

l Edition
rporation,
Batson -
; Borland
ianapolis,

Modele semantice de date

Prof.dr. Ion LUNGU, prof.dr. Ion Gh. ROȘCA, conf.dr. Constanța BODEA
Catedra de Informatică Economică, A.S.E. București

Modelele semantice ale datelor sunt utilizate în faza de modelare conceptuală a datelor, activitate extrem de importantă în procesul de realizare a bazelor de date. Articolul prezintă o serie de rezultate obținute de autori în cadrul activității de cercetare efectuată la contractul "Baze de date deductive multimedia pentru învățământ", încheiat cu CNCSSU în anul 1997.

Cuvinte cheie: baze de date, modelare semantică, entități, asocieri, proprietăți ale entităților, proprietăți ale asocierilor

1. Abordări în construirea modelelor semantice ale datelor

Modelele semantice ale datelor se obțin pe

baza abordărilor top-down sau bottom-up (fig. 1):

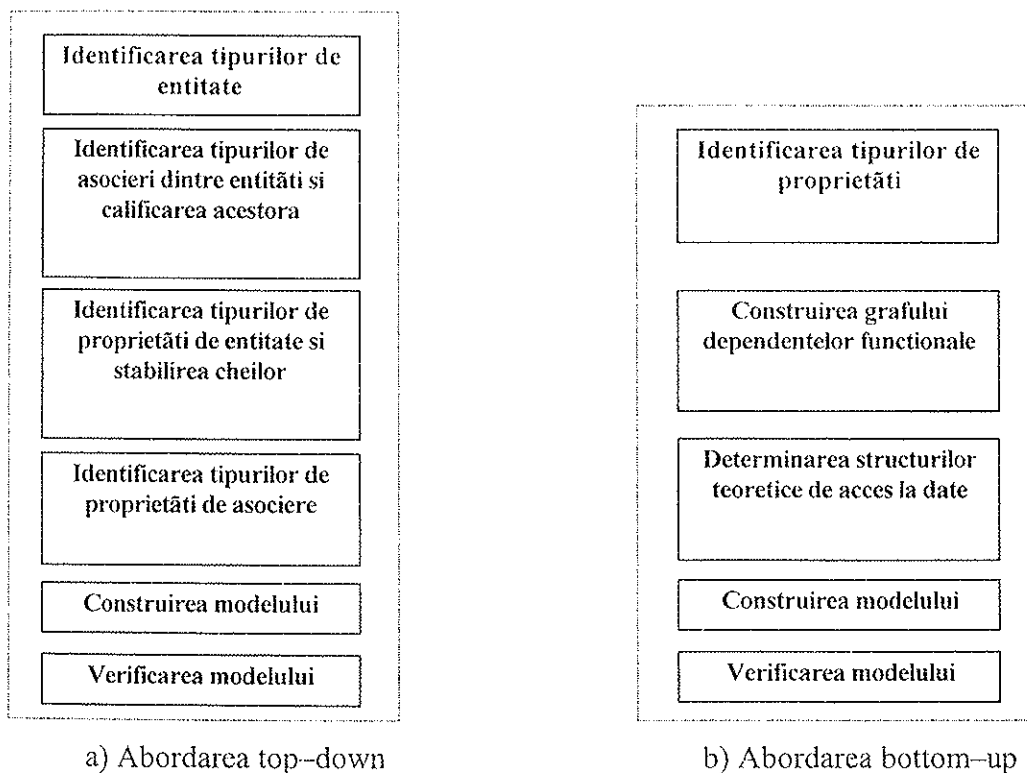


Fig. 1. Construirea modelelor semantice ale datelor

2. Abordarea top-down în construirea modelelor semantice ale datelor

Conform figurii 1, sunt parcurse următoarele etape:

- *Identificarea tipurilor de entitate.* Vor fi identificate acele componente ale sistemului real care sunt relevante, în sensul că utilizatorii sunt interesați atât de structura cât și de

comportarea lor. Entitățile identificate trebuie să aibă mai mult de o realizare și să prezinte cel puțin un atribut (proprietate) care să permită caracterizarea realizărilor.

- *Identificarea tipurilor de asocieri între entități și calificarea acestora.* O atenție deosebită se acordă identificării asocierilor de tipul generalizare/specializare și parte/întreg. Se consideră, pe rând, fiecare tip de

entitate și se analizează potențialul de specializare, în sensul posibilităților de realizare a unor specializări plecând de la tipul respectiv. O serie de asocieri între entități vor avea sens doar la nivelul unei specializări. De asemenea, specializarea se impune și atunci când nu toate atributele sunt valabile pentru toate realizările de entitate. Când numărul specializărilor este foarte mare, se reduc, menținându-se numai cele semnificative. După realizarea specializărilor se reactualizează lista entităților care urmează să fie incluse în modelul datelor. De asemenea, într-un mod asemănător este analizat potențialul generalizator al fiecărui tip de entitate, încercându-se identificarea generalizărilor posibile. Generalizarea se impune atunci când tipuri de entitate diferite prezintă atribute similare, sau asocieri analoge cu alte tipuri de entitate. Asocierile de tip întreg-parte se aplică în cazul entităților ansamblu / parte, container / conținut, colecție / exponat. Se va considera fiecare tip de entitate drept un întreg și se caută identificarea părților sale.

Totodată, fiecare dintre entități vor fi considerate drept potențiale părți ale unor entități ansamblu, și se încearcă identificarea acestora (dacă prezintă relevanță pentru utilizarea sistemului, adică dacă există cerințe informaționale referitoare la acestea). În urma identificării asocierilor de tip întreg-parte se

revizuieste lista entităților care urmează să fie incluse în model.

Sunt analizate și alte tipuri de asocieri ce se manifestă între entități și se realizează calificarea acestora, o atenție deosebită fiind acordată asocierilor "mulți la mulți".

Dificultățile în identificarea și calificarea asocierilor dintre entități

Asocierile puse în evidență în cadrul analizei structurate reflectă legături naturale care există între componentele sistemului (domeniului) real analizat. O aceeași realitate poate fi percepută diferit, de diferiți analiști de sistem, pentru un același sistem real putând fi obținute modele structurale distincte. Unele dintre modele pot încorpora aceeași semantică fiind, în același sens, echivalente. Altele, însă, pot reflecta numai parțial asocierile dintre entități sau pot exprima asocieri care nici nu există în realitate, fiind, prin urmare, modele eronate, obținute ca urmare a incapacității analistului de a percepe corect și complet realitatea sau de a o exprima corespunzător în cadrul unui model.

Dificultatea realizării unei corecte analize structurale provine, în primul rând din faptul că asocierile dintre entități posedă proprietatea de tranzitivitate, ceea ce face ca între entități să se manifeste nu numai legăturile directe, ci și indirecte (fig. 2).

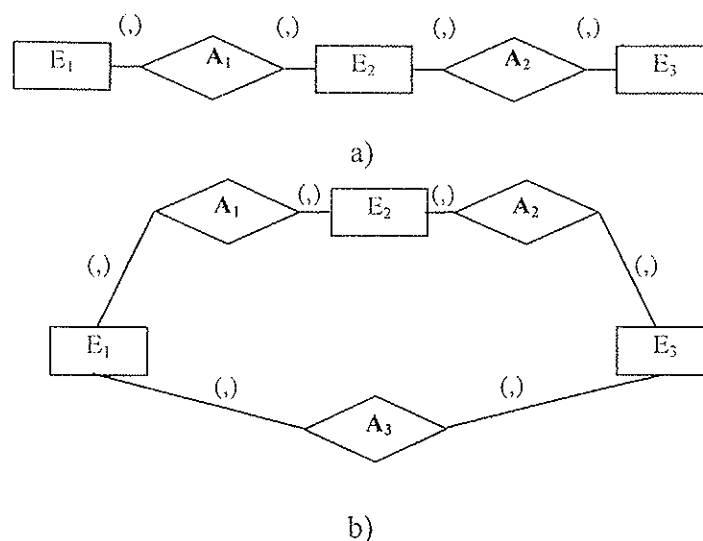


Fig. 2. Asocierile directe și indirecte între entități

Între entitățile E1 și E2, respectiv E2 și E3 din cadrul figurii 2.a există asocieri directe, A1 și A2. Între E1 și E3 există și o legătură indirectă. Uneori se preferă evidențierea legăturii între E1 și E3 și sub forma unei asocieri distincte A3 (fig.2.b.). A3 este redundantă atunci când pentru realizarea unei corecte legături în E1 și E3 este suficientă asocierea indirectă. Nu se reco-

mandă includerea de legături redundante în modelul structural, întrucât acestea complică modelul, făcându-l greu de utilizat. Există situații în care A3 nu este redundantă. Ele apar atunci când asocierile directe A1 și A2 nu sunt plasate într-o succesiune sau nu sunt definite într-un mod care să permită funcționarea corespunzătoare a legăturii indirecte.

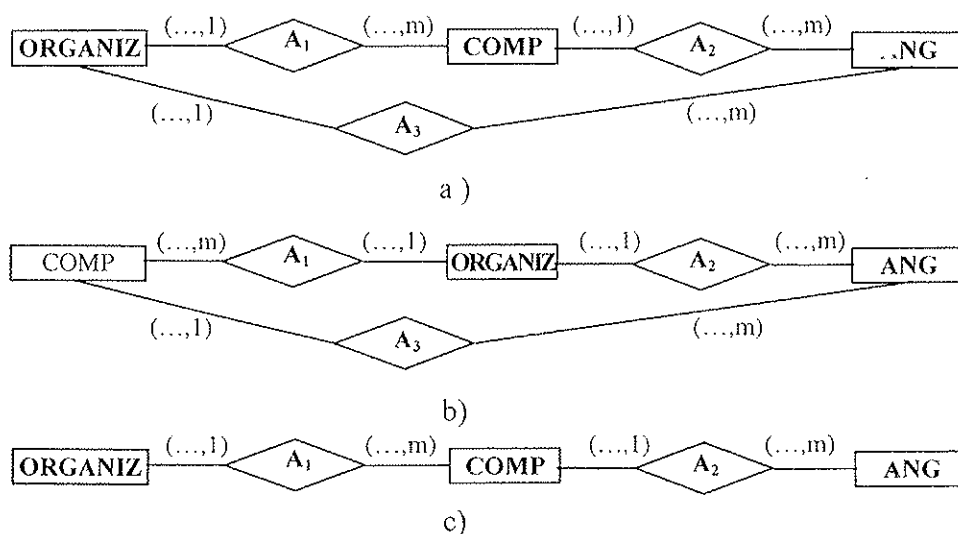


Fig. 3. Asocieri directe și indirecte între entitățile ORGANIZ, COMP și ANG

De exemplu, în figura 3.a, asocierea A3 este redundantă întrucât legătura între entitățile ORGANIZ (organizații economice) și ANG (angajați) poate fi asigurată prin intermediul asocierilor A1 și A2. Aceste asocieri permit determinarea angajaților unei anumite organizații, precum și organizația în care funcționează un anumit angajat.

În figura 3.b. asocierea A3 nu este redundantă, deoarece pe baza legăturilor A1 și A2 nu se pot determina angajații care lucrează într-un anumit compartiment sau compartimentul de care aparține un angajat. Întrucât asocierea indirectă nu funcționează corespunzător este necesară încorporarea unei legături directe între entitățile COMP și ANG. Figura 3.c. prezintă cea mai bună pentru reflectare a legăturilor dintre entitățile ORGANIZ, COMP și ANG în cadrul modelului structural.

Exemplul evidențiază importanța unei corecte înlănțuirii a legăturilor directe dintre entități, care să ducă la obținerea unor legături indirecte posibil de utilizat în mod eficient și care să permită simplificarea modelului structural, prin reducerea numărului de legături evidențiate explicit.

O altă dificultate întâmpinată în identificarea și calificarea asocierilor dintre entități constă în posibilitatea caracterizării aceleași legături în multe moduri.

De exemplu, legătura dintre entitățile ANGAJAT și COPII (fig. 4) poate fi calificat drept o asociere parțială de tip "mulți la mulți" (fig.4.a), ca asociere parțială de tip "doi la mulți" (fig.4.b.) sau drept o legătură exprimată prin două asocieri parțiale de tip "unu la mulți" (fig. 4.c) și care prezintă altceva decât descompunerea asocierii de tip "mulți la mulți" în asociere "unu la mulți"

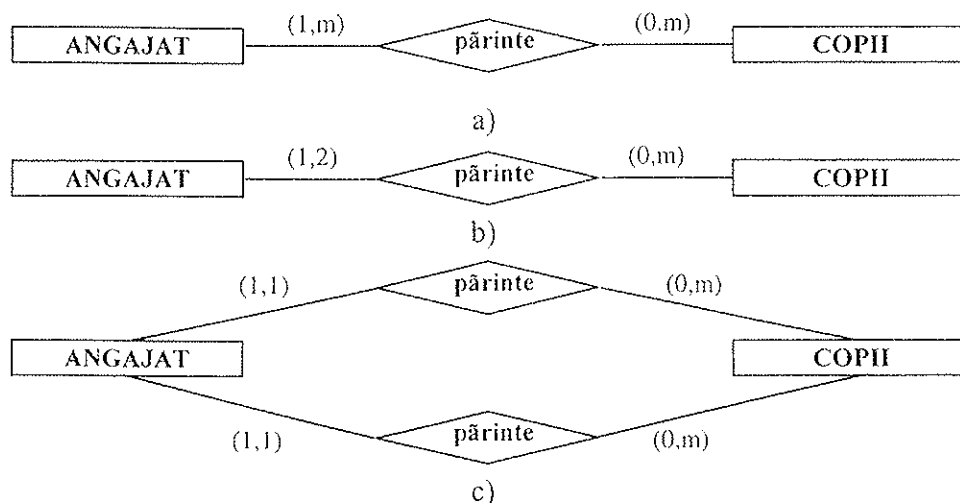
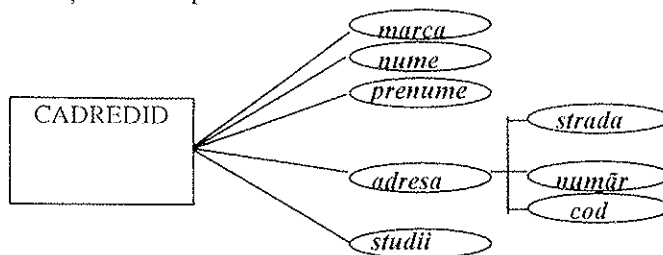


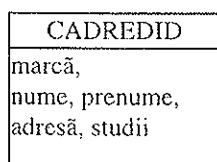
Fig. 4. Asocierea dintre entitățile ANGAJAT și COPII

• *Identificarea tipurilor de proprietăți de entitate și stabilirea cheilor.* Proprietățile (atributele) reprezintă caracteristici ale entităților pentru identificarea acestora. Pentru fiecare tip de entitate se caută stabilirea: modului în care se poate realiza descrierea tipului de entitate, informațiile care prezintă

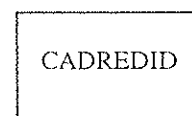
interes pentru tip, stările în care se poate afla tipul de entitate și modul în care este posibilă identificarea fiecărei stări în parte. Figura 5 prezintă entitatea CADREDID (CADRE DIDactice), cu menționarea proprietăților.



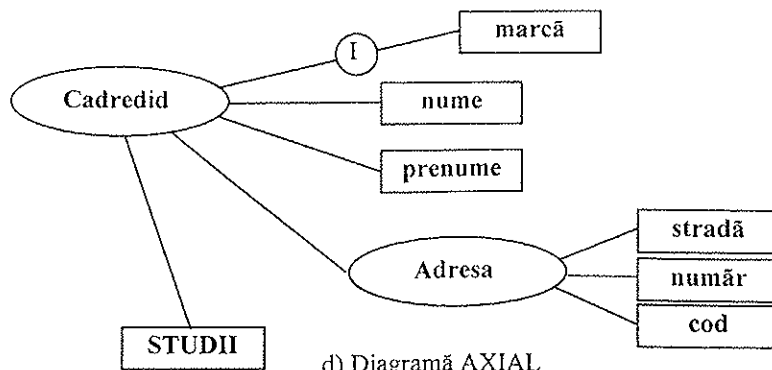
a) Diagramă entitate-asociere (Chen)



b) Diagramă MERISE



c) Diagramă SSADM



d) Diagramă AXIAL

Fig. 5. Menționarea proprietăților pentru entitatea CADREDID

Un atribut poate fi de identificare (cheie) dacă satisface o serie de cerințe. Pentru aceeași entitate pot exista mai multe atribute de identificare (mai multe chei candidate). Selectarea uneia dintre candidații cheie drept atribut de identificare a entității se realizează după anumite reguli precizate de literatura de specialitate.

• *Identificarea tipurilor de proprietăți ale asocierilor.* Se realizează într-o manieră similară etapei de identificare a tipurilor de asocieri între entități și calificare a acestora.

• *Construirea modelului.* Presupune completarea diagramelor cu tipurile de proprietăți de entitate și de asocieri. Diagramele pot deveni relativ complexe, ceea ce determină ca unele metode de modelare (de exemplu SSADM, YOURDON) să nu includă în reprezentări și proprietățile, ci numai entitățile și asocierile dintre acestea. Figura 6 prezintă modelul conceptual al datelor în formă completă.

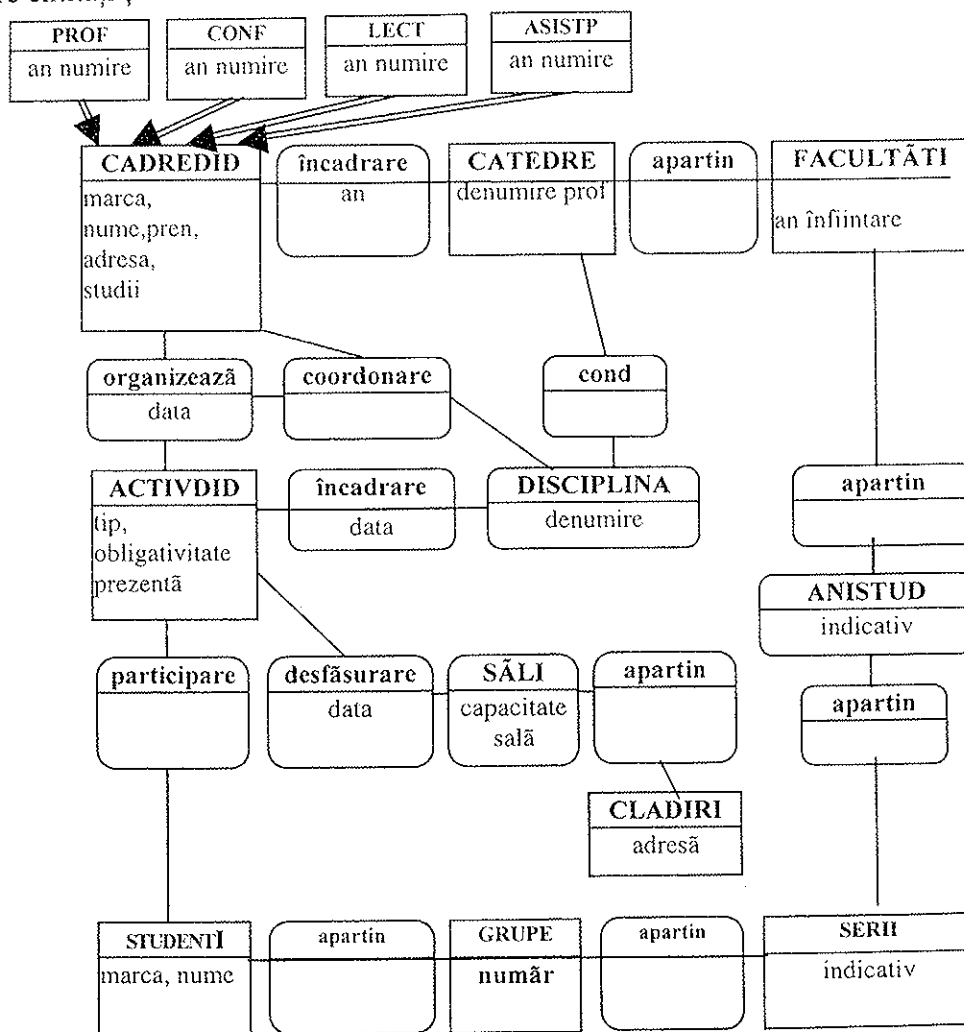


Fig. 6. Modelul semantic al datelor – forma completă – diagramă MERISE

Atunci când în cadrul sistemului există o bază de date, modelul semantic al datelor se poate obține plecând de la schema conceptuală a bazei de date, prin translatarea acesteia din model logic global în model conceptual al datelor. Trecerea prezintă diverse forme, după cum dorim sau nu ca modelul semantic să includă asocieri "mulți la mulți".

3. Abordarea bottom-up în construirea modelului semantic al datelor

A fost definită în cadrul MERISE sub forma următoarei succesiuni de etape: identificarea tipurilor de proprietăți; construirea grafului dependențelor funcționale; determinarea

ate afla
re este
n parte.
REDID
proprie-

structurii teoretice de acces la date; construirea modelului.

Arcele terminale obținute plecând de la proprietățile elementare definesc entitățile. Originile acestor arce constituie identificatorii de entitate. Arcele care rămân pun în evidență asocierile. Proprietățile nerezolvate ce rămân

sunt atribuite asocierilor. Proprietățile izolate vor defini entitățile izolate. Pe baza regulilor de gestiune sunt utilizate pentru stabilirea coordonatelor.

Figura 7 prezintă modul de constituire a unui model semantic, într-o abordare bottom-up.

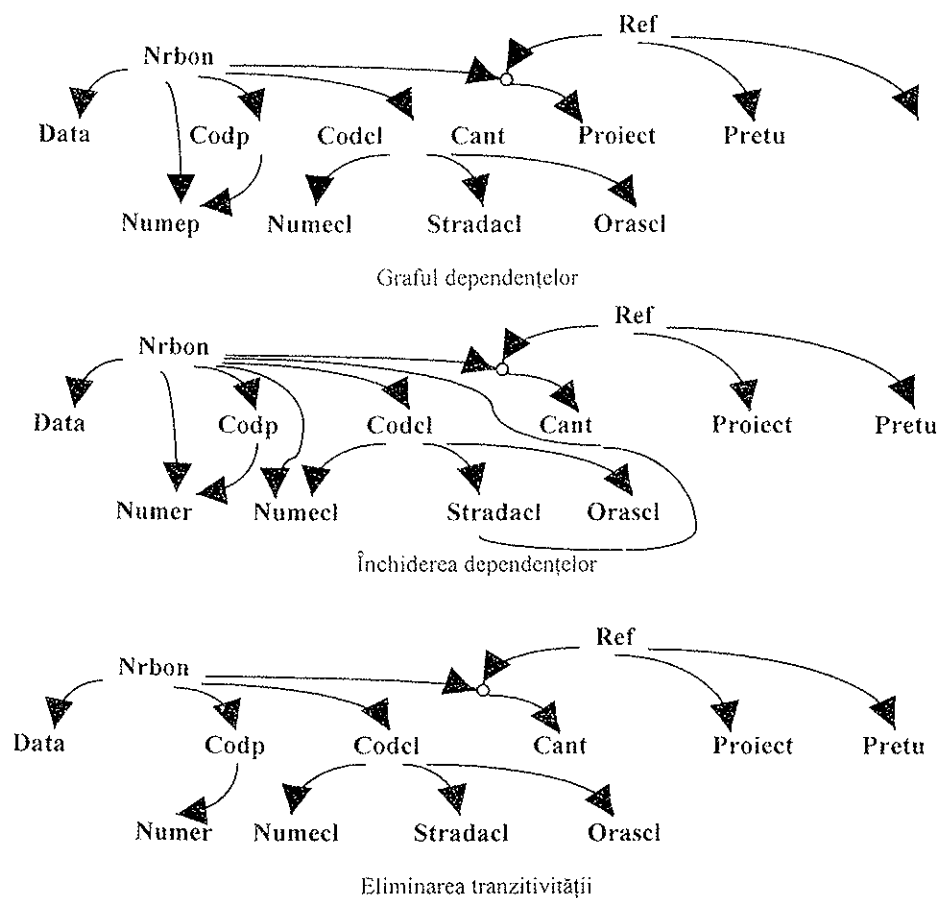


Fig. 7. Construirea modelului semantic al datelor - abordarea bottom-up

Dictionar de proprietăți
 Nrbon, Data, Numer, Codp
 Codel, Numecl, Stradacl, Orasel
 Cant, Ref
 Proiect, Pretu

Dependente funcționale
 $Nrbon \rightarrow Codp$, $Codp \rightarrow Numer$
 $Nrbon \rightarrow Codel$, $Codel \rightarrow Numecl$
 $Codel \rightarrow Stradacl$, $Codel \rightarrow Orasel$
 $Ref \rightarrow Proiect$, $Ref \rightarrow Pretu$

Bibliografie

• Lungu I., Sabău Gh, Surcel Tr., Bodea C.- Sisteme informatice pentru conducere, Editura SIAJ, 1995.

• Roșca Gh. I. și colectiv - Raport de Cercetare la Contractul nr. 7029/1997 cu CNCSU