Ratio Mathematica 9
(1995), 87 - 99

WHAT IS INSIDE A PROBABILISTIC EXPERT SYSTEM ?*

Giuseppe Di Biase**

SUNTO - Si illustra il funzionamento di un sistema esperto probabilistico senza soffermarsi sugli aspetti teorici che sono stati studiati in molti lavori, tra i quali vedi quelli citati in bibliografia. Ci si muove nell'ambito della teoria della probabilità soggettiva di de Finetti. Una volta verificata la coerenza di un assegnamento di probabilità su pochi eventi condizionati presi inizialmente in considerazione in merito ad un problema che si vuole studiare, il teorema fondamentale della probabilità soggettiva ci assicura l'esistenza di un prolungamento coerente della probabilità a nuovi eventi che rappresentano nuove informazioni che si acquisiscono durante l'analisi del problema. Il procedimento viene illustrato attraverso cinque esempi che possono rappresentare alcune situazioni tipiche in cui ci si può imbattere durante l'analisi di casi concreti collocati nell'ambito dell'incertezza.

ABSTRACT - This paper is an outline of how a probabilistic expert system works, therefore the theoretical details have been omitted. They have been widely dealt with e. g. in the quoted references. The framework is the one of de Finetti's theory. His fundamental theorem allows us to consider at first only the probability distribution on some of the events of interest and then, when new information gets into the system, to extend to other conditional events in a coherent way. Five examples that consider several typical situations one can come up against are illustrated. They allows us to have a general understanding of how a probabilistic expert system works.

---

* Work partly supported by C.N.R.- research grant "Matematica per la tecnologia e la società" (Resp. Prof. G. Coletti), contributed paper at Workshop "Modelli decisionali per sistemi intelligenti", Perugia 16-18 Novembre 1995.

** Dipartimento di Scienze - Viale Pindaro 42 - 65127 Pescara.

INTRODUCTION

A probabilistic expert system is a tool of Artificial Intelligence that should help, but not substitute, a decision maker with reference to an analyzing problem in various fields of science.

We assume that the judgements of the experts in their own fields of interest can be synthesized by means of conditional events. The framework is the subjective probability theory, introduced by de Finetti. His fundamental theorem allows us to consider at first only the probability distribution on some of the events of interest and then, when new information gets into the system, to extend to other conditional events in a coherent way. This is in opposition to the classical probability theory which considers the complete space of all possible cases.

THE FRAMEWORK

The knowledge base of the expert system is formed by a finite family of conditional events $\{E_i / H_i , i \in I\}$, $I = \{1, 2, ..., n\}$ plus the logical constraints (equivalences, implications, compatibility relations, etc.) among the family of standard events $\{E_i , H_i , i \in I\}$. Moreover a probability distribution on the family must be assessed. We denote by $p = (p_1, p_2, ..., p_n)$ such an assessment. As far as the notation is concerned we use the symbol without subscript to indicate a vector and the same symbol with subscript for the components of the vector.

We recall that a conditional event E / H (E given H, that is the event E supposing that the event H is true) is regarded as a three valued event. Denoting by EH the logical product between E and H and by -E the logical negation of the event E, it is true if EH occurs, false if (-E)H occurs and indefinite if H does not occur. If we consider the indicator function of E given H, it assumes the values 1, 0 or p(E / H) respectively in the three cases mentioned above.

It is helpful to introduce the concept of atoms associated with E / H. They are the elements of the set {EH, -(E)H, -H}. This is a partition of the certain event Ω. So the atoms associated with the family assessed can be represented by the logical products in all possible way. In others words each atom can be represented by $A_1 A_2 ... A_n$, different from the impossible event $\varnothing$, where, for each i belonging to I, A_i belongs to $\{E_i H_i, (-E_i)H_i, -H_i\}$.

The cardinality of the family of atoms is at most equal to 3^n. So it increases in an exponential way when the conditional events of interest rises and then a computational difficulty appears.

We have implemented an expert system on the Mathematica software that reduces this expansion because the knowledge base is updated step by step by considering each time the logical constraints among the standard events. Obviously the efficency is related to the size of the family of events considered, or even more to the logical constraints among them.

Since we handle conditional events in different probability spaces their consistency is not trivial. The coherence of the probability assessment must be checked. For the theoretical aspects see i. e. COLETTI [1], CRISMA [3], DE FINETTI [4], DI BIASE-MATURO [7], GILIO [9], HOLZER [14], REGAZZINI [16], SCOZZAFAVA [17], VIGIG [21].

RIGHT TO THE HEART OF THE PROCEDURE

Let $\{C_i\ , i \in J\}$, $J = \{1, 2, ..., s\}$, $s \leq 3^n$ the family of the atoms associated with our conditional events. Assessed p we need to see if there exist s real numbers q_i's such that

$$\begin{cases} \sum\limits_{C_k \subseteq E_i H_i} q_k - p_i \sum\limits_{C_k \subseteq H_i} q_k = 0\ , \ i \in I \\ \sum\limits_{k \in J-\{s\}} q_k + q_s = 1 \\ q_k \geq 0 \quad \forall\, k \in J \end{cases} \tag{1}$$

The system (1) is a constrained linear system of n+1 equations in s unknowns and s constraints. If the system (1) is compatible then the assessment p is coherent and the unknowns can be regarded as the probabilities of the atoms. Let S be the set of solutions q of system (1).

Notice that the unknown associated with the atom C_s appears only once in system (1). This variable is useful to run our algorithm, see DI BIASE [6] and DI BIASE-MATURO [8], and so it will not be dropped, although the probabilistic consistency imposes $q_s = 0$ in both cases $C_s = \varnothing$ and $C_s \neq \varnothing$.

Notice that the first n equations of (1) apply the multiplication theorem of probability which represents a necessary and sufficient condition for the coherence. But some equations could be satisfied in a trivial way, that is it could happen $p(H_i) = 0$ for some $i \in I$. In other words the set A defined below could be different from the empty one:

$$A = \left\{ i \in I : \exists\, q_k \in q : \sum_{C_k \subseteq H_i} q_k = 0 \right\} \neq \varnothing .$$

In this case the multiplication theorem for $i \in A$ still holds, but not in a signifying way because all assessments p are coherents. Nevertheless checking the coherence of $\{p_i, i \in A\}$ is enough, as proved by COLETTI [2] and GILIO [10]. Whereas their algorithms repeat the procedure by considering each time the subsystems formed by equations corresponding on the $i \in A$ and by finding each time the family of the atoms associated with $\{E_i|H_i, i \in A\}$, here many computational problems are solved; for details see DI BIASE [5] and DI BIASE-MATURO [7].

In order to do this we consider the null space associated with each solution q belonging to S:

$$K(q) = \{k \in J - \{s\} : q_k = 0\}$$

Obvioulsly if $K(q) = \varnothing$ then the components are all positive real numbers and the procedure stops, else we search for other solutions of system (1) by means of the following results proved in DI BIASE [6] and DI BIASE-MATURO [8].

Result 1. The trivial solution $q = (0, 0, ..., 1)$ always exists.

Result 2. Each convex combination of two vectors q belonging to S is a solution too.

Result 3. There exists a solution of system (1) with $q_s < 1$ if and only if another solution in which $q_s = 0$ exists.

From a numerical wiewpoint solving the following optimization problems is enough:

$$\begin{cases} \min\limits_{q \in S} q_s \\ \text{subject to system (1)} \end{cases} \tag{2}$$

By applying result 3, if such minimum is less than one we put it equal to zero. Let q_0 be that minimum.

In particular in DI BIASE [6] we coded a module which allows to determine two types of solutions: a q_1-type by solving a problem (2) and a q_2-type through a convex combination of solutions previously found. Let q_1 and q_2 be such solutions.

Since $K(q_2) = K(q_0) \cap K(q_1)$ it is $K(q_2) \subseteq K(q_1)$. It follows

$$\operatorname{card} K(q_2) \leq \operatorname{card} K(q_1) \tag{3}$$

Now if in (3) the stricly inequality holds the procedure will stop after a finite number of steps because we will find a $K(q_2)=\emptyset$. Else if equality holds, being $A\neq\emptyset$, we have

$$\{E_i|H_i, i\in A\}\subset\{E_i|H_i, i\in I\}$$

and then the algorithm will stop after a finite number of steps too.

In such a case we search for the solutions by updating the auxiliar atom putting

$$C_s = C_s \cup \left(\underset{i\in K(q)}{Y} C_i \right)$$

and minimizing q_s subject to system (1) in which we swap $i\in I$ with $i\in A$.

SOME EXAMPLES

For a theoretical solution of the following examples see GILIO [12] and SCOZZAFAVA [18]. Here we show their automatic solution implemented on the Mathematica software. For details of this implementation see DI BIASE [5] and [6]. The symbols && and ! indicate respectively the logical product and the logical negation, see i.e. MAEDER [15]. For explanation we write only some items of the implementation. Finally we would like to emphasize the use of Mathematica system especially for its ability to compute a large number of logic operations and problems of linear programming that are tedious to code and difficult to code correctly. The typical interactivity of this program, while handling expert systems, also is appreciated. Last but not least we find out that this software is very fit for checking the coherence of probability assessments, in fact its way of thinking is similar to our own, especially as far as the probabilistic and logical problems are concerned. Moreover it can simplify these problems identifying the events and their numerical values (the probabilities) with the same symbol.

Example 1. There are four conditional events e_1 / h_1, e_2 / h_2, e_3 / h_3, e_4 / h_4, such that $e_1 = e_2 = e_3 = e_4 = e$, $h_3 = h_1$ && h_2. Given p= (0.8, 0.01, 0.04, 0.1) it happens every time card $K(q_2) < K(q_1)$, then the module "solutions" run many time until it finds a solution with all positive components, see **Out[18]**. Notice that the items related to the building of the family of the atoms associated with the given events are left out. The atoms are collected in the vector d which appears beginning from In[12]. The same happens in all other reported examples.

In[1]:=
```
F={{e,h1},{e,h2},{e,h3},{e,h4}};
prob={0.8,0.01,0.04,0.1};
```

In[2]:=
```
h3=p[h1,h2];
```

In[12]:=
```
p=ConstrainedMin[ob1,equations,d];
sol1=d/.p[[2]]
```

Out[12] =
```
{0, 0, 0.0222222, 0.777778, 0, 0, 0.2, 0, 0, 0,
 0, 0, 0, 0, 0}
```

In[15]:=
```
solutions[x_,y_]:=Module[{ob,mx,sol,verasol},
ob=Apply[Plus,x];
mx=ConstrainedMax[ob,equations,d];
sol1=d/.mx[[2]]; verasol=(y+sol)/2]
```

In[16]:=
```
auto=solutions[nul[d,sol1],sol1]
```

Out[16]=
```
{0, 0, 0.0111111, 0.388889, 0, 0, 0.1,
 0, 0.005, 0, 0.045, 0.45, 0, 0, 0}
```

In[17]:=
```
While[Count[auto,0] > 0,
auto=solutions[nul[d,auto],auto]]; Print[auto]
```

Out[18]=
```
{0.000641026, 0.00128205, 0.0172676, 0.221675,
 0.0153846, 0.0307692, 0.0015625, 0.0125,
 0.000078125, 0.005, 0.144934, 0.502031,
 0.0015625, 0.0140625, 0.03125}
```

Example 2. There are three conditional events e_1 / h_1, e_2 / h_2, e_3 / h_3, such that $e_1 = e_3 = A$, $h_1 = e_2 = B$, $h_2 = h_3 = C$, where $A \Rightarrow B \Rightarrow C$. Given p= (0.2, 0.3, 0.1) it happens the probability of auxiliar atom is equal to one, see Out[15], then the assessment is not coherent.

In[1]:=
```
family={{A,B},{B,C},{A,C}};
prob={0.2,0.3,0.1};
```

In[6]:=
```
d={A, B && !A, C && !B, !B && !C}
```

In[12]:=
```
equations[ Append[equations, normalize]
```

Out[12]=
```
{0.2 (A + (B && !A)) == A,
 0.3 (A + (B && !A) + (C && !B))
 == A + (B && !A),
 0.1 (A + (B && !A) + (C && !B) == A,
 A+(B && !A)+(C && !B)+(!B && !C) == 1}
```

In[13]:=
```
obl = !B && !C
```

In[14]:=
```
p = ConstrainedMin[obl,equations,d];
```

In[15]:=
```
sol1=d/.p[[2]]
```

Out[15]=
```
{0, 0, 0, 1}
```

In[16]:=
```
If[Last[sol1]==0,Print["MUST FIND OTHER SOL."],
Print["THE ASSESSMENT IS NOT COHERENT"]]
```

Out[16]=
```
THE ASSESSMENT IS NOT COHERENT
```

Example 3. Same family of conditional events of previous example, but different probability assessment: p = (0.2, 0.3, 0.06). In this case the last component of vector q_0 in Out[29] is less than one, then assessment p could be coherent. Because we find all essential atoms with positive probabilities the procedure could be stopped, but for explanation we search for the null space associated. Since it is empty, see Out[34], there is coherence.

On the other hand it is easy to prove that the assessment is coherent if and only if $p_3 = p_1\, p_2$.

In[21]:=
```
family={{e1,h1},{h1,h2},{e1,h2}};
prob={0.2,0.3,0.06};
```

In[28]:=
```
p = ConstrainedMin[ob1,equations,d];
```

In[29]:=
```
sol1=d/.p[[2]]
```

Out[29]=
```
{0.06, 0.24, 0.7, 0}
```

In[30]:=
```
If[Last[sol1]==0,Print["MUST FIND OTHER SOL."],
Print["THE ASSESSMENT IS NOT COHERENT"]]
```

Out[30]=
```
MUST FIND OTHER SOL.
```

In[31]:=
```
sol1=Drop[sol1,-1];
```

In[32]:=
```
null[x_,y_]:=Part[x,Flatten[Position[y,0]]]
```

In[33]:=
```
facs=Table[d[[i]]->i, {Length[d]}];
ddfacs=Dispatch[facs];
```

In[34]:=
```
null[d,sol1] /. ddfacs
```

Out[34]=
```
{}
```

In[35]:=
```
If[Length[%]==0,Print["Coherent"],Print["Goes on"]]
```

Out[35]=
```
Coherent
```

Example 4. We update the family of example 3 with a fourth conditional event E / B such that E => !A && B where E is different from the impossible event. We have a collection of five atoms, see **In[10]**, usually denoted by d. Given p = (0.2, 0.3, 0.06, 0.5) we find a solution with their last component less than one, see **Out[12]**. Since the other components are all positive numbers the coherence is assured, see **Out[18]**.

In[1]:=
```
F={{A,B},{B,C},{A,C},{E,B}};
prob={0.2,0.3,0.06,0.5};
```

In[10]:=
```
d ={A, E, B && !A && !E, C && !B, !B && !C};
```

In[11]:=
```
p = ConstrainedMin[obl,equations,d];
```

In[12]:=
```
sol1=d/.p[[2]]
```

Out[12]=
```
{0.06, 0.15, 0.09, 0.7, 0}
```

In[15]:=
```
sol1=Drop[sol1,-1];
```

In[16]:=
```
null[x_,y_]:=Part[x,Flatten[Position[y,0]]]
```

In[17]:=
```
facs=Table[d[[i]]->i, {Length[d]}];
ddfacs=Dispatch[facs];
```

In[18]:=
```
null[d,sol1] /. ddfacs
```

Out[18]=
```
{}
```

In[19]:=
```
If[Length[%]==0,Print["COHERENT"],
Print["CONTINUE"]]
```

Out[19]=
```
COHERENT
```

Example 5. Same family of conditional events of example 4, but different probability assessment: p = (0.2, 0, 0, 0.5). In this case there exists a solution with the last component less than one, but many components are equal to zero, see Out[29]. We must find other solutions of q_1 -type and q_2 -type, see Out[33] and Out[34]. Since they have the same cardinality we go on and we find a control set A different from the empty one, see Out[35]. In this case we consider a control subfamily, denoted by symbol G, by imposing to the index i to belong to A, instead of to I. By analyzing the corresponding equations we can say that the assessment is coherent. On the other hand it is easy to prove that the assessment is coherent if and only if $p_3 = p_1 p_2$ and p_4 belong to the interval $[0, 1 - p_1]$.

In[21]:=
```
F={{A,B},{B,C},{A,C},{E,B}};
prob={0.2,0,0,0.5};
```

In[28]:=
```
p = ConstrainedMin[ob1,equations,d];
```

In[29]:=
```
sol1=d/.p[[2]]
```

Out[29]=
```
{0, 0, 0, 1., 0}
```

In[32]:=
```
solutions[x_,y_]:=Module[{ob,mx,sol,verasol},
ob=Apply[Plus,x];
mx=ConstrainedMax[ob,equations,d];
sol=d/.mx[[2]]; verasol=(y+sol)/2]
```

In[33]:=
```
autol=solutions[nul[d,sol1],sol1]
```

Out[33]=
```
{0, 0, 0, 0.5, 0.5}
```

In[34]:=
```
auto2=solutions[nul[d,auto1],auto1]
```

Out[34]=
```
{0, 0, 0, 0.75, 0.25}
```

In[35]:=
```
If[Count[auto1,0]==Count[auto2,0],
   typeA=nul[d,sol1];gg=Table[F[[i,2]],
   {i,1,Length[prob]}];
   AA={0,0,0,0};facs=Table[d[[i]]->i,
   {i,Length[d]}]; ddfacs=Dispatch[facs];
Do[If[Apply[Plus,
      auto2[[event[gg[[i]]]/.ddfacs]]]==0,
      AA[[i]]=1, AA[[i]]=0], {i,Length[gg]}];
AA=Flatten[Position[AA,1]];
Print["A not empty ======>",AA];
control=Part[equations,AA];
Hc=Complement[d,typeA];obj1=Apply[Plus,Hc];
ctrldfam=Flatten[Append[Drop[typeA,-1], Hc]];
ctrlnorm=Apply[Plus,ctrldfam]==1;
G=Append[control,ctrlnorm];
Print["G  not empty =======>", G];
pp=ConstrainedMin[obj1,G,ctrldfam];
solut=ctrldfam/.pp[[2]],
Print["find other solutions"]]
```

Out[35]=
```
A not empty =====> {1, 4}
G not empty =====>

{0.2 (A + E + (B && !A && !E)) == A,
 0.5 (A + E + (B && !A && !E)) == E,
 A + E + (C && !B) + (B && !A && !E) == 1}

{0.2, 0.5, 0.3, 0}
```

REFERENCES

1. G. COLETTI, *Coherent numerical and ordinal probabilistic assessments*, IEEE Transactions on Systems, Man and Cybernetics, Vol. 24, no. 12 (1994), 1747-1754.
2. G. COLETTI, *Numerical and qualitative judgement in probabilistic expert systems*, in Probabilistic Methods in Expert Systems, ed. R. Scozzafava, SIS, Roma, 1993, 37-55.
3. L. CRISMA, *Prolungamenti di probabilità condizionate localmente coerenti*, Quaderni del Dip. di Mat. Appl. "B. de Finetti", Università di Trieste, no. 5, 1993.
4. B. DE FINETTI, *Theory of probability*, Wiley & Sons, London, 1974.
5. G. DI BIASE, *Sulla ricerca di un insieme di probabilità condizionate coerenti nei sistemi esperti*, in AMASES/94, pp. 357-368, Atti del XVIII Convegno AMASES, Modena, Italia, Pitagora, Bologna, 1994.
6. G. DI BIASE, *Handling conditional events and probability assessments by Mathematica*, in Mathematics with Vision, eds. V. Keranen & P. Mitic, Computational Mechanics Publications, Southampton, 1995, 133-140.
7. G. DI BIASE, & A. MATURO, *Checking the coherence of conditional probabilities in expert systems: remarks and algorithms*, Mathematical Models for Handling Partial Knowledge in Artificial Intelligence, ed G. Coletti, R. Scozzafava & D. Dubois, Plenum Press, New York, 1994, 191-200.
8. G. DI BIASE, & A. MATURO. *Updating knowledge through conditional events: theory and implementation*, in Wupes/94, pp. 29-40, Proceedings of the 3rd Workshop on Uncertainty Processing in Expert Systems, Trest, Czech Republic, University of Economics, Praha, 1994.
9. A. GILIO, *Criterio di penalizzazione e condizioni di coerenza nella valutazione soggettiva della probabilità*, Bollettino UMI, n. 7- 4-B (1990), 645-660.
10. A. GILIO, *Probabilistic consistency of knowledge bases in inference systems*, in Lecture Notes in Computer Sciences, eds. M. Clarke, R. Kruse & S. Moral, pp. 160-167, Springer Verlag, Berlin and New York, 1993.
11. A. GILIO, *Classi quasi additive di eventi e coerenza di probabilità condizionate*, Rendiconti dell'Istituto di Matematica dell'Univ. di Trieste, XXI (1989), 22-38.
12. A. GILIO, *Algorithms for conditional probability assessments*, Journal of Information & Optimization Sciences, Analytic Publishing Co., Delhi, (1995), to appear.

13. A. GILIO, & R. SCOZZAFAVA, *Le probabilità condizionate coerenti nei sistemi esperti*, in Ricerca Operativa e Intelligenza Artificiale, Giardini, Pisa, 1988, 22-29.
14. S. HOLZER. *Sulla nozione di coerenza per le probabilità subordinate*, Rendiconti dell'Istituto di Matematica dell'Univ. di Trieste, XVI (1984), 46-62.
15. R. MAEDER, Programming in Mathematica, 2nd. ed., Addison Wesley, New York, 1991.
16. E. REGAZZINI, *Sulle probabilità coerenti nel senso di de Finetti*, CLUEB, Bologna, 1983.
17. R. SCOZZAFAVA, *A real life approach to the teaching of probability*, Ratio Mathematica, 1994, 7, 75-82.
18. R. SCOZZAFAVA, *How to solve some critical examples by a proper use of coherent probability*, in Uncertainty in Intelligent Systems, eds. B. Bouchon-Menier, L. Valverde and R. Yager, North Holland-Elsevier, Amsterdam, 1993, 121-132.
19. R. SCOZZAFAVA, *A survey of some common misunderstandings concerning the role and meaning of finitely additive probabilities in statistical inference*, Statistica, vol. 44 (1984), 21-45.
20. G. SHAFER, *Conditional probability*, International Statistical Review, vol. 53, no. 3 (1985), 261-277.
21. P. VICIC, *Proprietà di relazioni riguardanti il confronto fra eventi*, Quaderni del Dip. di Mat. Appl. "B. de Finetti", Università di Trieste, no. 11, 1993.