Outside and Inside of First Normal Form

Prof.dr. Marin FOTACHE
Catedra de Informatică Economică, Universitatea „Alexandru Ioan Cuza”, Iaşi

First Normal Form (1NF) is the only normal form required for a database to be accepted as relational. Further normalization - 2NF, 3NF, BCNF, 4NF, 5NF, DKNF - is not included in the relational model theory, but it has been considered as a necessary process for diminishing the storage space and for eliminating or reducing the anomalies in insert-update-delete operations. The paradox with 1FN is that it is simultaneously easy to define but not clear to explain in depth. The accepted perspectives on 1FN were seen as largely restrictive, mainly because of atomicity (scalability) of the domains. Recent contributions suggests that 1FN can be seen in a free manner, so that relational model could deal with complex data, especially with objects.
Keywords: *normalization, first normal form, repetitive groups, domain, type, relation type attribute.*

Introducere

Teoria clasică a normalizării este construită în jurul a cinci forme normale. Codd, părintele modelului relaţional, a definit iniţial trei forme normale, notate prin 1NF, 2NF şi 3NF. Întrucât, într-o primă formulare, definiţia 3FN ridica ceva probleme, Codd şi Boyce au elaborat o nouă variantă, cunoscută sub numele Boyce-Codd Normal Form (BCNF). Deşi este vorba, în principiu, de o formulare mai riguroasă a aceleaşi 3FN, BCNF este prezentată separat în majoritatea lucrărilor. Formele 4 şi 5 sunt legate de numele lui Ronald Fagin care, de asemenea, a definit şi o formală normală domenii/cheie. Dintre acestea, doar prima formă este obligatorie pentru o bază de date relaţională, celelalte urmărind "doar" să diminueze risipa de spaţiu de stocare şi să elimine anomaliile manifestate la inserare, modificare şi ştergere de înregistrări.

Atomicitate

Prima formă normală este, fără îndoială, un ciudat amestec de facil şi confuz. Putem spune chiar că mare parte din doza de facil, de superficialitate din cursurile universitare, cărţile dedicate bazelor de date sau chiar proiectării bazelor de date (inclusiv normalizării) îşi au sorgintea tocmai în confuzia care a învăluit, încă de la începuturile relaţionalului, definirea prime forme normale şi atomicităţii (există lucrări respectabile, precum *Atzeni s.a. 99*, care nici nu fac aluzie la prima formă normală sau atomicitatea valorilor într-o relaţie). Unii autori, precum Elmasri şi Navathe, consideră că restricţia de atomicitate este proprie modelului relaţional, restricţie care este eliminată din modelele relaţional imbricat (*nested relational model*) şi obiectual-relaţional ce permit relaţii nenormalizate (Elmasri & Navathe 00, p. 485).

Ceea ce se înţelege îndeobşte prin atomicitatea atributelor ţine de caracterul scalar al valorilor: întregi, şiruri de caractere, date calendaristice. Prin comparaţie, un atribut nonatomic este unul definit pe un domeniu de valori compozite, complexe. Fiecare componentă dintr-o valoare compozită poate fi, la rândul ei, compozită, ajungându-se la o structură ierarhică şi flexibilă, în funcţie de natura obiectului sau procesului modelat. Elmasri şi Navathe sunt cât se poate de limpezi: pentru atributele unei relaţii în 1FN, singurele valori permise sunt cele atomice sau indivizibile (Elmasri & Navathe 00, p. 485).

În ceea ce mă priveşte, materialul favorit legat de 1FN este cel al lui Chris Date (Date03) publicat în iunie 2003 pe site-ul http://www.dbdebunk. com, site conceput şi gestionat de un apropiat al lui Date şi un necruţător observator al ignoranţei comunităţii IT (şi academice) în materie de baze de date - Fabian Pascal. Materialul evocat constituie deopotrivă încununarea şi convergenţa lucră-

rilor scrise singur (cea mai celebră este cartea a cărei a opta ediție a fost tradusă şi în româneşte - Date04) sau împreună cu Hugh Darwen (vezi *The Third Manifesto*). Meritoriu este şi faptul că, din capul locului, Date se autodenunță ca fiind unul dintre cei care, de-a lungul timpului, au contribuit la confuzia în care se scaldă atomicitatea.

Prima definiție a 1FN aparține, firește, lui E.F. Codd pentru care o relație este în 1FN dacă nici unul dintre domeniile sale nu are elemente (valori) de tip set (Codd72. Preluare din Date03). Câțiva ani mai târziu, Codd afirmă că un domeniu este simplu dacă toate valorile sale sunt atomice, adică nedecompozabile de către SGBD (Codd79). Date identifică în cartea lui Codd publicată în 1990 (Codd90) două afirmații similare:

- valorile din domeniile pe care fiecare relație este definită sunt necesarmente atomice vis-a-vis de SGBD;
- datele atomice sunt cele care nu pot fi descompuse de către SGBD în subcomponente (cu excepția unor funcții speciale).

Din păcate, Codd elimină o neclaritate înlocuind-o cu alta. Şi Date se întreabă: ce înseamnă *cu excepția unor funcții speciale* ? Faptul că dintr-un şir de caractere (exemplul clasic de valoare atomică) pot fi extrase diferite componente prin funcții precum SUBSTR, funcții prezente în mai toate SGBD-urile actuale, face din valoarea de tip şir de caractere una neatomică? Sună cam hilar. Mai ales că, dacă raportăm atomicitatea la SGBD-uri, deseori vom fi în situația în care o valoare atomică într-o bază de date gestionată cu un SGBD să fie neatomică în cazul unei aceleaşi baze de date (structuri) implementată pe un alt SGBD. Ceea ce, să recunoaştem, nu e prea onorant pentru un model atât de riguros fundamentat aşa cum este relaționalul.

Date, împreună cu Darwen şi Pascal propun renunțarea la atomicitate ca o cerință a 1NF. După Date, atomicitatea nici nu are un semnificație absolută, întrucât depinde de ceea ce dorim să facem cu datele asupra atomicității cărora ne pronunțăm (Date03). Practic, toate tabelele care respectă principiile modelului relațional sunt în 1FN, iar această primă normală păstrează din relevanță doar în două privințe:

- indică faptul că relația la care se referă poate să nu fie într-o formă normalizată superioară (2NF, 3FN...);
- structură de date ce respectă 1FN este una relațională, cu alte cuvinte tabelele care nu sunt în 1FN sunt non-relaționale.

Acest din urmă aspect merită o discuție detaliată, deoarece ține de aspectul structural al modelului relațional.

Renunțarea la atomicitate nu este singura schimbare fundamentală în modelul relațional. Domeniile pot conține orice tip de valori: scalare sau compozite, şiruri de caractere, numere, dar şi vectori, liste, imagini, înregistrări audio/video, documente XML sau orice tip de definit de utilizatori. Dealtminteri, dacă inițial Codd s-a străduit să impună sintagma *domeniu* în detrimentul *tipului*, tocmai pentru a opera o distrincție netă dintre baze de date şi limbaje de programare, Date, Darwen şi Pascal sunt, în ultimul timp, mult mai apropiați de noțiunea de *tip*, pe care o consideră mai sugestivă şi, în plus, identic aplicabilă şi altor modele de organizare a datelor date. Atâta vreme cât există suport din partea SGBD-ului pentru definirea, stocarea şi accesarea sa, se poate folosi orice tip de date într-o BD.

Să luăm exemplul tabelei BIBLIOTECĂ din figura 1. Tabela gestionează informații despre cărțile existente în depozitul bibliotecii Facultății de Economie şi Administrarea Afacerilor (FEAA). De remarcat că în bliblioteca există două exemplare ale cărții cu ISBN-ul 973-683-709-2 şi trei exemplare de cea dedicată Visual FoxPro. Prima carte a fost scrisă de patru autori şi îi sunt asociate opt cuvinte-cheie, iar a doua are un singur autor. Relația nenormalizată conține trei grupuri repetitive: Cote, Autori şi CuvinteCheie.

Pe baza tipului şir de caractere, putem defini un domeniu de tip vectori de şiruri de caractere, iar cele trei atribute, Cote, Autori şi CuvinteCheie se vor declara pe acest nou domeniu de tip set. Chiar dacă noua relație, denumită BIBLIOTECĂ_SET (figura 2) seamănă izbitor cu forma denormalizată, este vorba de o schemă radical diferită.

BIBLIOTECĂ_SET

ISBN	Titlu	Cote_SET	Autori_SET
973-683-889-7	Visual FoxPro. Ghidul dezvoltării aplicațiilor profesionale	{III-13421,III-13422, III-13423}	{Marin Fotache, Ioan Brava, Cătălin Strâmbei, Liviu Crețu}
973-683-709-2	SQL. Dialecte DB2, Oracle şi Visual FoxPro	{III-10678, III-10679}	{Marin Fotache}

Editura	LocSediuEd	AnApariție	CuvinteCheie_SET
Polirom	Iaşi	2002	{baze de date, SQL, proceduri stocate, FoxPro, formulare, orientare pe obiecte, client-server, web}
Polirom	Iaşi	2001	{baze de date, algebră relațională, SQL}

Fig. 1. Relația universală nenormalizată BIBLIOTECĂ

BIBLIOTECĂ_SET

ISBN	Titlu	Cote_SET	Autori_SET
973-683-889-7	Visual FoxPro. Ghidul dezvoltării aplicațiilor profesionale	{III-13421,III-13422, III-13423}	{Marin Fotache, Ioan Brava, Cătălin Strâmbei, Liviu Crețu}
973-683-709-2	SQL. Dialecte DB2, Oracle şi Visual FoxPro	{III-10678, III-10679}	{Marin Fotache}

Editura	LocSediuEd	AnApariție	CuvinteCheie_SET
Polirom	Iaşi	2002	{baze de date, SQL, proceduri stocate, FoxPro, formulare, orientare pe obiecte, client-server, web}
Polirom	Iaşi	2001	{baze de date, algebră relațională, SQL}

Fig. 2. Domenii de tip set

Atribute de tip relații

A doua inovație o constituie atributele ale căror valori pot fi chiar relații (*relation-valued attributes*) - ATR. În exemplul nostru, în locul celor trei seturi putem folosi câte o relație. Câştigul este important, deoarece ingredientul folosit este pur relațional (relația !), iar mecanismul de declarare a restricțiilor şi cel de interogare (algebra relațională) este, în esență, acelaşi. Atributele Cote_REL, Autori_REL şi CuvinteCheie_REL din relația BIBLIOTECĂ_ATR (figura 3) sunt de acest tip, iar valorile lor pentru fi supuse operatorilor clasici: selecție, proiecție, joncțiune etc. Date şi Harwen propun câțiva noi operatori algebrici relaționali pentru a asigura comparabilitatea tabelelor cu şi fără ATR, GROUP şi UNGROUP (Date & Darwen00).

Din contră, dacă am folosi prima variantă - cea a seturilor - ar fi necesari operatori speciali: reuniune de seturi, intersecție de seturi etc.

Această a doua inovație a modelului relațional este o consecință directă a primeia. Practic, dacă un atribut X este de tip relație, iar D este domeniul lui X, atunci toate valorile lui D sunt relații. În plus, orice atribut de tip relație poate conține atribute care sunt, la rândul lor, de tip relații, astfel încât numărul nivelurilor de imbricare este nelimitat.

Interesant este că 1984 Serge Abiteboul şi Nicole Bidoit au publicat o lucrare ce poate fi considerată drept "antemergătoare" (Abiteboul & Bidoit 84). Deşi autorii afirmau că propun un model de date nou (l-au botezat Verso), în mare, ideea era ca valoarea unui atribut să fie nu numai atomică, ci şi o instanță a unui format (autorii au ezitat să folosească termenul relație). Ceea ce este notabil la modelul propus este că prin acest mecanism era obținută o ierarhie ce putea fi interogată folosind aceeaşi operatori algebrici relaționali. Cu atât mai mult cu cât valorile nule aveau o formă de reprezentare care ar mulțumi astăzi pe Date şi compania (relații vide).

Howard Dreizen şi Shi-Huo Chang propun în numărul din decembrie 1989 al *ACM Transactions on Database Systems* acceptarea, cu titlu restrictiv, a unor condiții excepționale în schema bazei de date pentru a rezolva o serie de situații practice relativ rare care însă crează anomalii. Cu acest prilej, autorii se pronunță pentru folosirea relațiilor incluse, altfel spus, includerea relațiilor în alte relații, de o manieră nerestrictivă; astfel, o relație R (D1, D2, ..., Dn) de n domenii este recursivă omogen dacă fiecare domeniu este fie (1) un set de valori atomice, fie (2) un set de relații recursive omegen cu scheme identice (Dreizen & Chang 89).

BIBLIOTECĂ_ATR

ISBN	Titlu	Cote_REL
973-683-889-7	Visual FoxPro. Ghidul dezvoltării aplicațiilor profesionale	Cote: III-13421, III-13422, III-13423
973-683-709-2	SQL. Dialecte DB2, Oracle şi Visual FoxPro	Cote: III-10678, III-10679

Autori_REL	Editura	LocSediuEd	AnApariție	CuvinteCheie_REL
Autori: Marin Fotache, Ioan Brava, Cătălin Strâmbei, Liviu Crețu	Polirom	Iaşi	2002	CuvinteCheie: baze de date, SQL, proceduri stocate, FoxPro, formulare, orientare pe obiecte, client-server, web
Autori: Marin Fotache	Polirom	Iaşi	2001	CuvinteCheie: baze de date, algebră relațională, stocate

Fig. 3. Atribute de tip relații (ATR)

Odată trecută euforia contactului cu atributele de tip relații, ne putem întreba retoric, precum Date: în proiectarea schemei bazei chiar avem nevoie de ATR-uri ? Întrebarea corectă ar fi, mai degrabă: când ATR-urile sunt mai avantajoase, prin comparație cu cele atomice (şi mecanismul de normalizare expus în acest capitol) ? Principalul avantaj ține, după cum am discutat, de enorma lor flexibilitate. Dintre limite sau, după caz, dezavantaje, ar merita început cu structura ierarhică a ATR, ca şi în lumea orientării pe obiecte, în general. Nu întotdeauna realitatea este ierarhică. Relațiile dintre obiecte pot fi cu mult mai complexe, iar de cele mai multe ori sunt necesare regrupări, resistematizări ale datelor, situații în care stângăciile structurilor ierarhice se manifestă în toată splendoarea lor.

Să luăm discuție relația STUDENȚI_EXAMENE din figura 4 pentru care încercăm să valorificăm ATR-urile.

STUDENȚI_EXAMENE

Matricol	NumePrenume	An	Specializare	CodDisc
EL13455	Popovici I Vasile	3	Informatică economică	AI3501
EL13456	Zăineanu W Ion	3	Informatică economică	AI3501
EL13457	Abălaşei R Zicu	3	Informatică economică	AI3501
EL13455	Popovici I Vasile	3	Informatică economică	AI3502
EL13456	Zăineanu W Ion	3	Informatică economică	AI3502
EL13457	Abălaşei R Zicu	3	Informatică economică	AI3502
EL13456	Zăineanu W Ion	3	Informatică economică	AI3501
EL13457	Abălaşei R Zicu	3	Informatică economică	AI3502
EL13458	Şpagă M Michael	3	Informatică economică	AI3503

	DenumireDisc	NrCredite	DataExamen	Nota
	Baze de date I	6	29/01/2004	8
	Baze de date I	6	29/01/2004	4
	Baze de date I	6	29/01/2004	9
	Programare vizuală şi RAD	7	01/02/2004	10
	Programare vizuală şi RAD	7	01/02/2004	8
	Programare vizuală şi RAD	7	01/02/2004	4
	Baze de date I	6	12/02/2004	8
	Programare vizuală şi RAD	7	15/02/2004	9
	Analiza sistemelor informaționale	6	04/02/2004	7

Fig. 4. Studenți, cursuri şi examene

Prima variantă - vezi figura 5 - foloseşte un atribut de tip relaţie numite Examene_ATR ce conţine informaţii despre fiecare examen susţinut de un student. Relaţia STUDENŢI_EXAMENE_ATR1 va conţine doar câte un tuplu pentru fiecare student.

STUDENŢI_EXAMENE_ATR1

Matricol	NumePrenume	An	Specializare
EL13455	Popovici I Vasile	3	Informatică economică
EL13456	Zăineanu W Ion	3	Informatică economică
EL13457	Abălaşei R Zicu	3	Informatică economică
EL13458	Şpagă M Michael	3	Informatică economică

Examene_ATR

CodDisc	DenumireDisc	NrCredite	DataExamen	Nota
AI3501	Baze de date I	6	29/01/2004	8
AI3502	Programare vizuală şi RAD	7	01/02/2004	10
CodDisc	**DenumireDisc**	**NrCredite**	**DataExamen**	**Nota**
AI3501	Baze de date I	6	29/01/2004	4
AI3502	Programare vizuală şi RAD	7	01/02/2004	8
AI3501	Baze de date I	6	12/02/2004	8
CodDisc	**DenumireDisc**	**NrCredite**	**DataExamen**	**Nota**
AI3501	Baze de date I	6	29/01/2004	9
AI3502	Programare vizuală şi RAD	7	01/02/2004	4
AI3502	Programare vizuală şi RAD	7	15/02/2004	9
CodDisc	**DenumireDisc**	**NrCredite**	**DataExamen**	**Nota**
AI3503	Analiza sistemelor informaţionale	6	04/02/2004	7

Fig. 5. Prima variantă de folosire a ATR

În egală măsură se poate însă folosi şi un ATR care să conţină toţi examinaţii pentru o disciplină dată într-o sesiune - StudenţiNote_ATR. Relaţia s-ar prezenta după cum este sugerat în figura 6.

Care dintre cele două variante reflectă mai bine realitatea ? Prima variantă prezintă avantajul grupării tuturor examenelor unui student sub "umbrela" tuplului care se referă la studentul respectiv. Ar fi relativ simplu de calculat media studentului, de aflat la ce examene a picat o singură dată, sau de două (trei...) ori etc. Dar dacă ne interesează studenţii care au luat la *Programare orientată pe obiecte* aceeaşi notă ca şi Şpagă Michael, atunci interogarea s-ar complica destul. A doua variantă este, ca structură, mai aproape de realitate, deoarece notele sunt preluate de pe cataloage care se întocmesc la fiecare examen. Avantajul obţinerii lejere a catalogului de la examen este "compensat" de dificultatea calculării mediilor pentru un student sau formaţie de studiu, pentru comparaţii între situaţiile şcolare ale studenţilor etc.

Din acest punct de vedere, renunţarea la ATR şi lucrul cu atribute atomice se materializează într-o structură, să-i spunem, mai neutră, care, deşi nu atât de intuitivă precum cea ierarhică, se pretează mult mai bine la prelucrări dintre cele mai diverse.

STUDENŢI_EXAMENE_ATR2

CodDisc	DenumireDisc	NrCredite	DataExamen
AI3501	Baze de date I	6	29/01/2004
AI3501	Baze de date I	6	12/02/2004
AI3502	Programare vizuală şi RAD	7	01/02/2004
AI3502	Programare vizuală şi RAD	7	15/02/2004
AI3503	Analiza sistemelor informaţionale	6	04/02/2004

StudenţiNote_ATR

Matricol	NumePrenume	An	Specializare	Nota
EL13455	Popovici I Vasile	3	Informatică economică	8
EL13456	Zăineanu W Ion	3	Informatică economică	4
EL13457	Abălaşei R Zicu	3	Informatică economică	9
Matricol	**NumePrenume**	**An**	**Specializare**	**Nota**
EL13456	Zăineanu W Ion	3	Informatică economică	8
Matricol	**NumePrenume**	**An**	**Specializare**	**Nota**
EL13455	Popovici I Vasile	3	Informatică economică	10
EL13456	Zăineanu W Ion	3	Informatică economică	8
EL13457	Abălaşei R Zicu	3	Informatică economică	4
Matricol	**NumePrenume**	**An**	**Specializare**	**Nota**
EL13457	Abălaşei R Zicu	3	Informatică economică	9
Matricol	**NumePrenume**	**An**	**Specializare**	**Nota**
EL13458	Şpagă M Michael	3	Informatică economică	7

Fig. 6. A doua variantă de folosire a ATR

Concluzii

După discuţia din această lucrare, în care domeniile pot fi definite cât se poate de flexibil, în funcţie de nevoile aplicaţiei, ne putem pune întrebarea: ce rost mai are să discutăm despre aducerea relaţii în 1FN, operaţiune care, uneori, presupune creşterea numărului de tupluri de câteva ori, astfel încât în loc să micşorăm redundanţa, mai degrabă o mărim, cel puţin aparent (vezi şi Fotache00 ?

Două ar fi argumentele în sprijinul celor prezentate. Mai întâi, în procesul normalizării, acesta e doar primul pas. Redundanţa pe care o introducem rezolvă o problemă extrem de importantă, cea a pierderilor de informaţii, fiind sigur că celelalte forme normalizate vor elimina aproape tot ce este redundant într-o relaţie.

Al doilea argument ţine de instrumentele software de care dispunem. La acest moment suportul SGBD-urilor pentru domenii de tip set, ca să nu mai vorbim de atribute de tip relaţie, este mai degrabă unul modest. Chiar dacă produsele importante au facilităţi importante în definirea de obiecte, când se pune problema mecanismului de declarare a integrităţii obiectelor, şi mai ales a celui de interogare, lucrurile se acutizează. Chiar dacă teoretic opţiunile sunt generoase, atunci când se pune problema punerii în operă a unei aplicaţii de lucru cu baze de date nu putem eluda "meandrele concretului", cu atât mai puţin "sinergia faptelor".

Bibliografie

[Abiteboul & Bidoit 84] Abiteboul, S., Bidoit, N. - *Non First Normal Form Relations to Represent Hiearchically Organized Data*, Proc. of the 3rd ACM SIGACT-SIGMOD symposium on Principles of database systems, Waterlov, Ontario, Canada, 1984

[Atzeni s.a.99] Atzeni, P., Ceri, S., Paraboschi, S., Torlone, R. - *Database Systems. Concepts, Languages and Architectures*, McGraw-Hill, London, 1999

[Codd72] Codd, E.F. – *Further Normalization of the Database Relational Model*, DataBase Systems, Courant Computer Science Symposia Series, Vol.6, Englewood Cliffs, N.J.,Prentice-Hall, 1972

[Codd79] Codd, E.F. - *Extending the Database Relational Model to Capture More*

Meaning, ACM Transactions on Database Systems, vol. 4, no.4, Dec. 1979

[Codd90] Codd, E.F. - *The Relational Model for Database Management. Version 2*, Addison-Wesley, 1990

[Date03] Date, C.J. (2003). "What First Normal Form Really Means", http:/www.dbdebunk.com, (ultima accesare - martie 2004)

[Date04] Date, C.J. – *An Introduction to Database Systems*, 8th edition, Pearson Addison-Wesley, Boston, 2004

[Date & Darwen00] Date, C.J., Darwen, H. - *Foundation for Database Systems. The Third Manifesto*, 2nd ed., Addison-Wesley, Reading, Massachusetts, 2000

[Dreizen & Chang 89] Dreizen, H.M., Chang, S.K. – *Imprecise Schema: A Rationale for Relations with Embedded Subrelations*, ACM Transactions on Database Systems, Vol. 14, no.4, December 1989

[Elmasri & Navathe 00] Elmasri, R., Navathe, S.R. - *Fundamentals of Database Systems*, Addison-Wesley, Reading, Massachussetts, 2000

[Fotache00] Fotache, M. – *Despre prima formă normală*, PC Report nr.92, mai, 2000