The Consistency Problem
in Belief and Probability Assessments

Andrea Capotorti
Dipartimento di Matematica
via Vanvitelli 1
06123 Perugia Italy
capot@gauss.dipmat.unipg.it

Barbara Vantaggi
Dipartimento di Matematica
via Vanvitelli 1
06123 Perugia Italy
vant@gauss.dipmat.unipg.it

Abstract

Among quantitative approaches for handling partial knowledge in Artificial Intelligence (AI), belief and probability functions are the most common. A domain expert can usually give a partial evaluation only on few events (generally without a particular logic frame), on which he has information at the moment. Hence, if we want to see this evaluation as a restriction of a particular kind of function representing uncertainty, we must prove consistency of these values. To do this we need to introduce computable consistency properties (as de Finetti's coherence principle for probability assessments). The main problem is computational complexity. In this paper we show cases with a right balance between dimensional complexity and frame richness.

PROBABILITY ASSESSMENTS

1 INTRODUCTION

One of the possible quantitative methods for handling partial knowledge in Artificial Intelligence (AI) is the probabilistic approach. The classical probability theory has been considered not a good tool because it requires a fixed structure (an algebra) and it asks an expert to evaluate all the elements of discernment. This is hard or impossible to obtain. For these reasons different schemes for reasoning with uncertainty have been developed in the qualitative (qualitative probabilistic networks [11], comparative probability [1]) and quantitative directions (Dempster-Shafer theory [8] cap.9, Fuzzy theory [12], probability theory) or in both (Coletti [2], Druzdzel and van der Gaag [5]).

All these models have specific problems and there is still a wide debate going on in literature about which would be the most appropriate for handling partial knowledge.

In this paper we deal with some problems that arise when we want to know if a partial assessment agrees with the basic properties of the functions chosen to manage partial knowledge, in particular probability and belief functions. These two approaches present similar difficulties both for checking the consistency of evaluations and for the numerical complexity of the entities involved the computation. The order of complexity of the latter is usually worse than the one of the former.

We provide meaningful examples, both with probability and belief assessments, to illustrate this “conflict” between freedom of evaluation and consistency. In the general case the proof of consistency is carried out by solving a linear system. We have identified particular logical situations where the problem of solving the systems can be skipped by checking further properties involving only the initial values.

2 PROBABILITY ASSESSMENTS

Let \( \mathcal{X} \) be a finite set of events that we will denote with \( E_i \), \( i = 1, \ldots, n \). The classical probability theory requires to endow \( \mathcal{X} \) with a particular frame, the algebra \( A \), so that a probability evaluation is a function

\[
P : A \rightarrow [0,1]
\]

De Finetti’s theory [4] does not require the algebra frame allowing to evaluate only the events of direct interest at the moment. This theory is particularly flexible also because it includes tools for reaching a coherent enlargement when the expert has to evaluate new propositions. For this reason some authors as Coletti et al. in [3] use this theory in the management of uncertainty in AI.

All possible assessments are those which satisfy a consistency condition called the “coherence principle”. Given the assessments \( P(E_i) = p_i \), \( i = 1, \ldots, n \), we need to introduce the atoms generated by the \( E_i \)’s. An atom \( C_k \) is an event of the form

\[
C_k = (\bigwedge_{j \in K} E_j) \wedge (\bigwedge_{j \in \{1, \ldots, n\} \setminus K} E_j^c)
\]

for some \( K \subseteq \{1, \ldots, n\} \). It is well known that the cardinality of atoms is \( s \leq 2^n \) and \( s \) becomes smaller when the number of logical relations among events increases.
The proof of coherence consists in checking the existence of a probability distribution \( w = (w_1, \ldots, w_s)^t \) on the atoms, being a solution of the following linear system:

\[
S1) \quad \begin{cases} \quad Aw = p^+ \\ \quad w \geq 0 \end{cases}
\]

where \( A \) is a \((n+1) \times s\) matrix composed of the \(n \times s\) matrix \([a_{ij}]\) bordered by an \((n+1)\)-st row of 1, with

\[
a_{ij} = \begin{cases} \quad 1 & \text{if } C_j \subseteq E_i \\ \quad 0 & \text{if } C_j \not\subseteq E_i \end{cases}
\]

and \( p^+ = (p_1, \ldots, p_n, 1)^t \).

We give now an example of inconsistent partial assessment.

**EXAMPLE 1**

Let \( A, B \) and \( C \) be three events such that:

\[
A \wedge B \wedge C = \emptyset \\
A \wedge B \neq \emptyset \\
C = (A \vee B) \wedge (A \wedge B)^c
\]

The values \( f(A) = 0.1, f(B) = 0.5, f(C) = 0.3 \) are not coherent because the associated system \( S1 \) has not a solution.

The concept of coherence has an interpretation: in an hypothetical bet, where the \( r_ip_i \) denote the stakes and the \( r_i \) the amounts, there is no sure loss or sure win. Equivalently, the \( p_i \)'s are coherent if

\[
\sup_{i=1}^n r_i(E_i - p_i) \geq 0 \quad \forall r_i \in \mathbb{R} \quad i = 1, \ldots, n
\]

where the sup is taken over all possible truth values of the \( E_i \) and we denote the event \( E_i \) and its truth value with the same symbol. Remember that, if the class of events is an algebra, any assessment such that at least the three basic properties of probability functions hold, is coherent. The main problem in this approach is that the number \( s \) of atoms is, in general, of \( O(2^n) \). The same problem has been pointed out by Hansen et al. [7]. They study the coherence problem (called PSAT) with linear programming techniques and they assert PSAT is NP-hard, as it is in NP and it is a generalization of the NP-hard problem SAT (Georgakopoulos et al. [6]). They propose some heuristics based on the column generation technique leading to numerical solution of large instances in a reasonable computing time. On the contrary we want to characterize situations leading to find analytical solution. In particular we look for cases in which, either the proof of coherence can be split into sub-problems, or it is not necessary at all. We first introduce another theory which has analogous problems.

### 3 BELIEF FUNCTIONS

An other well-known quantitative approach for handling partial knowledge is the Dempster-Shafer theory. This theory is a generalization of probability because super-additive functions (belief functions) are used to evaluate the uncertainty. Also belief functions need a well-defined frame. In fact a belief function is a monotone map defined on \( A = 2^x \), the algebra generated by a finite set of events \( X \), as

\[
Bel : A \longrightarrow [0, 1]
\]

such that

1. \( Bel(\emptyset) = 0 \) and \( Bel(\Omega) = 1 \)
2. \( \forall A_1, \ldots, A_m \in A, m \geq 2 \)

\[
Bel(\bigvee A_i) \geq \sum_{I \subseteq \{1, \ldots, m\}} (-1)^{|I|+1} Bel(\bigcap A_i) \quad (|I| \text{ is the cardinal of set } I)
\]

Every belief function can be characterized by a Möbius inverse (called basic assignment) which is the mapping \( m \) such that

i. \( m(\emptyset) = 0 \)
ii. \( Bel(A) = \sum_{B \subseteq A} m(B) \)

Note that the problem of knowing if a generic assessment is a belief function is equivalent to finding a basic assignment. In other words, numbering from 1 to \( 2^s \) the elements of \( A \), an evaluation \( b = (b_1, \ldots, b_{2^s})^t \) is a belief function iff the following linear system has a solution \( m = (m_1, \ldots, m_{2^s})^t \)

\[
S2) \quad \begin{cases} \quad \Delta m = b \\ \quad m \geq 0 \end{cases}
\]

where \( \Delta \) is a \((2^s) \times (2^s)\) matrix \([\delta_A(B)]\), with

\[
\delta_A(B) = \begin{cases} \quad 1 & \text{if } B \subseteq A \\ \quad 0 & \text{otherwise} \end{cases}
\]

Hence the complexity is \( O(2^s) \) that is much bigger than the complexity of the system \( S1 \). Moreover, in the examples of judgement through belief function, only partial assessment \( \alpha = (\alpha_1, \ldots, \alpha_k) \) with \( \alpha_i = f(A_i) \quad i = 1, \ldots, k \) and \( A_i \in A \) is given, without dealing at all with the problem of the proof of its consistency. Then, to prove consistency has to be found a solution of the following system

\[
S3) \quad \begin{cases} \quad \Delta' m = \alpha \\ \quad m \geq 0 \end{cases}
\]

where \( \Delta' \) is the \((k+1) \times 2^s\) matrix, sub-matrix of \( \Delta \). Regoli in [9] has shown that consistency is equivalent.
SURE CONSISTENCY CASES

In this section we show some cases in which the proof of consistency is made easier by particular logic relations. Obviously, if the values of an assessment are outside [0, 1] or they do not satisfy the monotonic property, the assessment is not coherent. For this reason we analyse only situations where these properties hold.

I) Let $E_1, \ldots, E_n$ a finite set of events such that the generated atoms are exactly $2^n$. This situation occurs when there is no any logic constraint among the $E_i$'s, hence they are logically independent. Any assessment $P(E_i) = p_i$ with $0 \leq p_i \leq 1$ $i = 1, \ldots, n$ is a coherent probability distribution. Because the number of atoms is $2^n$, the rank of matrix $A$ in system $S1$ is $(n + 1)$ and hence the existence of a solution is assured. Moreover, in this case, a non-negative solution $w$ may be built up. We can assume, without loss of generality, that $0 \leq p_1 \leq \ldots \leq p_n \leq 1$. Consider the following atoms

$$C_1 = E_1 \land \ldots \land E_n$$

$$C_i = E_i^{(1)} \land \ldots \land E_i^{(n-1)} \land E_i \land \ldots \land E_n$$

$$C_{n+1} = E_1^{(1)} \land \ldots \land E_{n+1}^{(1)}$$

and denote with

$$w_1 = p_1$$
$$w_2 = p_2 - p_1$$
$$w_i = p_i - p_{i-1}$$
$$w_n = p_n - p_{n-1}$$
$$w_{n+1} = 1 - p_n$$

Then, the vector $w = (w_1, \ldots, w_{n+1}, 0, \ldots, 0)^t$ is a solution of system $S1$ and hence the assessment $p = (p_1, \ldots, p_n)$ is coherent.

II) In the previous case we required the number of atoms to be exactly $2^n$. This condition can be relaxed if some particular conditions on the $p_i$'s hold. For example:

- if $p_n = 1$ then $C_{n+1}$ may be the empty set, and this happens when the disjunction of all the $E_i$'s is the sure event
- if $p_{i-1} = p_i$ then $C_i$ may be the empty set, and this happens when $E_i \subseteq E_{i-1}$
- more in general:

$$p_i = \ldots = p_i^{(n)}$$
$$p_{(i+1)} = \ldots = p_{(i+1)^{(n)}}$$

with $p_{(i+r)^{(n)}} < p_{(i+r+1)^{(n)}}$, $r = 0, \ldots, s$ then all the atoms $C_{(i+r)^{(n)}}, \ldots, C_{(i+r+s)^{(n)}}$, $r = 0, \ldots, s$ except one may be empty.
From the building procedure of solution \( w \) in I), it follows that, when \( p_1 < p_2 < \ldots < p_n \), the existence of the atoms \( C_1, \ldots, C_{n+1} \) is sufficient for the coherence. Otherwise, if \( p_{i-1} < p_i = \ldots = p_{i+r} < p_{i+r+1} \) then the existence of at least one of the \( C_i, \ldots, C_{i+r} \) is sufficient.

III) If available logic constraints are only \( E_i \in C_j \), for some \( i, j \in \{1, \ldots, n\} \), any monotonic assessment is coherent. In this case the only empty atoms are those with \( E_i \wedge E_j \) in their expression. Suppose to order the \( p_i s \) as in I), hence \( i < j \) and that the atoms \( C_1, \ldots, C_{n+1} \) are non-empty. Then the distribution \( w = (w_1, \ldots, w_{n+1}, 0, \ldots, 0)^t \) (obviously the length of \( w \) is less than \( 2^n \)) is a solution of S1.

IV) We analyse now properties similar to 2. of belief function when there is not the structure of algebra. In particular, starting from a finite number of events \( E_1, \ldots, E_n \), let \( \mathcal{E} \) be the multiplicative class

\[ \{ \bigwedge E_i ; I \subseteq \{1, \ldots, n\} \} \]

In the elements of \( \mathcal{E} \) complementary events \( E_i \) are not involved. Shafer in [10] gives a \( \infty \)-monotonicity condition that, with \( Bel(\emptyset) = 0 \) and \( Bel(\Omega) = 1 \), guarantees consistency for an evaluation on a generic multiplicative class. In our case that \( \mathcal{E} \) is generated by \( E_1, \ldots, E_n \) the property can be simplified in the following:

\[ Bel(\bigwedge E_i) = b(A_i) \forall I \subseteq \{1, \ldots, n\} \]

is a restriction of a belief function on \( \mathcal{E} \) iff

\[ P0) \quad 1 \geq \sum_{J \subseteq \{1, \ldots, n\}} (-1)^{|J|+1} b(A_J) \]

\[ P1) \quad b(A_I) \geq \sum_{J \subseteq \{1, \ldots, n\}} (-1)^{|J|+1} b(A_J) \forall I \subseteq \{1, \ldots, n\} \]

The proof is made solving the linear system S3 by forward-substitution. We now extend \( \mathcal{E} \) to the frame \( \mathcal{L} = \mathcal{E} \cup \mathcal{F}_1 \cup \ldots \cup \mathcal{F}_d \) where the \( \mathcal{F}_i = \mathcal{F}_{K_i} \) are classes composed by the disjunction of \( E_j \)s with \( j \in K_i \). Then an assessment

\[ Bel(\bigwedge E_i) = b(A_i) \forall I \subseteq \{1, \ldots, n\} \]

\[ Bel(\bigvee E_i) = b(V_E) \forall F_j \text{ s.t. } (\bigvee E_i) \in F_j \]

is a restriction of a belief function on \( \mathcal{L} \) iff

\[ P0') \quad 1 \geq \sum_{h=1}^d (-1)^{h+1} \sum_{N_h \subseteq \{1, \ldots, d\}} b_{V_E} \]

\[ \text{with } Z_h = \{ j \in N_h K_j \} \]

V) Properties P1 and P2 are useful also when we have not exact classes as \( \mathcal{E} \) or \( \mathcal{L} \). If some evaluation on \( \mathcal{E} \) or \( \mathcal{L} \) is missing, the properties are reduced to parametric inequalities systems, where the parameters are the unknown values. The existence of a solution of such a system is easy to check because we can solve by forward-substitution. For example:

Let \( \mathcal{E} \) be the multiplicative class generated by

\[ P1) \quad b(A_I) \geq \sum_{J \subseteq I} (-1)^{|J|+1} b(A_J) \]

\[ \forall I \subseteq \{1, \ldots, n\} \]

\[ P2) \quad b(V_E) \geq \sum_{F \in \mathcal{F}_i} (-1)^{|F|} b(V_E) + \]

\[ (+(-1)^{|F|+1} b(A_{E})) \]

\[ \forall F \text{ s.t. } (\bigvee E) \in F \]

Also in this case the proof is made solving the linear system S3 by forward-substitution.

Note that in EXAMPLE 2 and 3 it is easy to check that condition P1 and P2, respectively, do not hold. In EXAMPLE 2 we have the inconsistency

\[ 0.8 = f(A_4) \geq f(A_2 \cup A_3) \]

\[ = f(A_2) + f(A_3) - f(A_1) = 0.9 \]

Likewise in EXAMPLE 3 we have the inconsistency

\[ 0.05 = f(A_1 \cup A_2 \cup A_3) \geq f(A_1 \cup A_2) + f(A_1 \cup A_3) + f(A_2 \cup A_3) - f(A_1) - f(A_2) - f(A_3) + f(A_1 \cap A_2 \cap A_3) = 0.055 \]

even if property 2.

\[ 0.05 = f(A_1 \cup A_2 \cup A_3) \geq f(A_1) + f(A_2) + f(A_3) + f(A_1 \cap A_3) - f(A_1 \cup A_2) - f(A_2 \cup A_3) + f(A_1 \cap A_2 \cap A_3) = 0.04 \]

The power of these conditions is that for the proof of consistency we can use only the values of the assessment, without generating the matrix \( \Delta' \) in system S3. So we deal only with \( 2^n \) quantities rather then \( 2^{2^n} \).

Obviously analogous properties hold for probability distribution, the advantage being, in this case, that we can avoid to build the matrix A in system S1. However, there are \( 2^n \) numerical inequalities to check.
The following assessment is given
\[
\begin{align*}
I(E_1 \land E_2 \land E_3 \land E_4) &= 0.01 \\
I(E_i \land E_j \land E_k) &= 0.02 \forall i \neq j \neq k \\
I(E_1 \land E_2) &= 0.03 \\
I(E_2 \land E_3) &= 0.035 \\
I(E_2 \land E_4) &= 0.04 \\
I(E_1) &= 0.06 \ i = 1, 2, 3 \\
I(E_4) &= 0.041 \\
\end{align*}
\]

and
\[
\begin{align*}
\theta_{14} &= f(E_1 \land E_4) \\
\theta_{23} &= f(E_2 \land E_3) \\
\theta_{34} &= f(E_3 \land E_4) \\
\end{align*}
\]

The system given by inequalities of property \( P1 \) with \( \theta_{ij} \)'s is than a non-solvable, because the following sub-system
\[
\begin{align*}
\begin{cases}
\theta_{14} &\text{AND} &0.03 \\
\theta_{23} &\text{AND} &0.03 \\
\theta_{34} &\text{AND} &0.041 \\
\theta_{14} + \theta_{34} - 0.01 &\text{AND} &0.03 \\
\end{cases}
\end{align*}
\]

has not solution. This shows (it is possible to obtain the same result by solving the system \( S3 \) that the initial assessment is not a restriction of a belief function.

5 CONCLUSIONS

From the properties shown in this paper we can deduce a strong dependence between the proof of consistency and logical-numerical relations among events. The more complex is the frame of discernment, the more simple is the check of coherence. Facing the problems of complexity and consistency of both probability and belief assignments, we have investigated particular situations in which we can easily assert if an evaluation is consistent or not. Following this strategy, we will examine in the future if other such frames or numerical relations exist. In particular, we will examine situations in which, from the logical relations among events, we might determine particular sub-frames. Therefore, the problem will be split in sub-problems related by constraints. An elementary case of this kind happens when we can detect \( k \geq 2 \) families \( L_1, \ldots, L_k \) composed respectively by all conjunctions and disjunctions among \( E_{i1}, \ldots, E_{in} \), and with assessments \( f_i, i = 1, \ldots, k \), each of the \( f_i \) independently consistent, with the only constraint that
\[
\sum_{i=1}^{k} f_i(E_{i1} \lor \ldots \lor E_{in}) \leq 1.
\]

In this situations the frame \( L_i \) is restrictive, even if less than algebra, so we will investigate poorer logic relations.

References