

Utilizarea retelelor bayesiane în fiabilitatea software

Lect. Mihai POPESCU
Academia Tehnica Militara, Bucuresti

Retelele bayesiane sunt cea mai buna solutie pentru problemele care au un suport decizional incert. Cea mai recenta folosire a fost de catre Microsoft în wizardurile de help din Microsoft Office, dar si în aplicatii care necesita luarea deciziei într-un timp critic pentru sistemele de propulsie în spatiul aerian [HORV95].

Cuvinte cheie: fiabilitate, complexitate, retele bayesiene.

Retelele bayesiane sunt cunoscute si sub numele de retele probabilistice cauzale, retele de probabilitate grafica, modele probabilistice cauza-efect si diagrame de influenta probabilistica, fiind de actualitate datorita folosirii lor ca o posibila solutie pentru situatiile în care informatia avuta la dispozitie este incompleta.

Deoarece pentru estimarea si predictia fiabilitatii software este nevoie de modele de fiabilitate care sa poata manevra: procese si produse diverse; relatii cauza si efect adevarate; incertitudine; informatie incompleta, utilizarea retelelor bayesiane este adecvata, în special pentru faptul ca exista instrumentul software HUGIN, pe care autorul l-a achizitionat via Internet, care implementeaza calculul probabilistic sofisticat al probabilitatilor conditionate, fiind sprijinit si de o interfata grafica de top. Un BBN (*Bayesian Belief Networks*) [XXXX99] este un graf care reprezinta relatii probabilistice între variabile.

Nodurile grafului reprezinta variabile, iar arcele relatii cauzale între variabile.

Cu BBN este posibil sa se experimenteze dependente între variabile diferite si sa se propage impactul unui lucru evident în probabilitati ale unor rezultate nesigure, cum ar fi fiabilitatea viitoare a unui sistem. Cu BBN se pot reprezenta si manipula modele complexe ale fiabilitatii software, care pot sa nu fie implementate niciodata cu metodele conventionale.

Alt avantaj este acela ca poate prezice evenimente bazat pe date parțiale sau nesigure.

Se propune un exemplu despre cum se poate utiliza BBN pentru a prezice fiabilitatea unui produs bazat pe cunoasterea procesului si a produsului însusi (figura 1). Nodurile pot fi variabile discrete sau continue. De exemplu, nodul utilizând un standard de codificare este discret, având doua valori: "da" si "nu", în timp ce nodul fiabilitate poate fi continuu (cum este probabilitatea de defectare).

Arcele sunt relatii cauzale (de influenta) între variabile. De pilda, fiabilitatea software se defineste prin numarul de defecte latente si prin utilizarea operationala (frecventa cu care defectarile se pot declansa).

Aceasta relatie se modeleaza prin trasarea arcelor din nodurile "Numarul de defecte latente" si "Folosire operationala" înspre "Fiabilitate".

Daca am sti ca exista un numar mare de defecte si un nivel ridicat de folosire operationala, atunci probabilitatea ca fiabilitatea sa fie mare ar fi mai mica decât atunci când am sti ca sunt defecte putine si o utilizare operationala scazuta.

În BBN acest lucru se modeleaza cu tabela de probabilitate a unui nod, NPT (*Node Probability Table*).

Nodul de fiabilitate poate fi o variabila discreta, luând trei valori: joasa, medie si înalta.

NPT capteaza probabilitatile conditionale ale nodurilor parinte. Pentru nodurile fara parinti (ex.: folosirea standardului IEC 1508) tabela probabilitatii nodului e o simpla probabilitate marginala.

Probabilitatile pentru NPT pot fi de tipul (tabelul 1):

➤ probabilitati subiective (stabilite de experti);

➤ probabilitati bazate pe date obiective (colectate pe proiecte similare).

Tabelul 1. Tabela NPT pentru nodul "fiabilitate"

Folosire operationala		Scazuta			Medie			Intensa		
Numar de defecte		mic	mediu	ridicatat	Mic	mediu	ridicatat	mic	mediu	ridicatat
Fiabilitate	Mica	0,1	0,2	0,33	0,2	0,33	0,5	0,2	0,33	0,7
	Medie	0,2	0,3	0,33	0,3	0,33	0,3	0,3	0,33	0,2
	Înalta	0,7	0,5	0,33	0,5	0,33	0,2	0,5	0,33	0,1

Pentru relevanta folosirii instrumentului Hugin, se porneste de la urmatorul scenariu:

- nu se stie nimic despre proces;
- experienta echipa de conducere;

- complexitatea problemei;
- utilizarea standardului IEC 1508.
- nu se stie nimic despre folosirea operationala.

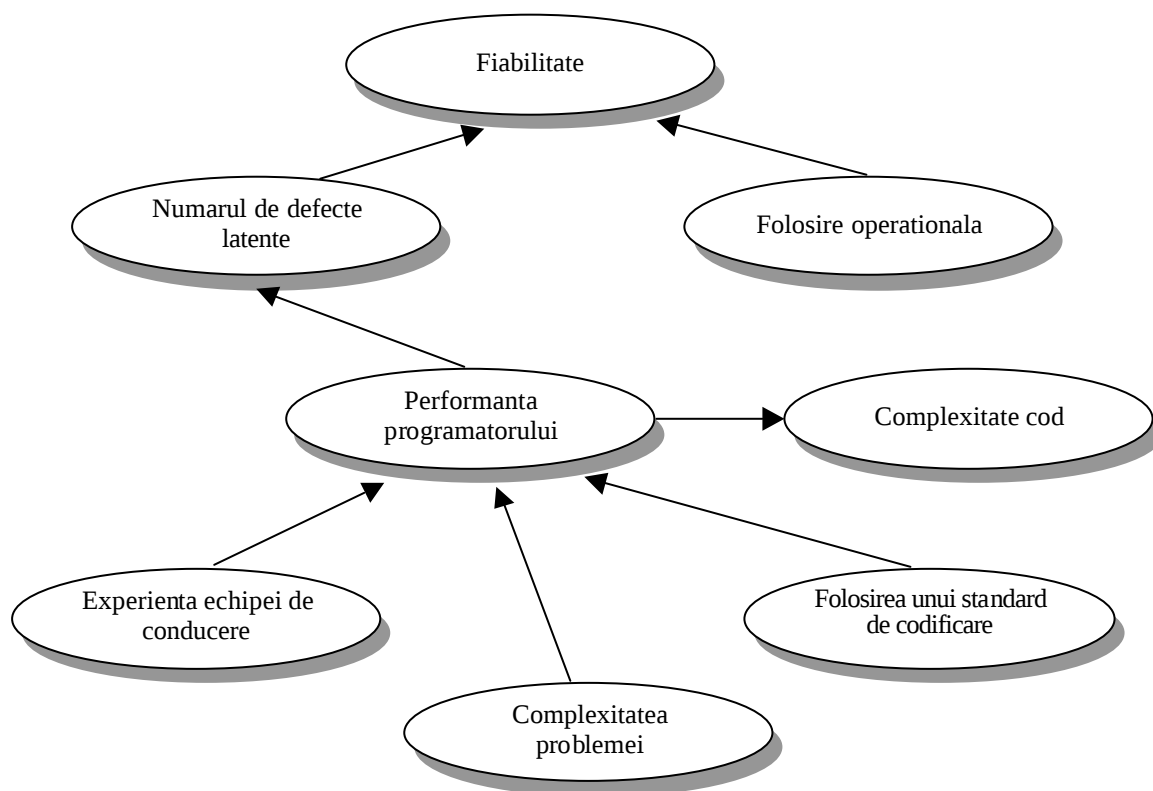


Fig.1. Utilizarea BBN pentru predictia fiabilitatii folosind claritatea procesului si a produsului

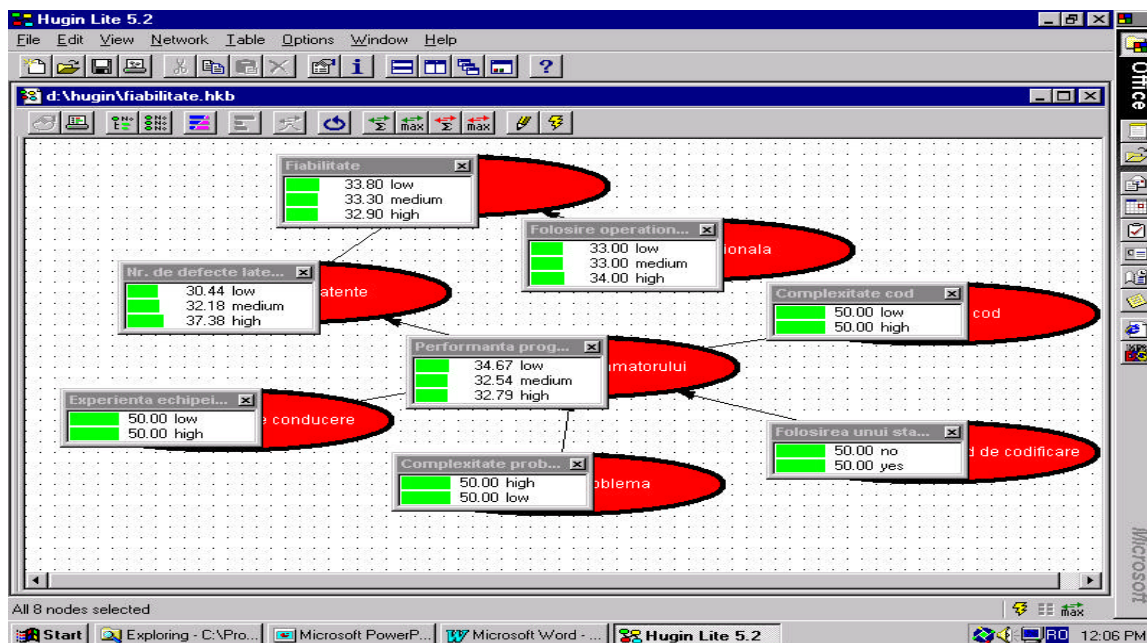


Fig. 2. BBN - starea initiala

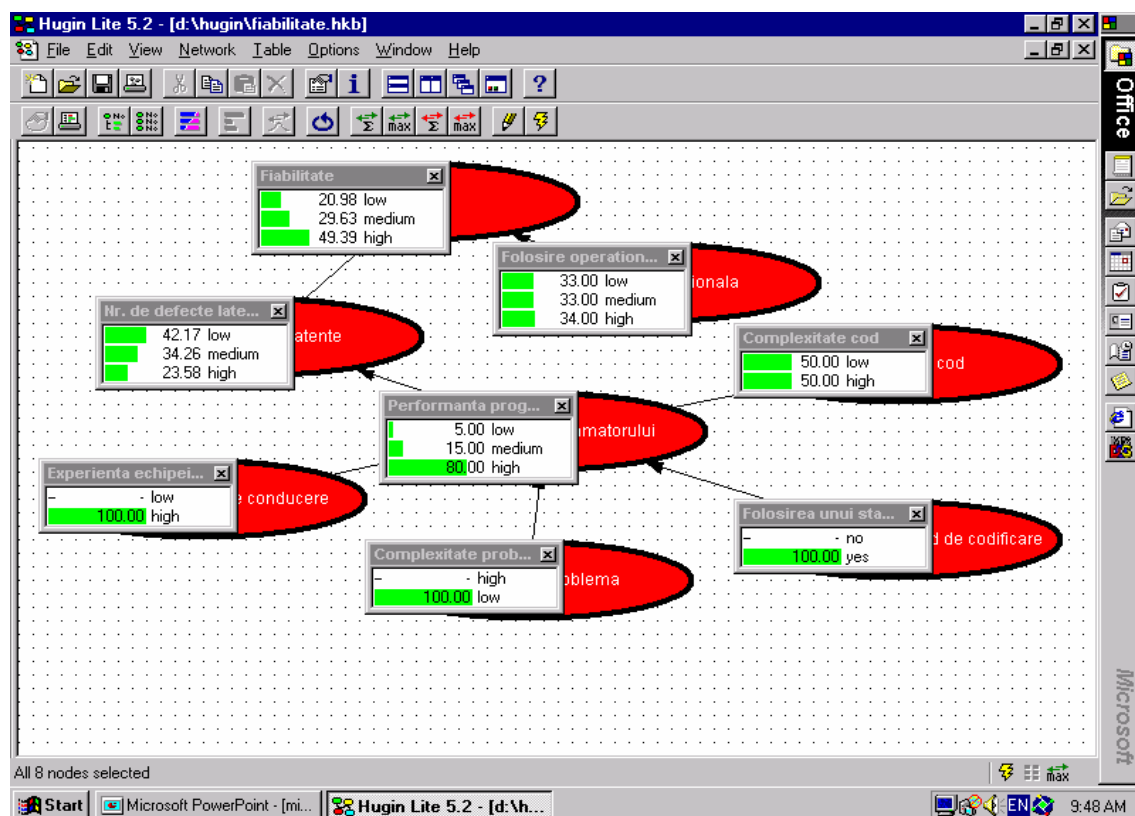


Fig. 3. BBN - proces controlat

Aceasta ipoteza se transpune în egalitatea probabilitatilor matematice pentru toate aceste noduri. Exista o tendinta de ignoranta care va duce la a se crede ca performanta programatorului e mult mai

probabil scazuta decât înalta si de aceea numarul de defecte latente e mult mai probabil sa fie mai mare decât mic, ceea ce are ca urmare o fiabilitate scazuta, chiar

daca nu stim nimic despre folosirea operationala a produselor (figura 2).

În continuare (figura 3), se face supozitia ca procesul e foarte bine controlat. Aceasta se materializeaza prin atribuirea valorii 1 probabilitatilor respective. Acest lucru va conduce la modificari semnificative asupra performantei codicatorului.

Impactul total asupra fiabilitatii nu e extraordinar (se schimba destul de modest fiabilitatea înalta).

Beneficii BBN:

— beneficiul major fata de modelele statistice clasice (care lucreaza cu nivele de încredere, nu cu probabilitati) este ca, dupa observarea rezultatului unui experiment (colectând date noi), actualizeaza convingerea avuta înaintea experimentului (probabilitatea anterioara), generând o probabilitate posterioara; nu este, însa, usor de

obtinut probabilitatea anterioara, chiar de experti în domeniu;

— beneficiaza de un algoritm inteligent de propagare care lucreaza eficient chiar pentru retele BBN mari (sute de noduri). Crearea unui instrument, numit Hugin, care ofera o interfata grafica performanta pentru motorul de propagare, a facut fezabila modelarea problemelor cu deductie nesigura.

Bibliografie

[HORV95] Horvitz, E., Barry, M., "Display of information for time-critical decision making", Proc. of 11th Conf in AI, Montreal, August 1995.

[XXXX99] Documentatie de produs, via INTERNET