W_i \cap Z$ si $\Sigma|\!\!-\!\!-_{\mathscr{R}_{FM}} X \rightarrow\rightarrow W_i - Z$.

Altfel spus, noua partitie satisface aceeasi proprietate ca vechea partitie.

Deoarece U este finita si multimea dependentelor functionale sau multivaluate este finita, rezulta ca algoritmul de mai sus este finit. (Numarul partitiilor lui $U - X$ este de asemenea finit).

Fie $P = \{Y_1, . . . , Y_k\}$ partitia finala obtinuta.

Rezulta prin inductie

dupa pasii folositi ın constructia lui P ca:

$\Sigma |\text{--}_{\mathscr{R}_{FM}} X \rightarrow\rightarrow Y_i$, $i = 1, k$.

Daca $Z \subseteq U - X$ este reuniune de Y_i, adica $Z = Y_{i1} \cup \ldots \cup Y_{ih}$, aplicand MVD5f se obtine $\Sigma \vdash_{\mathscr{R}_{FM}} X \twoheadrightarrow Z$.

Invers. Pentru $Z \subseteq U - X$, daca $\Sigma \vdash_{\mathscr{R}_{FM}} X \twoheadrightarrow Z$, atunci Z este reuniune de Y_i, pentru ca altfel s-ar putea rafina partitia P, ceea ce este o contradictie.

Definitia 3. Pentru Σ o multime de dependente functionale sau multivaluate si X o submultime de atribute, numim baza de dependenta pentru X cu privire la Σ, partitia $B(\Sigma,X) = \{\{A_1\} \ldots \{A_h\}, Y_1, \ldots, Y_k\}$, unde

$X = A_1 \ldots A_h$, iar $Y_1 \ldots Y_k$ este partitia construita ın teorema 1.

Observatia 4. Avem $\Sigma|\text{--}_{\mathscr{R}_{FM}} X \to\to Z$ iff Z este o reuniune de elemente din partitia $B(\Sigma,X)$.

In adevar, daca Z este reuniune de elemente din $B(\Sigma,X)$, atunci fie

$Z = A_{i1} \ldots A_{it} \cup Y_{j1} \cup \ldots \cup Y_{jl}$.

Avem : $\Sigma|\text{--}_{\mathscr{R}_{FM}} X \to\to A_{ip}$, $p = 1, t$ dupa MVD1f si $\Sigma|\text{--}_{\mathscr{R}_{FM}} X \to\to Y_{jq}$, $q = 1, l$ dupa teorema 1.

Aplicand acestora MVD5f se obtine

$\Sigma|\text{--}_{\mathscr{R}\mathrm{FM}}\ X \to\to Z.$

Invers, presupunem c˘a avem $\Sigma|\mathscr{R}\mathrm{FM}$

$X \to\to Z$. Fie $Z = X1 \cup Z1$,

unde $X1 \subseteq X$ si $Z1 \subseteq U - X$. $(X1 \cap Z1 = \emptyset)$.

Dup˘a MVD6f rezulta

$\Sigma|\mathscr{R}_{\mathrm{FM}}\ X \to\to Z1$. De aici dupa teorema 1,

$Z1 = Y_{j1} \cup \ldots \cup Y_{jl}$.

Daca $X1 = A_{i1} \ldots A_{it}$, atunci Z este reuniunea elementelor $A_{i1} \ldots A_{it}$, $Y_{j1} \ldots Y_{jl}$ din $B(\Sigma,X)$.

Observatia 5. Fie $X_\Sigma = \{A|\Sigma|\mathscr{R}_{FM} \; X \to A\}$.

Atunci pentru orice $A \in X*_\Sigma$ avem $\{A\} \; B(\Sigma,X)$.

In adevar, pentru $A \in X*_\Sigma$ dupa FD-MVD1f, obtinem $\Sigma|\mathscr{R}_{FM} \; X \to\to A$ si aplicand teorema 1, rezulta $\{A\} \in B(\Sigma,X)$.

1 Studiul dependentelor functionale si multivaluate utilizand calculul propozitional

Pentru fiecare atribut $A \in U$ asociem o variabila propozitionala notata a. Pentru o dependenta functionala σ: $A_1 \ldots A_m \to B_1 \ldots B_h$ asociem formula (numita implicatie) $\sigma : a_1 \ldots a_m \Rightarrow b_1 \ldots b_h$ (ca ın cazul studiului dependentelor functionale, utilizand calculul propozitional, capitolul II, §2).

Daca $\sigma : A_1 \ldots A_m \rightarrow\rightarrow B_1 \ldots B_h$ este o dependenta multivaluata

si $U - A_1 \ldots A_m B_1 \ldots B_h = C_1 \ldots C_p$, atunci formula asociata lui σ va fi

σ: $a_1 \ldots a_m \Rightarrow b_1 \ldots b_h + c_1 \ldots c_p$.

Semnul + noteaza disjunctia logica.

Daca $\sigma : X \rightarrow Y$, atunci _ se noteaza si prin $x \Rightarrow y$. Daca $\sigma : X \rightarrow\rightarrow Y$, atunci _ se noteaza si prin $x \Rightarrow\Rightarrow y$.

Multimile X, Y si $Z = U - XY$ pot fi vide. Facem conventia ca pentru multimea vida $\varnothing$, formula asociata are valoarea true (conjunctia unei multimi vide de variabile propozitionale este true).

Propozitia 1.1 $x \Rightarrow\Rightarrow y$ este true iff $x \Rightarrow\Rightarrow y - x$ este true.

Demonstratie. Daca $X = A_1 \ldots A_m$, $Y = B_1 \ldots B_h$ si $Z = U - XY = C_1 \ldots C_p$, atunci $x \Rightarrow\Rightarrow y$ este

$a_1 \ldots a_m \Rightarrow b_1 \ldots b_h + c_1 \ldots c_p$.

Fie $Y - X = B_1 \ldots B_t$ si $Y \cap X = B_{t+1} \ldots B_h$.

Daca δ este o asignare, astfel ıncat

$\delta(x \Rightarrow\Rightarrow y) = true$, atunci putem avea:

a) exista i, $i \in \{1, 2, \ldots, m\}$, astfel ıncat $\delta(a_i) =$ false.

In acest caz $\delta(x \Rightarrow y - x) = true$.

b) $\forall i$, $i \in \{1, 2, \ldots m\}$, $\delta(a_i) = true$.

De aici $\delta(a_1 \ldots a_m) = true$.

Rezulta $\delta(b_1 \ldots b_h) = true$ sau $\delta(c_1 \ldots c_p) = true$.

b1) $\delta(b_1 \ldots b_h) = true$ implica $\delta(b_1 \ldots b_t) = true$, deci $\delta(x \Rightarrow (y - x) + c_1 \ldots c_p) = true$.

b2) $\delta(c_1 \ldots c_p) = true$ atunci $\delta(x \Rightarrow\Rightarrow y - x) = true$.

Invers rezulta similar.

Pentru simplitatea scrierii vom considera valoarea de adevar true notata prin 1, iar false prin 0.

Ca si ın cazul dependentelor functionale, intentia noastra este de a stabili o legatura ıntre notiunea de consecinta din domeniul dependentelor functionale si multivaluate si notiunea de consecinta logica din calculul propozitional.

Exemplul 1.1 Fie $U = \{A,B,C,D\}$, si $\Sigma = \{A \to\to B, C \to B\}$ si

$\sigma: A \to B$. Atunci _ : $a \Rightarrow b$,

$\Sigma = \{a \Rightarrow b + cd, c \Rightarrow b\}$.

Aratam ca $\Sigma \models$ _. In adevar, fie r o relatie ce satisface dependentele

$A \to\to B$ si $C \to B$. Sa aratam ca r satisface $A \to B$. Fie $t_1, t_2 \in r$ cu $t_1[A] = t_2[A]$ si fie

$t_1 = (a1, b1, c1, d1)$ si $t_2 = (a2, b2, c2, d2)$.

Deoarece r satisface $A \to\to B$, rezulta ca exista $t_3, t_4 \in r$, astfel ıncat $t_3 = (a1, b1, c2, d2)$ si $t_4 = (a1, b2, c1, d1)$. Deoarece $t1, t4 \in r$ si r satisface $C \to B$, rezulta ca

$b1 = b2$, adica $t_1[B] = t_2[B]$, deci r satisface $A \rightarrow B$. Aratam acum ca $\Sigma \models_{c.l.}$ _.

Fie δ o asignare, astfel ıncat $\delta(a \Rightarrow b + cd) = 1$ si $\delta(c \Rightarrow b) = 1$.

Sa aratam ca $\delta(a \Rightarrow b) = 1$.

Daca $\delta(a) = 0$, atunci am terminat.

Daca $\delta(a) = 1$, atunci $\delta(b+cd) = 1$, deci $\delta(b) = 1$ sau $\delta(cd) = 1$; ın cazul $\delta(b) = 1$ am terminat.

In cazul $\delta(cd) = 1$ rezulta $\delta(c) = 1$ si cu $\delta(c \Rightarrow b) = 1$ se obtine $\delta(b) = 1$, deci iarasi $\delta(a \Rightarrow b) = 1$.

Teorema 1.1 Teorema de echivalenta.

Fie Σ o multime de dependente functionale sau multivaluate si σ o dependenta functionala sau multivaluata.

Urmatoarele afirmatii sunt echivalente:

a) σ este o consecinta a lui Σ.

b) σ este o consecinta a lui Σ pe multimea relatiilor cu 2 uple.

c)   este consecinta logica a lui Σ.

Vom da ıntai o demonstratie sintactica a teoremei de echivalenta. Pentru aceasta vom considera reguli de inferenta pentru implicatii ce se construiesc

pornind de la regulile de inferenta din $\mathscr{R}$FM =
{FD1f-FD3f, MVD0f-MVD3f, FD-MVD1f-FD-MVD3f}

In capitolul II am considerat regulile de inferenta A1′, A2′, A3′ pentru
implicatii din calculul propozitional, reguli ce corespund axiomelor lui Armstrong.

Enuntam regulile de inferenta asociate celor din $\mathscr{R}$FM:

FD1′: $y \subseteq x$

$$\frac{y \subseteq x}{x \Rightarrow y}$$

$y \subseteq x$ noteaza faptul ca orice variabila ce apare ın y, apare de asemenea ın x.

FD2′: $x \Rightarrow y,\ z \subseteq w$

$xw \Rightarrow yz$

FD3′: $x \Rightarrow y,\ y \Rightarrow z$

$x \Rightarrow z$

MVD0′: $xyz = u,\ y \cap z \subseteq x,\ x \Rightarrow\Rightarrow y$

$x \Rightarrow\Rightarrow z$

u este conjunctia tuturor variabilelor asociate lui U.

MVD1′: $y \subseteq x$

$x \Rightarrow\Rightarrow y$

MVD2′: $z \subseteq w,\ x \Rightarrow\Rightarrow y$

$xw \Rightarrow\Rightarrow yz$

MVD3′: $x \Rightarrow\Rightarrow y,\ y \Rightarrow\Rightarrow z$

$x \Rightarrow\Rightarrow z - y$

FD-MVD1′: $x \Rightarrow y$

$x \Rightarrow\Rightarrow y$

FD-MVD2′: $x \Rightarrow\Rightarrow z,\ y \Rightarrow z',\ z' \subseteq z,\ y \cap z = \varnothing$

$$x \Rightarrow z'$$

FD-MVD3′: $x \Rightarrow\Rightarrow y,\ xy \Rightarrow z$

$$\frac{\qquad\qquad\qquad\qquad}{x \Rightarrow z - y}$$

Observatia 1.2 Deoarece sistemul de reguli {A1′, A2′, A3′} este valid si {A1′, A2′, A3′} este echivalent cu {FD1′,FD2′,FD3′} (ın virtutea faptului ca $\{A1, A2, A3\}$ este echivalent cu $\{FD1f, FD2f, FD3f\}$ si $\Sigma\text{--}_{\{FD1f\text{-}FD3f\}}\ \sigma$ iff $\Sigma\text{--}|_{\{A1,A2,A3\}}\ \sigma$) rezulta ca $FD1'$, $FD2'$, $FD3'$ sunt valide. Aceasta afirmatie rezulta desigur si direct.

Observatia 1.3 In virtutea propozitiei 5 ne vom dispensa de una din regulile FD-MVD2′ sau FD-MVD3′. Vom renunta la ultima.

Lema 1.4 Regulile de inferenta FD1′-FD3′, MVD0′-MVD3′, FD-MVD1′,FD-MVD2′ sunt valide. Fie $\mathscr{R}$FM′ multimea acestor reguli.

Demonstratie. Primele 3 sunt valide dupa observatia 1.2.

MVD0′ este valida: Consideram reprezentarea multimilor de variabile

din $x, y, z, u - xy, u - xz$:

Din ipotezele respective se obtine: 6 = $\emptyset$ si 8 = $\emptyset$.

Fie t1 = u − xy si t2 = u − xz.

Sa notam prin i conjunctia variabilelor din domeniul i, i = 1, 5, 7.

Formula x ⇒⇒ y se scrie astfel: 1 2 3 4 ⇒ 2 3 5 + 7, iar x ⇒⇒ z devine

1 2 3 4 ⇒ 3 4 7 + 5.

Fie δ o asignare astfel ıncat $\delta(x \Rightarrow\Rightarrow y) = 1$.

Daca $\delta(1\ 2\ 3\ 4) = 0$, atunci $\delta(x \Rightarrow\Rightarrow z) = 1$.

Daca $\delta(1\ 2\ 3\ 4) = 1$, atunci $\delta(2\ 3\ 5) = 1$ sau $\delta(7) = 1$.

Cand avem $\delta(2\ 3\ 5) = 1$, atunci $\delta(5) = 1$, deci $\delta(x \Rightarrow\Rightarrow z) = 1$.

Cand avem $\delta(7) = 1$, atunci $\delta(3\ 4\ 7) = 1$, deci $\delta(x \Rightarrow\Rightarrow z) = 1$.

Observatia 1.4 Fie Σ o multime de formule din calculul propozitional si Σ^+ multimea formulelor ce pot fi derivate din Σ, utilizand regulile de inferenta FD1′-FD3′, MVD0′- MVD3′, FD-MVD1′, FD-MVD2′.

Atunci dupa lema 1.4 rezulta ca $\Sigma \models_{c.l.} \Sigma^+$, adica formulele din Σ^+ sunt consecinte

logice ale formulelor din Σ.

Lema 1.5 Fie Σ o multime de formule asociate multimii Σ de dependente functionale sau multivaluate si X o multime de atribute.

Fie $X^+ = \{A | \Sigma \text{--}_{FM} X \rightarrow A\}$. Fie $B(\Sigma,X)$ baza de dependenta pentru X cu privire la Σ si $W \in B(\Sigma,X)$, astfel ıncat $W \cap X+ = \emptyset$. Consideram asignarea δ_0 definita astfel: $\delta_0(a) = 0$ iff $A \in W$.

Atunci avem: $\delta_0(\_) = 1$ pentru $\_ \in \Sigma$.

Teorema 1.2 (Teorema de completitudine pentru formule).

Fie Σ multimea de formule asociate multimii Σ de dependente functionale sau multivaluate si _ formula asociata dependentei _ (functionala sau multivaluata). Atunci _ este consecinta logica a lui Σ daca si numai daca _ poate fi demonstrata ın Σ utilizand regulile de inferenta

$\mathscr{R}$FM′. Pe scurt: Σ $|=_{c.l.}$ _ iff Σ|FM′ _.

In continuare vom da demonstratia sintactica a teoremei de echivalenta.

Teorema 1.3 (Teorema de echivalenta). Fie Σ o multime de dependente functionale sau multivaluate si _ o dependenta functionala sau multivaluata.

Atunci sunt echivalente urmatoarele afirmatii:

a) _ este o consecinta a lui Σ.

b) _ este o consecinta a lui Σ pe domeniul relatiilor cu 2 uple.

c) _ este o consecina logica a lui Σ.

Demonstratie.

a) implica b) rezulta imediat.

Aratam ca b) implica c). Fie b) adevarata si presupunem c) falsa.

Atunci exista o asignare δ, astfel ıncat δ(_) = 1 pentru orice _ $\in \Sigma$ si δ(_) = 0.

Consideram relatia r cu 2 uple t_1 si t_2 definite astfel:

$t_1[A] = 1$ pentru orice $A \in U$, $t_2[A] = 1$ iff $\delta(a) = 1$.

Dupa lema 2.5 obtinem:

r satisface _, orice _ $\in \Sigma$ si r nu satisface _,

ceea ce contrazice b).

Aratam acum: c) implica a). Fie _ consecinta logica a lui Σ. Dupa teorema 1.1 rezulta ca $\Sigma|\mathscr{R}'FM$ _, de unde $\Sigma|\mathscr{R}FM$ _.

Avem atunci Σ |= _,

deoarece regulile din $\mathscr{R}$FM sunt valide.