Marginal Problem in Different Calculi of AI*

Milan Studený**

Czech Academy of Sciences, Inst. Inform. Theory and Automation,
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic

Abstract. By the marginal problem we understand the problem of the existence of a global (full-dimensional) knowledge representation which has prescribed less-dimensional representations as marginals. The paper deals with this problem in several calculi of AI: probabilistic reasoning, theory of relational databases, possibility theory, Dempster-Shafer's theory of belief functions, Spohn's theory of ordinal conditional functions. The following result, already known in probabilistic framework and in the framework of relational databases, is shown also for the other calculi: the running intersection property is the necessary and sufficient condition for pairwise compatibility of prescribed less-dimensional knowledge representations being equivalent to the existence of a global representation. Moreover, a simple method of solving the marginal problem in the possibilistic framework and its subframeworks is given.

1 Introduction

Dealing with integration of knowledge in probabilistic expert systems one encounters the problem of consistency well-known as the *marginal problem* [8]: having prescribed a collection of less-dimensional probability measures (which represent pieces of knowledge given by experts - see [11]) one should recognize whether there exists a joint multidimensional probability measure having the prescribed less-dimensional measures as marginals (such a joint measure then could represent global knowledge kept by an expert system).

Of course, an analogous problem can be expected when one tries to model expert knowledge within another calculus for uncertainty management. Concretely, this paper is concerned with the marginal problem in the following branches of AI:

- probabilistic reasoning
- theory of relational databases
- theory of ordinal conditional functions
- possibility theory
- theory of belief functions.

** E-mail: studeny@utia.cas.cz

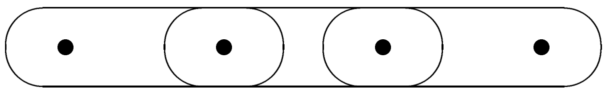
* Supported by the grant n. 201/94/0471 "Marginal problem and its application" of the Grant Agency of Czech Republic.

As concerns the probabilistic framework[3] no direct method of solving the marginal problem is known but there exists an asymptotic method. Using the collection of prescribed less-dimensional measures one can define by means of the so-called *iterative proportional fitting procedure* [3] a sequence of multidimensional probability measures which is proved in [2] to converge iff there exists a joint measure having the prescribed measures as marginals. The limit measure then has the prescribed marginals and minimizes I-divergence within the class of such joint measures.

Nevertheless, one can sometimes evade this iterative procedure as the global consistency is under a certain structural condition put on the collection of underlying attribute sets[4] equivalent to the condition of pairwise compatibility which is easy to verify or disprove. Kellerer [9] showed that the global consistency is equivalent with the pairwise compatibility iff the collection $\mathcal{S}$ of underlying sets satisfies the *running intersection property* (see also [10, 8]):

there exists an ordering $S_1, \ldots, S_n$ of $\mathcal{S}$ such that
$\forall j \geq 2 \quad \exists i \quad 1 \leq i < j \qquad S_j \cap (\bigcup_{k<j} S_k) \subset S_i$.

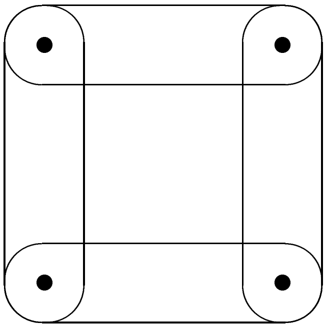
This condition has a meaning of acyclicity of the hypergraph $\mathcal{S}$ (this terminology was accepted in [1]). For example, the chain figured below satisfies the running intersection property.



On the other hand, the cycle below does not satisfy it.

[3] In this paper we restrict our attention to probability measures on finite sets.

[4] By an attribute we understand an elementary symptom or variable in consideration of an expert system. The corresponding knowledge is represented differently in different calculi. In probabilistic reasoning it is represented by a one-dimensional probability measure. Every expert can give evidence concerning certain area, i.e. his statements refer to a small set of attributes. Thus, the piece of knowledge given by an expert is represented by a less-dimensional probability measure embracing exactly attributes from the mentioned set (= the underlying attribute set).



Similar results were later shown in the theory of relational databases. A simple direct method of solving the marginal problem in this framework is known – see [10]. It consists in verifying whether a certain multidimensional database relation, made of the prescribed less-dimensional database relations by a concrete procedure, has the prescribed relations as marginals. Moreover, the running intersection property was also shown to be a necessary and sufficient condition for pairwise compatibility being equivalent to global consistency in this framework [1].

The aim of this paper is to show the results concerning the running intersection property for the other mentioned calculi as well. Note that this result for the calculus of ordinal conditional functions was recently proved in [16]. The procedure from [16] can also be used in all the other calculi – one only has to give several basic constructions (specific for a calculus). This is done in this paper in order to make a comprehensive survey. Moreover, the above mentioned simple direct method of solving the marginal problem for relational databases is extended to the possibilistic calculus in this paper.

The next section recalls how knowledge is represented in all the calculi. The third section then describes the mentioned method of solving the marginal problem in the possibilistic calculus and gives an example showing that the global consistency of a collection of possibility measures is indeed strictly stronger than its global consistency in sense of Dempster-Shafer theory. In the fourth section the main constructions allowing us to show the main results are given.

2 Basic definitions

This section recalls how knowledge is represented in all the mentioned calculi of AI. The most of these calculi are constructed to be as general as possible and some readers can object that the definitions below are restrictive. But they express the essence of these calculi and make possible certain unifying point of view: the calculus of Dempster-Shafer theory is considered to be the most general framework and the other approaches are incorporated as its subframeworks.

All above mentioned calculi will have some common setting in this paper. In the following we suppose that N is a nonempty finite set of *attributes.* Moreover, a nonempty finite set X_i called the *frame*[5] corresponds to each attribute $i \in N$. Whenever $\emptyset \neq S \subset N$ the symbol X_S denotes the cartesian product $\Pi_{i \in S}\, X_i$. Having an element $x \in X_N$ and $\emptyset \neq S \subset N$ the projection of x to X_S (i.e. the element of X_S whose components coincide with the components of x) will be denoted by x_S. The power set of a set Y will be denoted by $\exp Y$.

We start our definition survey with the probabilistic calculus which is probably the most developed approach for dealing uncertainty in AI.

Definition 1 (probability measure)
A *probability distribution* over N is a nonnegative real function $P : X_N \to \langle 0, \infty)$ satisfying $\sum\{P(x);\, x \in X_N\} = 1$. The formula

$P(A) = \sum\{P(x);\, x \in A\}$ (for $A \subset X_N$)

then defines a set function (on $\exp X_N$) called a *probability measure* over N. Whenever $\emptyset \neq S \subset N$ and P is a probability measure over N, then its *marginal* on S is a probability measure P^S over S defined as follows (of course $P^N \equiv P$):

$P^S(A) = P(A \times X_{N \setminus S})$ for $A \subset X_S$, $\emptyset \neq S \neq N$.

Note that a marginal on a set of attributes will be always denoted by the symbol of the original (knowledge representation) having moreover as the upper index the symbol of the attribute set (the same principle will be followed also in the other calculi).

Another framework where the marginal problem has already been studied is the theory of relational databases.

Definition 2 (database relation)
A *database relation* over N is a nonempty subset of X_N. Whenever $\emptyset \neq S \subset N$ and R is a database relation over N, then its *marginal* on S is a database relation R^S over S defined as follows ($\mathsf{R}^N = \mathsf{R}$):

$u \in \mathsf{R}^S \Leftrightarrow [\ (u, v) \in \mathsf{R}$ for some $v \in X_{N \setminus S}]$
whenever $u \in X_S$.

A further theory in our focus is Spohn's theory of ordinal conditional functions [15]. This theory gives a tool for mathematical description of dynamic handling of deterministic epistemilogy and in this sense it constitutes a counterpart of the probabilistic approach. Researchers in AI paid attention especially to a special class of natural conditional functions [7, 14]:

Definition 3 (natural conditional function)
Having a nonnegative integer function $\kappa : X_N \to \{0, 1, \ldots\}$ satisfying $\min\{\kappa(x);\, x \in X_N\} = 0$, the formula

$\kappa(A) = \min\{\kappa(x);\, x \in A\}$ (for $\emptyset \neq A \subset X_N$)

[5] The frame is the set of "possible" values for the considered attribute.

defines a set function on $(\exp X_N) \setminus \{\emptyset\}$ called a *natural conditional function* (NCF) over N. Whenever $\emptyset \neq S \subset N$ and κ is an NCF over N, then its *marginal* on S is an NCF κ^S over S defined as follows ($\kappa^N \equiv \kappa$):

$$\kappa^S(A) = \kappa(A \times X_{N \setminus S}) \qquad \text{for } \emptyset \neq A \subset X_S,\ \emptyset \neq S \neq N.$$

The next calculus is possibility theory which was proposed by Zadeh [17] as a model for quantification of judgements on the basis of fuzzy theory and later developed by Dubois and Prade [5].

Definition 4 (possibility measure)
A *possibility distribution* over N is a real function $\pi : X_N \to \langle 0, 1\rangle$ satisfying $\max\{\pi(x);\ x \in X_N\} = 1$. The formula

$$\pi(A) = \max\{\pi(x);\ x \in A\} \qquad (\text{for } \emptyset \neq A \subset X_N)$$

then defines a set function on $(\exp X_N) \setminus \{0\}$ called a *possibility measure* over N. Whenever $\emptyset \neq S \subset N$ and π is a possibility measure over N, then its *marginal* on S is a possibility measure π^S over S defined as follows ($\pi^N \equiv \pi$):

$$\pi^S(A) = \pi(A \times X_{N \setminus S}) \qquad \text{for } \emptyset \neq A \subset X_S,\ \emptyset \neq S \neq N.$$

One of the most popular approaches for dealing with uncertainty in AI is Dempster-Shafer theory [3, 13]. Knowledge can be described here in several equivalent ways (belief or respectively commonality or plausibility function), we chose the concept of basic probability assignment.

Definition 5 (basic probability assignment)
A *basic probability assignment* (BPA) over N is a real function $\mathbf{m} : \exp X_N \to \langle 0, \infty)$ satisfying $\sum\{\mathbf{m}(A); A \subset X_N\}$ and $\mathbf{m}(\emptyset) = 0$. Whenever $\emptyset \neq S \subset N$ and $\mathbf{m}$ is a BPA over N, then its *marginal* on S is a BPA $\mathbf{m}^S$ over S defined as follows (of course $\mathbf{m}^N \equiv \mathbf{m}$):

$$\mathbf{m}^S(A) = \sum\{\mathbf{m}(\mathsf{R});\ \mathsf{R} \subset X_N \quad \mathsf{R}^S = A\} \qquad \text{for } A \subset X_S, \text{ (see Definition 2 for the symbol } \mathsf{R}^S).$$

Focal elements are the sets $A \subset X_N$ with $\mathbf{m}(A) > 0$.

The calculi above can be compared each other. For example one can assign a posssibility measure π_κ to every NCF κ by means of the formula:

$$\pi_\kappa(A) = e^{-\kappa(A)} \qquad \text{for } \emptyset \neq A \subset X_N.$$

The mapping $\kappa \to \pi_\kappa$ is injective and respects marginals i.e.

$$(\pi_\kappa)^S = \pi_{(\kappa^S)} \qquad \text{for NCF } \kappa \text{ over } N,\ \emptyset \neq S \subset N.$$

If there exists an injective mapping respecting marginals from a calculus to another calculus we shall say that the former calculus is a *subframework* of the latter one.
Thus, database relations are a subfamework of possibility measures since a possibility measure π_R is assigned to every database relation R:

$$\pi_{\mathsf{R}}(A) = \begin{cases} 1 & \text{if } A \cap \mathsf{R} \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \qquad (\emptyset \neq A \subset X_N)$$

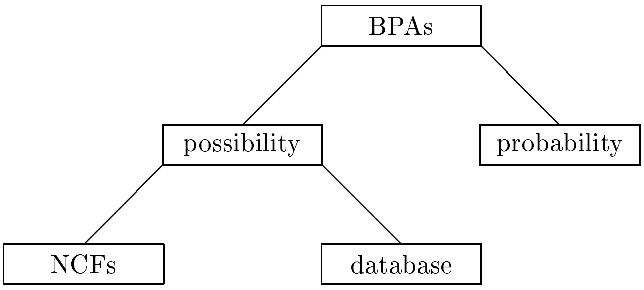
and probability measures are a subframework of BPAs since a BPA $\mathbf{m}_P$ is assigned to every probability measure P:

$$\mathbf{m}_P(A) = \begin{cases} P(x) & \text{if } A = \{x\} \text{ for } x \in X_N \\ 0 & \text{otherwise} \end{cases}$$

and possibility measures are a subframework of BPAs, since every possibility measure π can be identified with a BPA $\mathbf{m}_\pi$ whose collection of focal elements is a nest[6] and satisfies the relation

$$\pi(x) = \sum\{\mathbf{m}_\pi(B);\ x \in B \subset X_N\} \qquad \text{for } x \in X_N.$$

We left to the reader to verify that $\mathbf{m}_\pi$ is determined uniquely by these two conditions and that all the mappings are injective and respect marginals. Thus, the situation can be illustrated by the following picture.



Remark The reader may think that database relations are a subframework of probability measures since one can assign a probability distribution P_{R} to every database relation R:

$$P_{\mathsf{R}} = \begin{cases} (\operatorname{card} \mathsf{R})^{-1} & \text{if } x \in \mathsf{R} \\ 0 & \text{otherwise.} \end{cases}$$

Nevertheless, this mapping does not respect marginals and therefore it is not interesting from our 'marginal problem' point view. Indeed, one can take $N = \{1,2\},\ X_1 = X_2 = \{0,1\},\ \mathsf{R} = \{(0,0),(0,1),(1,0)\}$ and have $(P_{\mathsf{R}})^{\{1\}} \neq P_{(\mathsf{R}^{\{1\}})}$.

3 Marginal problem in possibility theory

This section gives a simple direct method of solving the marginal problem in the possibilistic framework. Moreover, an example shows that a collection of possibility measures may be globally consistent within the BPA-framework but not within the possibilistic framework.

Firstly, we give exact definitions of concepts connected with the marginal problem. They are shared by all the calculi we deal with in this paper.

[6] i.e. $A \subset B$ or $B \subset A$ for every two focal elements A, B

Definition 6 (compatibility, consistency)
Let us have in mind any of the calculi mentioned in section 2. Suppose that $\{k_S;\ S \in \mathcal{S}\}$ is a collection of knowledge representations within that calculus (where the lower index S in k_S denotes the nonempty set of underlying attributes). The collection $\{k_S; S \in \mathcal{S}\}$ is called *pairwise compatible* if and only if

$\forall S, T \in \mathcal{S}$ with $S \cap T \neq \emptyset$ $\ (k_S)^{S\cap T} = (k_T)^{S\cap T}$.

Moreover, $\{k_S;\ S \in \mathcal{S}\}$ is called *globally consistent* iff there exists a global knowledge representation k (having the set of underlying attributes N) such that $\forall\, S \in \mathcal{S} \quad k^S = k_S$.

It is evident that global consistency implies pairwise compatibility but the converse is not true. The following example shows it for all the mentioned calculi.

Example 1 (compatibility $\not\Rightarrow$ consistency)
Consider the global attribute set $N = \{1,\ldots,n\}$ where $n \geq 3$, the frames $X_i = \{0,1\}$ for $i \in N$ and the collection of attribute sets $\mathcal{S} = \{\,\{1,2\},\{2,3\},\ldots,\{n-1,n\},\{n,1\}\,\}$. Further details depend on calculi:

$\boxed{\text{a}}$ probability measures
Define probability distributions $\{P_S; S \in \mathcal{S}\}$ as follows:
$P_{\{n,1\}}(00) = P_{\{n,1\}}(11) = 0$
$P_{\{n,1\}}(01) = P_{\{n,1\}}(10) = 0.5$
and for remaining $S \in \mathcal{S}$
$P_S(00) = P_S(11) = 0.5 \quad P_S(01) = P_S(10) = 0$.
This collection is not globally consistent: having a probability distribution P over N with these prescribed marginals, the inequality $P(x) \leq P^S(x_S)$ implies $0 \leq P(x) \leq \min_{S\in\mathcal{S}} P_S(x_S) = 0$ for all $x \in X_N$ and this contradicts $\sum\{P(x);\ x \in X_N\} = 1$.

$\boxed{\text{b}}$ database relations
Put $\mathsf{R}_{\{n,1\}} = \{(0,1),(1,0)\}$ and $\mathsf{R}_S = \{(0,0),(1,1)\}$ for remaining $S \in \mathcal{S}$. This collection is not globally consistent as there is no database relation R over N having these prescribed marginals. Indeed, no $x \in X_N$ satisfies $\forall\, S \in \mathcal{S} \quad x_S \in \mathsf{R}_S$ and this implies a contradictory conclusion $\mathsf{R} = \emptyset$.

$\boxed{\text{c}}$ NCFs
Let us define NCFs as point functions on X_S:
$\kappa_{\{n,1\}}(00) = \kappa_{\{n,1\}}(11) = 1$
$\kappa_{\{n,1\}}(01) = \kappa_{\{n1\}}(10) = 0$
and for remaining $S \in \mathcal{S}$
$\kappa_S(00) = \kappa_S(11) = 0 \quad \kappa_S(01) = \kappa_S(10) = 1$.
Supposing an NCF κ over N has these marginals, the inequality $\kappa(x) \geq \kappa^S(x_S)$ implies $\kappa(x) \geq \max_{S\in\mathcal{S}}\ \kappa_S(x_S) = 1$ for all $x \in X_N$ and this contradicts $\min\{\kappa(x); x \in X_N\} = 0$.

$\boxed{\text{d}}$ possibility measures
Take $\pi_S \equiv \pi_{\mathsf{R}_S}$, $S \in \mathcal{S}$ where $\{\mathsf{R}_S; S \in \mathcal{S}\}$ are database relations from $\boxed{\text{b}}$. Supposing π is a possibility distribution having these prescribed marginals, the inequality $\pi(x) \leq \pi^S(x_S)$ implies $\pi(x) \leq \min_{S\in\mathcal{S}}\ \pi_S(x_S) = 0$ for all $x \in X_N$ and this contradicts $\max\{\pi(x); x \in X_N\} = 1$.

$\boxed{\text{e}}$ BPAs

Take $\mathbf{m}_S \equiv \mathbf{m}_{\pi_S}$, $S \in \mathcal{S}$ where $\{\pi_S; S \in \mathcal{S}\}$ are possibility measures from $\boxed{\text{d}}$. To disprove global consistency realize that every $\mathbf{m}_S$ has only one focal element (with assigned value 1): $\mathbf{m}_{\{n,1\}}$ has $\{(01),(10)\}$ as its focal element, any other $\mathbf{m}_S$ has $\{(00),(10)\}$. If $\mathbf{m}$ is a BPA over N having $\{\mathbf{m}_S; S \in \mathcal{S}\}$ as marginals and R is one of its focal elements, then its marginal R^S has to be a focal element of $\mathbf{m}_S$ (for every $S \in \mathcal{S}$). Hence by the procedure from $\boxed{\text{b}}$ derive $\mathsf{R} = \emptyset$ and this contradicts the definition of BPA.

As mentioned in Introduction, there is no direct method of testing global consistency within the probabilistic framework. Nevertheless, there exists such a method in the possibilistic framework.

Proposition 1 Suppose that $\{\pi_S; S \in \mathcal{S}\}$ is a collection of possibility distributions (when the lower index denotes the set of underlying attributes). Then $\{\pi_S; S \in \mathcal{S}\}$ is globally consistent if and only if the formula

$$\pi_*(x) = \min_{S \in \mathcal{S}} \pi_S(x_S) \qquad \text{for } x \in X_N$$

defines a possibility distribution whose marginals are $\{\pi_S; S \in \mathcal{S}\}$.

Proof: The sufficiency is evident. For necessity suppose that π is a posibility measure having $\{\pi_S; S \in \mathcal{S}\}$ as marginals. Then

$\boxed{1.}$ $\pi \leq \pi_*$

For each $x \in X_N$ and $S \in \mathcal{S}$ write $\pi(x) \leq \pi^S(x_S) = \pi_S(x_S)$ and use the definition of π_* .

$\boxed{2.}$ π_* is a possibility distribution

Evidently $0 \leq \pi_* \leq 1$; having $x_0 \in X_N$ with $\pi(x_0) = 1$ the preceding step gives $\pi_*(x_0) = 1$.

$\boxed{3.}$ $\forall S \in \mathcal{S} \quad (\pi_*)^S = \pi_S$

Having fixed $S \in \mathcal{S}$ and $z \in X_S$ by the definition of marginal find $y \in X_N$ with $z = y_S$ and $(\pi_*)^S(z) = \pi_*(y)$. Using the definition of π_* write $\pi_*(y) \leq \pi_S(y_S) = \pi_S(z)$ i.e. $(\pi_*)^S(z) \leq \pi_S(z) = \pi^S(z)$. On the other hand $\pi^S(z) \leq (\pi_*)^S(z)$ follows from $\boxed{1.}$ and therefore $(\pi_*)^S(z) = \pi_S(z) = \pi^S(z)$. □

The result above gives already published criteria for subframeworks[7]. A collection of database relations $\{\mathsf{R}_S; S \in \mathcal{S}\}$ is globally consistent iff the set $\mathsf{R}_* = \bigcap_{S \in \mathcal{S}} \mathsf{R}_S \times X_{N \setminus S}$ is a database relation having $\{\mathsf{R}_S; S \in \mathcal{S}\}$ as marginals [10]. A collection of NCFs $\{\kappa_S; S \in \mathcal{S}\}$ is globally consistent iff the function $\kappa_*(x) = \max_{S \in \mathcal{S}} \kappa_S(x_S)$ (for $x \in X_N$) determines an NCF over N having $\{\kappa_S; S \in \mathcal{S}\}$ as marginals [16]. These criteria can be derived from Proposition 1 owing to the following principle: supposing that possibility measures $\{\pi_S; S \in \mathcal{S}\}$ correspond to database relations (resp. NCFs) π_* gives a possibility measure corresponding to a database relation (resp. an NCF).

Therefore every collection of database relations (resp. NCFs) is globally consistent iff it is consistent within the possibilistic framework. Similarly, one can

[7] As concerns the test in the possibilistic framework I only found in [6] p. 6–7 a procedure where π_* was defined as one step of a procedure of approximate reasoning [18]. But the mentioned procedure computes one-dimensional marginals of π_* and the authors of [6, 18] are not interested in the connection to the starting possibility measures which are not supposed to be pairwise compatible.

show that a collection of probability measures is globally consistent iff it is consistent within the BPA-framework[8]. Nevertheless, a collection of possibility measures need not be globally consistent although it is globally consistent within the BPA-framework as the following example shows.

Example 2 Put $N = \{1,2,3\}, X_i = \{0,1\}$ for $i \in N$ and $\mathcal{S} = \{\,\{1,2\},\{1,3\},\{2,3\}\,\}$.
Consider a collection of possibility distributions $\{\pi_S; S \in \mathcal{S}\}$ defined as follows ($S \in \mathcal{S}$):
$\pi_S(00) = \pi_S(11) = \frac{2}{3} \quad \pi_S(01) = \pi_S(10) = 1$.
This collection is not globally consistent as the function $\pi_*(x) = \min_{S\in\mathcal{S}} \ \pi_S(x_S) = \frac{2}{3}$ is not a possibility distribution. But, the collection of BPAs $\{\mathbf{m}_{\pi_S}; S \in \mathcal{S}\}$ is globally consistent as one can consider the BPA $\mathbf{m}$ with three focal elements:
$\{\,(010),(011),(100),(101)\,\}$,
$\{\,(001),(011),(100),(110)\,\}$,
$\{\,(001),(101),(010),(110)\,\}$
with assigned values $\frac{1}{3}$.

4 Solvable collections

In this section the collections of attribute sets for which pairwise compatibility is equivalent to global consistency are studied. We show that these collections are characterized within all mentioned calculi by means of the running intersection property.

We start with some definitions which are, of course, shared by all the studied calculi.

Definition 7 (solvable collection, reduced collection)
Having in mind any of the calculi mentioned in section 2 a collection of nonempty attribute sets $\mathcal{S}$ will be called *solvable* within that calculus iff every pairwise compatible collection of knowledge representations, whose collection of attribute sets is $\mathcal{S}$, is globally consistent.
$\mathcal{S}$ will be called *reduced* iff
$\forall\, A, B \in \mathcal{S}$ neither $A \subset B$ nor $B \subset A$.
If $\emptyset \neq T \subset N$, then the *contraction* of $\mathcal{S}$ to T denoted by $\mathcal{S} \wedge T$ is defined as the collection of maximal sets[9] of $\{S \cap T;\ S \in \mathcal{S} \ \ S \cap T \neq \emptyset\}$.

The method used in [16] to show the necessity of the running intersection property for solvable collections within the NCF-framework in fact does not depend on a particular calculus (see Lemmas 8,9 and Theorem 2 in [16]). One only needs to show that a collection of attribute sets $\mathcal{S}$ is not solvable in two following cases:

$$[\mathbf{a}] \left\{ \begin{array}{l} \mathcal{S} \text{ contains a sequence } S_1,\ldots,S_n \ (n \geq 3) \\ \ \text{such that } \ \forall\, i = 1,\ldots,n \qquad (S_{n+1} \equiv S_1) \\ S_i \cap S_{i+1} \setminus \bigcup(\,\mathcal{S} \setminus \{S_i, S_{i+1}\}) \neq \emptyset. \end{array} \right.$$

[8] Hint: Without loss of generality suppose $N = \bigcup \mathcal{S}$. Projections (=database marginals) of every focal element of a global BPA $\mathbf{m}$ having the prescribed marginals must be focal elements of marginals i.e. singletons. Hence, every focal element of $\mathbf{m}$ is a singleton.

[9] $A \in \mathcal{T}$ is maximal in $\mathcal{T}$ iff $[B \in \mathcal{T},\ A \subset B] \Rightarrow B = A$.

[b] $\begin{cases} \mathcal{S} \text{ is reduced, } \mathrm{card}\,\mathcal{S} \geq 2 \text{ and} \\ \forall i,j \in \bigcup \mathcal{S} \ \ \exists\, S \in \mathcal{S} \ \text{ with } \ i,j \in S. \end{cases}$

Also the proof of sufficiency requires only to show that a collection $\mathcal{S}$ with $\mathrm{card}\,\mathcal{S} = 2$ is solvable – see [16]. Thus, in the sequel we only verify these facts for all the mentioned calculi.

Lemma 1 Supposing $\mathcal{S}$ is a solvable collection of attribute sets and $\emptyset \neq T \subset N$ the contraction $\mathcal{S} \wedge T$ is also solvable.

Proof: Supposing $\{k_L;\ L \in \mathcal{S} \wedge T\}$ is a pairwise compatible collection of knowledge representations we are to show that it is globally consistent. To this end we construct (and this step depends on a calculus) a pairwise compatible collection of knowledge representations $\{k'_S;\ S \in \mathcal{S}\}$ such that $\forall\, L \in \mathcal{S} \wedge T \quad [L = S \cap T$ for some $S \in \mathcal{S}]$ implies $(k'_S)^L = k_L$.

In the sequel we give the corresponding constructions for all studied calculi.

$\boxed{\text{a}}$ probability measures

Having a pairwise compatible collection of probability distributions $\{P_L;\, L \in \mathcal{S} \wedge T\}$ select for each $i \in N \setminus T$ a probability distribution Q_i on X_i. Then put (for $x \in X_S$, $S \in \mathcal{S}$):

$P'_S(x) = (P_L)^{S \cap T}(x_{S \cap T}) \cdot \prod_{i \in S \setminus T} Q_i(x_i)$,

where $L \in \mathcal{S} \wedge T$ with $S \cap T \subset L$ is arbitrarily chosen (take $(P_L)^{\emptyset}(-) = 1 = \prod_{i \in \emptyset} Q_i(-)$).

$\boxed{\text{b}}$ database relations

Having pairwise compatible database relations $\{\mathsf{R}_L;\ L \in \mathcal{S} \wedge T\}$, for each $S \in \mathcal{S}$ with $S \cap T \neq \emptyset$ find $L \in \mathcal{S} \wedge T$ with $S \cap T \subset L$ and put $\mathsf{R}'_S = (\mathsf{R}_L)^{S \cap T} \times X_{S \setminus T}$ (if $S \cap T = \emptyset$ then $\mathsf{R}'_S = X_S$).

$\boxed{\text{c}}$ NCFs

Having pairwise compatible NCFs $\{\kappa_L;\ L \in \mathcal{S} \wedge T\}$, for each $S \in \mathcal{S}$ with $S \cap T \neq \emptyset$ find $L \in \mathcal{S} \wedge T$ with $S \cap T \subset L$ and put $\kappa'_S(x) = (\kappa_L)^{S \cap T}(x_{S \cap T})$ for $x \in X_S,\ S \in \mathcal{S}$ (where $(\kappa_L)^{\emptyset}(-) = 0$).

$\boxed{\text{d}}$ possibility measures

Having pairwise compatible possibility distributions $\{\,\pi_L;\ L \in \mathcal{S} \wedge T\,\}$, put:

$\pi'_S(x) = (\pi_L)^{S \cap T}(x_{S \cap T}) \qquad$ for $x \in X_S,\ S \in \mathcal{S}$

(L has the same meaning as in $\boxed{\text{c}}$, $(\pi_L)^{\emptyset}(-) = 1$).

$\boxed{\text{e}}$ BPAs

Supposing $\{\mathbf{m}_L,\ L \in \mathcal{S} \wedge T\}$ are pairwise compatible BPAs define for each $S \in \mathcal{S}$:
$\mathbf{m}_S{}'(F) = (\mathbf{m}_L)^{S \cap T}(E)$, whenever $F = E \times X_{S \setminus T}$ with $E \subset X_{S \cap T}$ (L has the same meaning as in preceding steps and $(\mathbf{m}_L)^{\emptyset}(-) = 1$),
and $\mathbf{m}_S{}'(F) = 0$ for remaining $F \subset X_S$. $\square$

Consequence 1 Supposing **[a]** a collection of attribute sets $\mathcal{S}$ is not solvable.

Proof: Put $T = \{z_i;\ i = 1, \ldots, n\}$ where we chose $z_i \in S_i \cap S_{i+1} \setminus \bigcup(\mathcal{S} \setminus \{S_i, S_{i+1}\})$. Then use Lemma 1 and Example 1 in section 3 to get the desired conclusion. $\square$

Lemma 2 Supposing **[b]** a collection of attribute sets $\mathcal{S}$ is not solvable.

Proof: Without loss of generality suppose $\bigcap \mathcal{S} = \emptyset$: otherwise put $T = N \setminus \bigcap \mathcal{S}$ and consider $\mathcal{S} \wedge T$ instead of $\mathcal{S}$. In all constructions below we put $X_i = \{0, 1\}$.

$\boxed{\text{a}}$ probability measures

Denote $m = card \bigcup \mathcal{S} - 1$ and define for each $S \in \mathcal{S}$ a probability distribution P_S ($x \in X_S$):

$$P_S(x) = \begin{cases} (m - card\ S) \cdot m^{-1} & \text{if } \sum_{i \in S} x_i = 0 \\ m^{-1} & \text{if } \sum_{i \in S} x_i = 1 \\ 0 & \text{if } \sum_{i \in S} x_i \geq 2. \end{cases}$$

It is no problem to verify pairwise compatibility of $\{P_S; S \in \mathcal{S}\}$. Now, suppose by contradiction that P is a probability distribution over N having $\{P_S; S \in \mathcal{S}\}$ as marginals. Then

(a) $P^{\cup\mathcal{S}}(x) = 0$ for $x \in X_{\cup\mathcal{S}}$, $\sum_{i \in \cup\mathcal{S}} x_i \geq 2$.

Indeed: for fixed $x \in X_{\cup\mathcal{S}}$ find $i, j \in \bigcup \mathcal{S}$ with $x_i = x_j = 1$ and then $S \in \mathcal{S}$ with $i, j \in S$. Hence $0 \leq P^{\cup\mathcal{S}}(x) \leq P^S(x_S) = 0$.

(b) $P^{\cup\mathcal{S}}(x) = m^{-1}$ for $x \in X_{\cup\mathcal{S}}$, $\sum_{i \in \cup\mathcal{S}} x_i = 1$.

Indeed: for fixed $x \in X_{\cup\mathcal{S}}$ take the only $i \in \bigcup \mathcal{S}$ with $x_i = 1$ and consider $S \in \mathcal{S}$ with $i \in S$. Then $m^{-1} = P^S(x_S) = P^{\cup\mathcal{S}}(x) + \sum \{ P^{\cup\mathcal{S}}(x_S, y);\ y \in X_{\cup\mathcal{S}\setminus S}\ \ \sum_{i \in \cup\mathcal{S}\setminus S} y_i \geq 1 \}$, where the latter sum is zero by (a).

Evidently (b) gives a contradictory conclusion $\sum \{ P^{\cup\mathcal{S}}(x); x \in X_{\cup\mathcal{S}} \} \geq (m+1) \cdot m^{-1} > 1$.

$\boxed{\text{b}}$ database relations

Let us put: $\mathsf{R}_S = \{x \in X_S;\ \sum_{i \in S} x_i = 1\}$ (for $S \in \mathcal{S}$). As $(\mathsf{R}_S)^V = \{x \in X_V;\ \sum_{i \in V} x_i \leq 1\}$ for every proper subset $V \subset S$, $\{\mathsf{R}_S;\ S \in \mathcal{S}\}$ are pairwise compatible. Nevertheless, no database relation R over N has $\{\mathsf{R}_S; S \in \mathcal{S}\}$ as marginals:

(a) whenever $x \in X_N$ with $\sum_{i \in \cup\mathcal{S}} x_i \geq 2$ then $x \notin \mathsf{R}$

Indeed: choose $i, j \in \bigcup \mathcal{S}$ with $x_i = x_j = 1$ and $S \in \mathcal{S}$ with $i, j \in S$. Then $x_S \notin \mathsf{R}_S$ implies $x \notin \mathsf{R}$.

(b) whenever $x \in X_N$ with $\sum_{i \in \cup\mathcal{S}} x_i \leq 1$ then $x \notin \mathsf{R}$

Indeed: find $S \in \mathcal{S}$ with $\sum_{j \in S} x_j = 0$ (fix contingent $i \in \bigcup \mathcal{S}$ with $x_i = 1$ and by $\bigcap \mathcal{S} = \emptyset$ find $S \in \mathcal{S}$ with $i \notin S$). Evidently $x_S \notin \mathsf{R}_S$ implies $x \notin \mathsf{R}$.

$\boxed{\text{c}}$ NCFs

Define an NCF κ_S over N for each $S \in \mathcal{S}$ as follows:

$$\kappa_S(x) = \begin{cases} 0 \text{ if } \sum_{i \in S} x_i = 1 \\ 1 \text{ otherwise} \end{cases} \qquad \text{for } x \in X_S.$$

The collection $\{\kappa_S;\ S \in \mathcal{S}\}$ is pairwise compatible as for $\emptyset \neq V \subset S \in \mathcal{S}$, $V \neq S$ it holds:

$$(\kappa_S)^V(y) = \begin{cases} 0 \text{ if } \sum_{i \in V} y_i \leq 1 \\ 1 \text{ otherwise} \end{cases} \qquad \text{for } y \in X_V.$$

To disprove global consistency use the criterion mentioned below Proposition 1 and compute $\kappa^*(x) = \max_{S \in \mathcal{S}}\ \kappa_S(x_S)$ (for $x \in X_N$). Supposing $\kappa^*(x) = 0$ we get $\forall\, S \in \mathcal{S}\ \ \kappa_S(x_S) = 0$ i.e. $x_S \in \mathsf{R}_S$ where R_S is from $\boxed{\text{b}}$ - but it was shown there that no $x \in X_N$ satisfies this requirement. Therefore $\kappa^* \equiv 1$.

$\boxed{\text{d}}$ possibility measures

One can use for example the collection $\{\pi_{\mathsf{R}_S};\ S \in \mathcal{S}\}$ where $\{\mathsf{R}_S;\ S \in \mathcal{S}\}$ are database relations from the item $\boxed{\text{b}}$ (see the reasoning before Example 2).

$\boxed{\text{e}}$ BPAs

One can use $\{\mathbf{m}_{P_S};\ S \in \mathcal{S}\}$ where probability measures $\{P_S;\ S \in \mathcal{S}\}$ are from $\boxed{\text{a}}$. $\square$

Lemma 3 A collection $\{I, J\}$ where $I, J \subset N$ is solvable within all mentioned calculi.

Proof: The constructions depend on calculi.

$\boxed{\text{a}}$ probability measures

Having $\{P_I, P_J\}$ a compatible collection of probability distributions we put:
$P(x) = 0$ if $(P_I)^{I\cap J}(x_{I\cap J}) = 0$, and
$P(x) = P_I(x_I)\cdot P_J(x_J)\cdot[(P_I)^{I\cap J}(x_{I\cap J})]^{-1}\cdot\prod_{i\in N\setminus I\cup J} Q_i(x_i)$ otherwise (where $(P_I)^{\emptyset}(-) = 1$ and Q_i are arbitrarily chosen one-dimensional probability measures).

$\boxed{\text{b}}$ database relations

Having $\{\mathsf{R}_I, \mathsf{R}_J\}$ a compatible collection of database relations put $\mathsf{R} = (\mathsf{R}_I \times X_{N\setminus I}) \cap (\mathsf{R}_J \times X_{N\setminus J})$.

$\boxed{\text{c}}$ NCFs

Having a couple of compatible NCFs $\{\kappa_I, \kappa_J\}$ put: $\kappa(x) = \max\{\kappa_I(x_I), \kappa_J(x_J)\}$ (for $x \in X_N$).

$\boxed{\text{d}}$ possibility measures

Having compatible possibility distributions $\{\pi_I, \pi_J\}$ use the formula from Proposition 1:
$\pi_*(x) = \min\{\pi_I(x_I),\ \pi_J(x_J)\}$ (for $x \in X_N$).

$\boxed{\text{e}}$ BPAs

Having a compatible collection of BPAs $\{\mathbf{m}_I, \mathbf{m}_J\}$ we can define a BPA $\mathbf{m}$ over N having them as marginals as follows: focal elements of $\mathbf{m}$ will have the form $\mathsf{G} = (\mathsf{E} \times X_{N\setminus I}) \cap (\mathsf{F} \times X_{N\setminus J})$ where $\mathsf{E} \subset X_I$ is a focal element of $\mathbf{m}_I$, $\mathsf{F} \subset X_J$ is a focal element of $\mathbf{m}_J$ and $\mathsf{E}^{I\cap J} = \mathsf{F}^{I\cap J}$ (in case $I \cap J = \emptyset$ automatically $\mathsf{E}^{\emptyset} = \mathsf{F}^{\emptyset}$). Put $\mathbf{m}(\mathsf{G}) = \mathbf{m}_I(\mathsf{E}) \cdot \mathbf{m}_J(\mathsf{F}) \cdot [(\mathbf{m}_I)^{I\cap J}(\mathsf{E}^{I\cap J})]^{-1}$. □

Hence, one can conclude using [16]:

Proposition 2 A collection of (nonempty) attribute sets $\mathcal{S}$ is solvable within any of the mentioned calculi iff it satisfies the *running intersection property*:
there exists an ordering $S_1, \ldots, S_n$ of $\mathcal{S}$ such that
$\forall j \geq 2 \quad \exists i \ \ 1 \leq i < j \quad S_j \cap (\bigcup_{k<j} S_k) \subset S_i$.

5 Conclusion

The results proved in this paper have mainly theoretical significance. The study of the marginal problem was so far limited to probability measures and dabase relations (resp. to NCFs). How, the horizons in this respect were broadened also to possibility measures and to the calculus of Dempster-Shafer theory.

The reader can object that the study of "ideal" consistency of input knowledge may be unrealistic, but I think it is useful to be aware of these results. For example, the results concerning the running intersection property highlight the significance of *decomposable models* [11] which correspond uniquely to collections satisfying this condition. If one is interested in the "internal coherence" of his (her) procedures (i.e. whether the "input" pieces of knowledge and the "output" knowledge are coherent) one should take advantage of these models no matter which calculus one decided on to represent knowledge.

I hope that the method of testing global consistency for possibility measures is of some benefit, as well.

References

1. Beeri, C., Fagin, R., Maier, D., Yannakis, M.: On the desirability of acyclic database schemes. J. Assoc. Comp. Mach. **30** (1983) 479–513.
2. Csiszár, I.: I-divergence geometry of probability distributions and minimization problems. Ann. Prob. **3** (1975) 146–158.
3. Deming, W.E., Stephan, F.F.: On a least square adjustment of a sampled frequency table when expected marginal totals are known. Ann. Math. Statis. **11** (1940) 427–444.
4. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Statis. **38** (1967) 325–339.
5. Dubois, D., Prade, H.: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press, New York - London (1988).
6. Dubois, D., Lang, J., Prade, H.: Possibilistic logic. report IRIT/97–98/R, Institut de Recherche en Informatique de Toulouse (1991).
7. Hunter, D.: Graphoids and natural conditional functions. Int. J. Approx. Reasoning **5** (1991) 485–504.
8. Jiroušek, R.: Solution of the marginal problem and decomposable distributions. Kybernetika **27** (1991) 403–412.
9. Kellerer, H.G.: Verteilungsfunktionen mit gegebenem Marginalverteilungen (in German). Z. Wahrsch. Verw. Gebiete **3** (1964) 247–270.
10. Malvestuto, F.M.: Existence of extensions and product extensions for discrete probability distributions. Discr. Math. **69** (1988) 61–77.
11. Pearl, J.: Probabilistic Reasoning in Intelligent Systems. Morgan Kaufman, San Mateo CA (1988).
12. Perez, A., Jiroušek, R.: Constructing an intensional expert system INES. In: J.H. von Bemmel, F. Grmy, J. Zvrov (eds.): Medical Decision Making: Diagnostic Strategies and Expert Systems. North-Holland, Amsterdam (1985), 307–315.
13. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton - London (1976).
14. Shenoy, P.P.: On Spohn's rule for revision of beliefs. Int. J. Approx. Reasoning **5** (1991) 149–181.
15. Spohn, W.: Ordinal conditional functions: a dynamic theory of epistemic states. In: W.L. Harper, B. Skyrms (eds.): Causation in Decision, Belief Change, and Statistics vol. II. Kluwer, Dordrecht (1988), 105–134.
16. Studený, M.: Conditional independence and natural conditional functions. Int. J. Approx. Reasoning (to appear).
17. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Set and Systems **1** (1978) 3–28.
18. Zadeh, L.A.: A theory of approximate reasoning. In: (J.E. Hayes, D. Michie, L.I. Mikulich (eds.): Machine Intelligence 9. Ellis Horwood, Chichester (1979), 149–194.

This article was processed using the LaTeX macro package with LLNCS style