International Journal of Approximate Reasoning 51 (2010) 832–845

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Measures of inconsistency and defaults

Dragan Doder [a,*], Miodrag Rašković [b], Zoran Marković [b], Zoran Ognjanović [b]

[a] Faculty of Mechanical Engineering, University of Belgrade, Kraljice Marije 16, 11120 Belgrade, Serbia
[b] Mathematical Institute of the Serbian Academy of Sciences and Arts, Kneza Mihaila 36, 11000 Belgrade, Serbia

ARTICLE INFO

Article history:
Received 21 July 2009
Received in revised form 11 May 2010
Accepted 27 May 2010
Available online 31 May 2010

Keywords:
n-Consistency
n-Probability
Probability measure
Approximate probability
Default reasoning

ABSTRACT

We introduce a method for measuring inconsistency based on the number of formulas needed for deriving a contradiction. The relationships to previously considered methods based on probability measures are discussed. Those methods are extended to conditional probability and default reasoning.

1. Introduction

The need for handling inconsistent knowledge bases has been recognized by the Artificial Intelligence community in recent years. Although it is the overruling opinion that inconsistencies are undesirable, they still appear in practice, mostly for two reasons: unreliable sources of information and presence of rules with exceptions.

Even if every set of formulas T_i which represents the ith source of information (or expert opinion) is consistent it is possible that $\bigcup T_i$ is inconsistent – a typical example is a working schedule list of demands which overlap. Hunter and Konieczny [1] also points to the following examples: a group of clinicians advising on a patient, a group of witnesses of some incident and a set of newspaper reports covering some event.

Another common reason that leads to inconsistencies is the presence of default rules (rules with possible exceptions, for example "birds fly") in knowledge bases. If a knowledge base consists of a consistent set of defaults ("quakers are pacifists", "republicans are not pacifists") formalized as entailments, and a consistent set of facts ("Nixon is a quaker", "Nixon is a republican"), it is still possible that the whole base is inconsistent ("Nixon is a pacifist", "Nixon is not a pacifist").

Contrary to common opinion, some authors claim that inconsistencies may be useful – Gabbay and Hunter [2,3] give an example of overbooking in airline booking systems. Many logical formalisms are developed for reasoning under inconsistency, like paraconsistent logics, default reasoning, possibility theory, belief revision and formal argumentation (see, for example [4–12]). Unlike classical logic, they enable inferring non-trivial conclusions from inconsistent knowledge bases, so that two different inconsistent sets can lead to different sets of conclusions.

Development of those techniques points out to the need for analyzing and comparing inconsistent sets. As it is pointed out in [13], defining degree of inconsistency turns out to be important in software specifications, databases, decision support, ontologies and information merging.

* Corresponding author. Tel.: +381 64 2970369; fax: +381 11 2186105.
E-mail addresses: ddoder@mas.bg.ac.rs (D. Doder), miodragr@mi.sanu.ac.rs (M. Rašković), zoranm@mi.sanu.ac.rs (Z. Marković), zorano@mi.sanu.ac.rs (Z. Ognjanović).

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2010.05.007

In several papers, the measure of inconsistency depends on the proportion of the language that is affected by the inconsistency in a theory ([14–16]). The limitation of this approach (except in [15]) is the fact that it doesn't consider the distribution of contradiction among the formulas.

The second approach considers the number of formulas needed to produce a contradiction. According to Sorensen ([17]), an inconsistent set A is better than an inconsistent set B, if the shortest derivation of a contradiction requires more members of A than B. This idea implies that the set of formulas $\{\phi_1,\ldots,\phi_n\}$ is not equivalent to the singleton $\{\phi_1\wedge\cdots\wedge\phi_n\}$. This property is valid in non-adjunctive logics (see [18,19]), special class of paraconsistent logics, and it can be meaningful in some practical applications.

For instance, we often have to deal with an inconsistent database. If we have to choose between several inconsistent databases, this way of measuring can be appropriate for answering the question which one is the least inconsistent. A more tolerable database is the one with a larger number of entries required for a contradiction. Similarly, in a court proceeding, it makes the difference whether each two testimonies are contradictory, or only several of them taken together produce a contradiction.

Our idea was to study this simple syntactic measure of inconsistency and to see how it compares with a more semantic notion of existence of a probability measure which assigns a high probability to each formula of the theory. In particular, we defined a theory T to be n-consistent if each n-element subset of T is was consistent, and wanted to connect this number n with the existence of a probability higher than $1-\frac{1}{n}$ for every formula of T, in which case we called T n-probable. However we soon discovered that Knight [20] had already investigated extensively this semantic notion based on the existence of appropriate probability, but he called it η-consistency. Our syntactic notion also appears in one of Knight's theorems ([20, Theorem 3.5]) but he does not have a name for it as he is concerned mainly with semantic side. Since our aim was to study in parallel both the syntactic and semantic approach and to try to find precise relationship between them, we needed two expressions. We choose to keep our definitions because consistency is a syntactical notion based on the non-existence of a formal proof of contradiction, so n-consistency seems more appropriate to denote the non-existence of a proof of contradiction from n hypotheses. For the semantic notion we decided to use the expression "n-probable". Even though this might confuse a reader familiar with Knight's work, we though that calling the semantic notion "n-consistency" and inventing some new name for the syntactic notion might be even more confusing. More precise comparison of our results to Knight's is given in Remark 3.1. Our approach differs significantly from the approach in [14,16] in the sense that their measures of inconsistency do not differentiate between $\{\phi_1,\ldots,\phi_n\}$ and $\{\phi_1\wedge\cdots\wedge\phi_n\}$ so singletons containing different contradictions may have different measures.

Section 2 introduces the syntactic notions of n-consistency and strict n-consistency, and establishes some basic facts. Section 3 introduces the semantical notion of n-probability (there exists a probability measure assigning to all formulas probability greater then $1-\frac{1}{n}$). It is demonstrated that this property is stronger than n-consistency, so a weaker property of local n-probability is introduced, which is shown to be equivalent to n-consistency. Relations to Knight's results are discussed and a generic example is given, which demonstrates that the obtained equivalence is the best possible. Section 4 uses conditional probability to introduce the notion of n-probability modulo some "certain knowledge". Connections to n-probability and n-consistency are proved and two examples are given indicating possible applications. A syntactic notion of n-consistency modulo a set of formulas is also introduced. In Section 5 we modify the approach from Section 4 by using non-standard probability measure. We define the notions of strong n-probability and strong probability for a non-standard measure μ which are preserved under μ-consequence, defined as conditional probability infinitesimally close to 1. Finally, in Section 6, two applications of results of Sections 4 and 5 to defaults are given. First, using the connection between defaults and non-standard conditional probability infinitesimally close to 1 established in [9], we show that strong n-consistency (syntactic analogue of strong n-probability introduced in Section 5) is preserved under default derivation. Second application is quite opposite in spirit. Using the results of Section 4 we define new relations which may be regarded as a finite approximation of default consequence. It is shown that these relations satisfy a weakened version of the Kraus et al.'s system of rules for defaults P ([8]). We also consider a system which combines these new relations with the standard defaults, similar to the system of Lukasiewicz [25].

The main contributions of this paper can be summarized as follows:

- We introduced new syntactic and semantic notions (one of them is very similar to Knight's) for measuring inconsistency, and investigated relations between them. (Sections 2 and 3).
- We generalized the introduced notions to measure inconsistency of theories resulting from beliefs of different agents. (Sections 4 and 5).
- We investigated relations between our results and default reasoning. (Section 6).

2. *n*-Consistency

Let $\mathcal{P}$ denotes a nonempty set of propositional letters. In this section, we assume that $\mathcal{P}$ is finite, while this restriction will not be imposed in the rest of the paper.

Definition 2.1. Let $\mathcal{P}=\{p_1,\ldots,p_m\}$. An atom is any formula of the form

$$\pm p_1 \wedge \cdots \wedge \pm p_m,$$

where $+p$ is p and $-p$ is $\neg p$.

The term "atom" comes from Boolean algebras where it denotes an element whose intersection with any other element gives 0 or itself. In the Lindenbaum algebra of formulas over $\mathcal{P}$, considered as a Boolean algebra, formulas of the form $\pm p_1 \wedge \cdots \wedge \pm p_m$ are exactly the atoms. For each $\phi \in For_{\mathcal{P}}$ there are atoms $\alpha_1, \ldots, \alpha_k$ such that ϕ is equivalent with $\alpha_1 \vee \cdots \vee \alpha_k$, which is called the disjunctive normal form of ϕ.

Definition 2.2. A theory T is n-consistent if each $T' \subseteq T$ of cardinality n is consistent.

For instance, each consistent theory T is also n-consistent, for all n. On the other hand, there are many examples of inconsistent theories that are n-consistent for some, sometimes quite large, n.

Example 2.1 (*Lottery Paradox*). The standard example of this kind is the so called lottery paradox with n players. If we denote by ϕ_i the claim "ith ticket will win the lottery", it is common sense to believe in $\neg\phi_i$, for all $i \leqslant n$. However, the fact is that some ticket will win the lottery. Putting these beliefs together, we get the following inconsistent, but n-consistent theory $\{\neg\phi_1, \ldots, \neg\phi_n, \phi_1 \vee \cdots \vee \phi_n\}$.

Suppose that $\mathcal{P} = \{p_1, \ldots, p_m\}$. Then there are 2^m atoms and 2^{2^m} different equivalence classes of formulas (elements of the corresponding Lindenbaum algebra). Thus, we may assume that there are only 2^{2^m} different formulas. Further on we will assume that sets of formulas we consider consist of representatives of equivalence classes, i.e. that they contain no equivalent formulas. As representatives of equivalence classes we take the formulas in disjunctive normal form.

Note that if all the formulas in a set of formulas have a common atom, this atom defines a valuation which satisfies the whole set, i.e., demonstrates its consistency. Thus, maximal consistent theories over the finite language are sets of formulas containing a common atom (in disjunctions). Consequently, they are of cardinality 2^{2^m-1}.

For example, if $\mathcal{P} = \{p, q\}$, fixing the atom $p \wedge q$ we obtain maximal consistent theory $T = \{p \wedge q, (p \wedge q) \vee (p \wedge \neg q), (p \wedge q) \vee (\neg p \wedge q), (p \wedge q) \vee (\neg p \wedge \neg q), (p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge q), (p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge \neg q), (p \wedge q) \vee (\neg p \wedge q) \vee (\neg p \wedge \neg q), (p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge q) \vee (\neg p \wedge \neg q)\}$ of cardinality 8. On the other hand, adding any formula ϕ to the theory T makes it highly inconsistent: it contains both ϕ and $\neg\phi$ (for some ϕ), so it is not even 2-consistent.

That leads to the following question: if the theory is inconsistent, but n-consistent, how large n can be? In order to answer the question, we introduce a more precise notion of "strict n-consistency" which should enable us to better classify inconsistent theories.

Definition 2.3. A theory T is strictly n-consistent if it is n-consistent and it is not $(n+1)$-consistent.

Note that,unlike n-consistent theories, strictly n-consistent theories are always inconsistent. Observe that the theory from Example 2.1 is strictly n-consistent. The main advantage of the new notion is that a theory may be strictly n-consistent for only one n, while it will be k-consistent for every $k \leqslant n$.

Although we saw that there are consistent theories of cardinality 2^{2^m-1}, the next theorem says that there are no strictly n-consistent theories for large n.

Theorem 2.1. *Let $\mathcal{P} = \{p_1, \ldots, p_m\}$. Then, there exists a strictly n-consistent theory $T \subseteq For_{\mathcal{P}}$ iff $n < 2^m$.*

Proof. The set S of negations of all atoms (i.e. set of complements of atoms in Lindenbaum algebra) forms a $(2^m - 1)$-consistent theory that is not 2^m-consistent. Indeed, a negation of an atom is equivalent to the disjunction of the remaining $2^m - 1$ atoms, and so, a conjunction of any k of formulas from the set S is equivalent to a disjunction of $2^m - k$ atoms. Especially, a conjunction of the length $2^m - 1$ is equivalent to some atom, and a conjunction of the length 2^m is a contradiction. Similarly, for $k < 2^m$, the set formed by the negations of any k atoms and the conjunction of negations of the remaining $2^m - k$ atoms will be a strictly k-consistent theory.

Conversely, suppose that $T = \{\phi_1, \phi_2, \ldots, \phi_{2^m}, \ldots\}$ is strictly n-consistent, for some $n \geqslant 2^m$. Without loss of generality, we may assume that $\{\phi_1, \ldots, \phi_{n+1}\}$ is a minimal inconsistent subset of T, and that all formulas in T are disjunctions of atoms. Obviously, ϕ_1 is not a tautology, so ϕ_1 contains at most $2^m - 1$ atoms.

We claim that there exists an atom in ϕ_1 that does not appear in ϕ_2 and vice versa, that there exists an atom in ϕ_2 that does not appear in ϕ_1. Otherwise, ϕ_1 implies ϕ_2 or ϕ_2 implies ϕ_1, so $\{\phi_1, \ldots, \phi_{n+1}\}$ is not minimal inconsistent. Therefore, ϕ_1 and ϕ_2 contain together at most $2^m - 2$ atoms.

Similarly, for any $k < 2^m$, there is an atom that appears in $\phi_1 \wedge \cdots \wedge \phi_k$, but not in ϕ_{k+1}, and an atom that appears in ϕ_{k+1} but not in $\phi_1 \wedge \cdots \wedge \phi_k$, therefore $\phi_1, \ldots, \phi_{k+1}$ contain together at most $2^m - (k+1)$ atoms. It follows that there is no common atom in the formulas $\phi_1, \ldots, \phi_{2^m}$ and that $\{\phi_1, \ldots, \phi_{2^m}\}$ is inconsistent, which contradicts the assumption. □

This theorem shows that strict n-consistency is a sharper tool for classifying inconsistent theories, because strictly n-consistent theories exist only for a quite limited number of n's.

3. *n*-Probability

Recall that $\mathcal{P}$ denotes a nonempty set of propositional letters. From now on, $\mathcal{P}$ can be finite or countable. By $For_{\mathcal{P}}$ we will denote the set of all propositional formulas over $\mathcal{P}$. A theory is any subset of $For_{\mathcal{P}}$.

Definition 3.1. A function $\mu : For_{\mathcal{P}} \to [0,1]$ which satisfies

1. $\mu(\phi) = 1$, whenever ϕ is a tautology;
2. $\mu(\phi) = \mu(\psi)$, whenever $\phi \leftrightarrow \psi$ is a tautology;
3. $\mu(\phi \vee \psi) = \mu(\phi) + \mu(\psi)$, whenever $\phi \wedge \psi$ is a contradiction;

will be called a probability measure. A probability measure μ is said to be neat if $\mu(\phi) = 0$ implies that ϕ is a contradiction.

Lemma 3.1. *For all $\phi, \psi \in For_{\mathcal{P}}$ and for any probability measure μ, the following holds:*

1. $\mu(\neg\phi) = 1 - \mu(\phi)$.
2. $\mu(\phi) \leqslant \mu(\psi)$, whenever $\phi \to \psi$ is a tautology;
3. $\mu(\phi \vee \psi) = \mu(\phi) + \mu(\psi) - \mu(\phi \wedge \psi)$.

As a consequence of Lemma 3.1, we have that equivalent formulas have the same measure. By induction on *n*, one can easily show that

$$\mu(\phi_1 \wedge \cdots \wedge \phi_n) \geqslant \mu(\phi_1) + \cdots + \mu(\phi_n) - (n-1). \quad (1)$$

If μ is any probability measure and ϕ is a formula equivalent to the disjunction of atoms $\alpha_1 \vee \cdots \vee \alpha_k$, then

$$\mu(\phi) = \mu(\alpha_1) + \cdots + \mu(\alpha_k).$$

Definition 3.2. A theory $T \subseteq For_{\mathcal{P}}$ is *n*-probable for the probability measure $\mu : For_{\mathcal{P}} \to [0,1]$ if

$$\mu(\phi) > 1 - \frac{1}{n}$$

for all $\phi \in T$.

A theory *T* is *n*-probable if there exists a probability measure μ such that *T* is *n*-probable for μ.

Lemma 3.2. *Each n-probable theory is n-consistent.*

Proof. Suppose that *T* is *n*-probable. If *T* is not *n*-consistent, then there is an inconsistent subset $\{\phi_1, \ldots, \phi_n\}$ of *T*, so

$$\mu(\phi_1 \wedge \cdots \wedge \phi_n) = 0,$$

for any probability measure μ. It follows that

$$\mu(\neg\phi_1) + \cdots + \mu(\neg\phi_n) \geqslant \mu(\neg\phi_1 \vee \cdots \vee \neg\phi_n) = 1.$$

On the other hand, *T* is *n*-probable, so $\mu(\neg\phi_i) < \frac{1}{n}$, for some μ; a contradiction. □

The following example shows that the converse implication in Lemma 3.2 need not be true.

Example 3.1. We will define a strictly *n*-consistent theory of cardinality $n+2$, that is not *n*-probable. To do that, we will construct the formulas of the theory in such a way that any *n* formulas of the theory have a common atom (of some finite language).

For that purpose, let *n* be an arbitrary positive integer and let $\mathcal{P}'$ be an arbitrary subset of $\mathcal{P}$ of cardinality *m* such that $2^m > \frac{(n+2)(n+1)}{2}$. Furthermore, let $\{\alpha_{ij} \mid 1 \leqslant i < j \leqslant n+2\}$ be some set of atoms over $\mathcal{P}'$ and let

$$T = \{\phi_1, \ldots, \phi_{n+2}\},$$

where ϕ_k is the formula $\bigvee_{k \neq i,j} \alpha_{ij}$ (in other words, α_{ij} appears as a disjunct of every formula of the theory *T* except *i*th and *j*th). *T* is strictly *n*-consistent, since any *n* formulas of *T* have a common atom which defines a model for this set of formulas, and there is no common atom for any $n+1$ formulas of *T*. (For illustration, if $n = 3$ we have $\phi_1 = \alpha_{23} \vee \alpha_{24} \vee \alpha_{25} \vee \alpha_{34} \vee \alpha_{35} \vee \alpha_{45}$, $\phi_2 = \alpha_{13} \vee \alpha_{14} \vee \alpha_{15} \vee \alpha_{34} \vee \alpha_{35} \vee \alpha_{45}$, $\phi_3 = \alpha_{12} \vee \alpha_{14} \vee \alpha_{15} \vee \alpha_{24} \vee \alpha_{25} \vee \alpha_{45}$, $\phi_4 = \alpha_{12} \vee \alpha_{13} \vee \alpha_{15} \vee \alpha_{23} \vee \alpha_{25} \vee \alpha_{35}$, $\phi_5 = \alpha_{12} \vee \alpha_{13} \vee \alpha_{14} \vee \alpha_{23} \vee \alpha_{24} vee \alpha_{34}$. For example, the atom α_{45} secures the consistency of the set $\{\phi_1, \phi_2, \phi_3\}$.)

Suppose that $\{\phi_1, \ldots, \phi_{n+1}\}$ is *n*-probable for a probability measure μ. We will show that the number

$$s = \mu(\phi_{n+2}) = \sum_{j \neq n+2} \mu(\alpha_{ij})$$

has to be very small. Without loss of generality, we may assume that

$$\sum_{1 \leqslant i < j \leqslant n+2} \mu(\alpha_{ij}) = 1.$$

Since, by assumption, $\mu(\phi_i) > 1 - \frac{1}{n},\ i \in \{1, \dots, n+1\}$, it follows that

$$\sum_{i=1}^{n+1} \mu(\phi_i) > (n+1)\left(1 - \frac{1}{n}\right),$$

so

$$n(1-s) + (n-1)s > (n+1)\left(1 - \frac{1}{n}\right).$$

Consequently,

$$s < n - (n+1)\left(1 - \frac{1}{n}\right) = \frac{1}{n} \tag{2}$$

and, therefore, T is not n-probable (for $n > 2$).

Since the notion of n-consistency is strictly weaker then the notion of n-probability, we define a new notion that turns out to be the probabilistic analogue of n-consistency.

Definition 3.3. A theory $T \subseteq For_{\mathcal{P}}$ is locally n-probable if each subset of T of cardinality $n + 1$ is n-probable.

It is easy to see that each n-probable theory is also locally n-probable. Obviously, the converse is not true.

Theorem 3.3. *A theory T is locally n-probable if and only if it is n-consistent.*

Proof. Suppose that T is n-consistent. Let $T_0 = \{\phi_1, \dots, \phi_{n+1}\}$ be an arbitrary finite subset of T of the cardinality $n + 1$. Let $\mathcal{P}_{T_0} = \{p_1, p_2, \dots, p_m\}$ be the set of all propositional letters from T_0, and let $\mathcal{P}'_{T_0} = \{r_1, r_2, \dots, r_n\}$ be another set of propositional letters such that $\mathcal{P}_{T_0} \cap \mathcal{P}'_{T_0} = \emptyset$. Now, we consider the set of atoms over $\mathcal{P}_{T_0} \cup \mathcal{P}'_{T_0}$. Since T is n-consistent, the formulas $\phi_2, \dots, \phi_{n+1}$ will contain at least one atom in common. We choose one of then and denote it by α_1. In a similar way we define the atom α_j for the set $\{\phi_1, \dots, \phi_{n+1}\} \backslash \{\phi_j\}$ so that $\alpha_j \neq \alpha_i$, for every $i < j$. Note that we can always find such an atom since we consider the set of atoms over $\mathcal{P}_{T_0} \cup \mathcal{P}'_{T_0}$.

Let $\mu : For_{\mathcal{P}} \to [0, 1]$ be any probability measure such that:

- $\mu(\alpha_i) = \frac{1}{n+1}, i \in \{1, \dots, n+1\}$.
- $\mu(\alpha_i) = 0, i \in I \backslash \{1, \dots, n+1\}$.

Then, $\mu(\phi_j) \geqslant \frac{n}{n+1}$, for $j \in \{1, \dots, n+1\}$. Note that, if $\{\phi_1, \dots, \phi_{n+1}\}$ is inconsistent, then $\mu(\phi_j) = \frac{n}{n+1}$, for each j.

The proof of the converse implication is identical to the proof of Lemma 3.2. □

Corollary 3.4. *A theory is strictly n-consistent iff it is locally n-probable and not locally (n + 1)-probable.*

Theorem 3.5. *Let T be a strictly n-consistent theory of cardinality n + 1. Then there exists a probability measure* μ *such that for every* $\phi \in T, \mu(\phi) \geqslant 1 - \frac{1}{n+1}$, *but T is not (n + 1)-probable.*

Proof. This theorem is an immediate consequence of the proof of Theorem 3.3. Note that, since T is strictly n-consistent, it is not $(n + 1)$-consistent, so it is not $(n + 1)$-probable (by Lemma 3.2). □

Remark 3.1. Theorem 3.5 is actually identical to Knight's theorem [20, Theorem 3.5], only translated in our terminology. It is the closest contact between our results and Knight's. We stated it primarily in order to clarify relations between this paper and Knight's [20]. As noted in the Introduction, the main difference between our approach and Knight's is that his principal concern is semantic while our idea was to study the interplay between syntax and semantics. Our notion of "n-probable" is almost like a special case of his "η-consistency" (for $\eta = 1 - \frac{1}{n}$) except that in our case the probability is $> 1 - \frac{1}{n}$ and in his it is $\geqslant \eta (\eta \in [0,1])$. However, our interest is only in high probability, so $1 - \frac{1}{n}$ is good enough, while providing a nice symmetry between syntax (n-consistent) and semantics (n-probable). Also, our first concern was to find a semantic analogue of

n-consistency and for that purpose even our n-probability is too strong. As Theorem 3.3 shows, the appropriate notion is the even weaker "local n-probability".

A natural question concerning Definition 3.3 would be whether the size of the subsets ($n + 1$) is necessarily connected with the restriction on the measure $\left(\mu(\phi) > 1 - \frac{1}{n}\right)$, i.e., do we need the same number n in both places. Example 3.2 shows that if we want to preserve the equivalence in Theorem 3.3, the size of subsets cannot be increased.

Example 3.2. Let $T = \{\phi_1, \ldots, \phi_{n+2}\}$ be the theory defined in Example 3.1. We will compute the maximal $r \in [0,1]$ such that there exists a probability measure μ which satisfies

$$\mu(\phi_i) \geqslant r, \quad i \in \{1, \ldots, n+2\}.$$

In the same way as we derived the inequality (2) we can conclude

$$s \leqslant n - (n+1)r. \tag{3}$$

On the other hand, since

$$\mu(\phi_{n+2}) \geqslant r,$$

we obtain

$$s \geqslant r. \tag{4}$$

From (3) and (4) we can conclude that

$$r \leqslant \frac{n}{n+2}.$$

The previous example illuminates the following Knight's theorem [20, Corollary 4.12] which we reformulate according to our terminology.

Theorem 3.6. *Suppose that T is a n-consistent theory of cardinality m. Then, there exists a probability measure μ such that $\mu(\phi) \geqslant \frac{n}{m}$, for all $\phi \in T$.*

Example 3.2 shows that Theorem 3.6 cannot be strengthened in the sense that $\frac{n}{m}$ is maximal, i.e. there is no $r > \frac{n}{m}$ such that there exists a probability measure μ such that $\mu(\phi) \geqslant r$ for all $\phi \in T$.

Finally, Example 3.2 demonstrates also that n-consistency and local n-probability are in better accord with the Sorensen's criterion: "as the number of beliefs required to obtain the inconsistency increases, the inconsistency becomes more tolerable". Namely, theory T from Example 3.1 is n-consistent but not even $(n-1)$-probable. On the other hand, it would be easy to construct a theory T' which would be strictly $(n-1)$-consistent and also $(n-1)$-probable (it is enough to take strictly $(n-1)$-consistent theory T' of cardinality n, so local $(n-1)$-probability becomes $(n-1)$-probability). In accordance with Sorensen's criterion, it would seems that theory T is better than T', since in T' we can derive a contradiction from n-formulas. On the other hand, there is a probability measure which assigns higher probabilities to formulas of T' than any probability for T. Therefore, n-consistency might be considered as a more precise tool for measuring inconsistency than n-probability.

4. Conditional *n*-probability and *n*-consistency

If we have some body of knowledge which might be inconsistent it makes sense to split this body into two parts: one part would consist of some "certain facts" believed to be certainly true and the second part would consist of some statements believed to be probable, on the basis of the "certain facts". Returning to the Lottery Paradox, we may say that, due to the rules of the game, it is a certain fact that some ticket will win ($\phi_1 \vee \cdots \vee \phi_n$), while, on the basis of this fact, we believe that $\neg\phi_i$ is probable. Thus, assuming $\phi_1 \vee \cdots \vee \phi_n$, the theory $T = \{\neg\phi_1, \ldots, \neg\phi_n\}$ would be strictly $(n-1)$-consistent and $(n-1)$-probable. This motivates us to introduce a notion of "n-probable modulo ϕ". We will use the results from the previous section and the notion of conditional probability.

Let $\phi \in For_{\mathcal{P}}$ and let μ be a probability measure on $For_{\mathcal{P}}$ such that $\mu(\phi) > 0$. Then, a conditional probability measure is defined by

$$\mu(\psi|\phi) = \frac{\mu(\psi \wedge \phi)}{\mu(\phi)}.$$

It is well known that a function $\mu_\phi : For_{\mathcal{P}} \to [0,1]$, defined by $\mu_\phi(\psi) = \mu(\psi|\phi)$ is a probability measure.

Example 4.1 (*Lottery Paradox – revisited*). Returning to the lottery paradox, it seems reasonable to grant special status to the fact that someone will win the lottery and consider this as "certain knowledge". Then we may ask how many elements of the set $T = \{\neg p_1, \ldots, \neg p_n\}$ will be consistent with the fact $p_1 \vee \cdots \vee p_n$.

A natural probability measure μ (assuming that the lottery is fair) should satisfy: $\mu(\neg p_i) = 1 - \frac{1}{n}$ for every $i \in \{1, \ldots, n\}$ and, of course, $\mu(p_1 \vee \cdots \vee p_n) = 1$. This would make T $(n-1)$-probable "assuming" $p_1 \vee \cdots \vee p_n$. If some other probability measure would satisfy $\mu(\neg p_i) > 1 - \frac{1}{n}$ for $i \in \{1, \ldots, n\}$, this would imply $\mu(\neg p_1 \wedge \cdots \wedge \neg p_n) > 0$ and therefore $\mu(p_1 \vee \cdots \vee p_n) < 1$.

Definition 4.1. A theory T is n-probable modulo ϕ if there exists a probability measure μ such that $\mu(\phi) > 0$ and $\mu(\psi|\phi) > 1 - \frac{1}{n}$, for all $\psi \in T$.

Lemma 4.1. *If theory T is n-probable modulo ϕ, then T is n-probable and n-consistent.*

Proof. According to Lemma 3.2, it is sufficient to show that T is n-probable. The latter is an immediate consequence of Definition 4.1 and the fact that μ_ϕ is a probability measure. □

We may ask for which formulas we can go in the opposite direction: from n-probability to n-probability modulo ϕ.

Theorem 4.2. *Let T be a finite n-probable theory and let $\phi \in For_{\mathcal{P}}$. If ϕ is consistent with every consistent subset of T, then T is n-probable modulo ϕ.*

Proof. Let μ be the probability measure such that $\mu(\psi) > 1 - \frac{1}{n}$, for all $\psi \in T$.

Here we will use the proof of [20, Lemma 4.7], where it is shown that if there exists a probability measure μ such that $\mu(\psi) \geqslant r$ for all $\psi \in T$ and some $r \in [0,1]$, and if ϕ is consistent with every consistent subset of T, then there exists probability measure ν such that $\nu(\psi) \geqslant \mu(\psi)$ for $\psi \in T$ and $\nu(\phi) = 1$.

Now, for $\psi \in T$ we have

$$\nu(\psi|\phi) = \nu(\psi) \geqslant \mu(\psi) > 1 - \frac{1}{n}. \qquad \square$$

We further generalize this to the notion of "n-probable modulo $\{\phi_1,\ldots,\phi_k\}$". This would correspond to a situation where we have k agents, each knowing some fact ϕ_i, and each formula in our (inconsistent) theory is believed to be probable by at least one agent. It turns out that this can make some sense only in case that the agent's beliefs are highly compatible.

Definition 4.2. A theory T is n-probable modulo $\{\phi_1,\ldots,\phi_k\}$, if there exists a probability measure μ such that $\mu(\phi_1 \wedge \cdots \wedge \phi_k) > 0$, for all $i,j \in \{1,\ldots,k\}$, $\mu(\phi_i|\phi_j) > 1 - \frac{1}{n}$, and, for all $\psi \in T$, there exists an index $i \in \{1,\ldots,k\}$ such that

$$\mu(\psi|\phi_i) > 1 - \frac{1}{n}.$$

The following example illustrates the need for the "facts" $\phi_1,\ldots,\phi_k$ to support each other.

Example 4.2. Let us use the variation of the usual bird – penguin example. We introduce propositional letters for the following facts:

b – "is a bird".
p – "is a penguin".
f – "flies".
s – "lives on South Pole".
e – "the female sits on eggs".
d – "dives".

Let $T_1 = \{f, \neg s, e, \neg d\}$ and $T_2 = \{\neg f, s, \neg e, d\}$. Let μ be a measure such that for some large n:

- $\mu(p \wedge b) > 0$.
- $\mu(\phi|b) > 1 - \frac{1}{n}$ for $\phi \in T_1$.
- $\mu(\psi|p) > 1 - \frac{1}{n}$ for $\psi \in T_2$.
- $\mu(b|p) > 1 - \frac{1}{n}$, and
- $\mu(p|b) < \frac{1}{2}$.

Clearly, any reasonable measure will satisfy these conditions. Note that all requirements from Definition 4.2, except $\mu(p|b) > 1 - \frac{1}{n}$, hold. However, $T_1 \cup T_2$ is not even 2-consistent. So, if μ testified to the n-probability of $T_1 \cup T_2$ modulo $\{p, b\}$, this notion would be in total discord with the notion of n-consistency. Furthermore, there is no measure satisfying Definition 4.2 in this situation, which can be seen from Corollary 4.4.

Theorem 4.3. *If theory T is n-probable modulo $\{\phi_1,\ldots,\phi_k\}$, then T is $(n - k + 1)$-probable modulo $\phi_1 \wedge \cdots \wedge \phi_k$.*

Proof. Let $\nu(\psi) = \mu(\psi|\phi_1 \wedge \cdots \wedge \phi_k)$. For each $\psi \in T$ there exists $i \in \{1,\ldots,k\}$ such that

$$\mu(\psi|\phi_i) > 1 - \frac{1}{n}.$$

Choose an arbitrary $\psi \in T$. Without loss of generality, we may assume that the corresponding index i in the previous formula is equal to 1. Then:

$$\nu(\psi) = \frac{\mu(\psi \wedge \phi_1 \wedge \cdots \wedge \phi_k)}{\mu(\phi_1 \wedge \cdots \wedge \phi_k)} = \frac{\mu(\psi \wedge \phi_1 \wedge \cdots \wedge \phi_k)/\mu(\phi_1)}{\mu(\phi_1 \wedge \cdots \wedge \phi_k)/\mu(\phi_1)} = \frac{\mu(\psi \wedge \phi_2 \wedge \cdots \wedge \phi_k|\phi_1)}{\mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1)} \geqslant \frac{\mu(\psi|\phi_1) + \mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1) - 1}{\mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1)}$$
$$= 1 - \frac{1-\mu(\psi|\phi_1)}{\mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1)} > 1 - \frac{1-\left(1-\frac{1}{n}\right)}{\mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1)} = 1 - \frac{1}{n\mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1)}.$$

Applying (1), we obtain that

$$\mu(\phi_2 \wedge \cdots \wedge \phi_k|\phi_1) \geqslant \mu(\phi_2|\phi_1) + \cdots + \mu(\phi_k|\phi_1) - (k-2) > (k-1)\left(1-\frac{1}{n}\right) - (k-2) = \frac{n-k+1}{n}.$$

It follows that

$$\nu(\psi) > 1 - \frac{1}{n \cdot \frac{n-k+1}{n}} = 1 - \frac{1}{n-k+1}. \qquad \square$$

The next example, a generalization of "3 prisoners example", shows that Theorem 4.3 cannot be improved, i.e. that the bound $(n-k+1)$ is the best possible.

Example 4.3. Let there be $n+2$ prisoners (enumerated by $1,2,\ldots,n, n+1, n+2$) on the death row and let there be k guards, where n is "much" larger then k. By random choice one prisoner will be pardoned and this fact is confirmed by every guard.

Let the prisoners be divided into k groups. We may assume, without the loss of generality, that ith prisoner is not in the ith group. Assume that the ith group of prisoners asks the ith guard to tell them the name of one randomly selected prisoner which will be executed, and that his answer is "the prisoner number i".

After that, each prisoner sees his chance of survival increase from $\frac{1}{n+2}$ to $\frac{1}{n+1}$.

Let ψ_i be the sentence "ith prisoner will be executed", for $i \in \{1,\ldots,n+2\}$.

Let θ be the sentence "exactly one prisoner will be pardoned", and let, for $i \in \{1,\ldots,k\}, \phi_i$ be the answer of the ith guard, i.e., $\phi_i = \psi_i \wedge \theta$, and let μ be a probability measure that intuitively corresponds to the fact that a pardoned prisoner is chosen by the random choice.

Let $T = \{\psi_1,\ldots,\psi_{n+2}\}$. For each $\psi \in T$ there exists some $i \in \{1,\ldots,k\}$ such that

$$\mu(\psi|\phi_i) > 1 - \frac{1}{n}.$$

Furthermore, for all $i,j \in \{1,\ldots,k\}, \mu(\phi_i|\phi_j) > 1 - \frac{1}{n}$, so the conditions of Theorem 4.3 are satisfied.

It is easy to see that, for $i \leqslant k$

$$\mu(\psi_i|\phi_1 \wedge \cdots \wedge \phi_k) = 1$$

and for $i > k$

$$\mu(\psi_i|\phi_1 \wedge \cdots \wedge \phi_k) = 1 - \frac{1}{n-k+2}.$$

This shows that T is $(n-k+1)$-probable modulo $\phi_1 \wedge \cdots \wedge \phi_k$.

However, if for some μ, T were $(n-k+2)$-probable modulo $\phi_1 \wedge \cdots \wedge \phi_k$, we would have, for that μ, that

$$\mu(\psi_{k+1} \wedge \cdots \wedge \psi_{n+2}|\phi_1 \wedge \cdots \wedge \phi_k) > 0$$

which is impossible since $\psi_1 \wedge \cdots \wedge \psi_{n+2} \wedge \theta$ is a contradiction.

Corollary 4.4. *If theory T is n-probable modulo $\{\phi_1,\ldots,\phi_k\}$, then T is $(n-k+1)$-probable, hence $(n-k+1)$-consistent.*

We may define also a syntactic analogue of "n-probable modulo a set of formulas".

Definition 4.3. If T and S are sets of formulas, T is n-consistent modulo S if for any $\psi_1,\ldots,\psi_n \in T$, $\{\psi_1,\ldots,\psi_n\} \cup S$ is consistent.

Obviously, if T is n-consistent modulo S it is also n-consistent.

Corollary 4.5. *If theory T is n-probable modulo $\{\phi_1,\ldots,\phi_k\}$, then T is $(n-k+1)$-consistent modulo $\{\phi_1,\ldots,\phi_k\}$.*

Proof. For the probability measure ν from the proof of Theorem 4.3 and for any $\psi \in T$ we have $\nu(\phi_1 \wedge \cdots \wedge \phi_k) = 1$ and $\nu(\psi) > 1 - \frac{1}{n-k+1}$. Hence, $\nu(\psi_1 \wedge \cdots \wedge \psi_{n-k+1} \wedge \phi_1 \wedge \cdots \wedge \phi_k) > 0$ for all $\psi_1 \wedge \cdots \wedge \psi_{n-k+1} \in T$. $\square$

Remark 4.1. In [20] Knight defines a finite theory $T = \{\psi_1,\ldots,\psi_m\}$ to be *η-consistent given b* $= (\beta_1,\ldots,\beta_m)$ $(\eta \in [0,1], b \in [0,1]^m)$, if there exists a probability measure μ such that $\mu(\psi) \geqslant \eta$ for all $\psi \in T$ and $\mu(\psi_i) \geqslant \beta_i$. However, Knight considers only a very special case of the constraint vectors where $b \in \{0,1\}^m$. In that case, each b represents a characteristic function, i.e., it defines a subset $\{\psi_i | \beta_i = 1\}$ of T. We may obtain that a very particular case of n-probability modulo $\{\phi_1, \ldots, \phi_k\}$ is still stronger then Knight's property. Namely, if T is n-probable modulo $\{\phi_1, \ldots, \phi_k\}$, and if we restrict T to be finite and $\{\phi_1, \ldots, \phi_k\} \subseteq T$, the measure v from the proofs of Theorem 4.3 and Corollary 4.5 testifies that T is $\left(1 - \frac{1}{n-k+1}\right)$-consistent given b, where b is the vector corresponding to the set $\{\phi_1, \ldots, \phi_k\}$.

In [20, Theorem 7.4] Knight obtains a similar conclusion from syntactical premise (that T is minimally inconsistent theory of cardinality $n + 1$), which illustrates one of the differences between our approaches: while he is mainly interested in semantical conclusions, we start from semantical notion (n-probability modulo $\{\phi_1, \ldots, \phi_k\}$) and obtain syntactical conclusions (n-consistency in Corollary 4.4 and n-consistency modulo $\{\phi_1, \ldots, \phi_k\}$ in Corollary 4.5).

It is interesting that "conditioning" a theory T on the conjunction $\phi_1 \wedge \cdots \wedge \phi_k$ rather than on the set $\{\phi_1, \ldots, \phi_k\}$ results in a significant downgrade of its consistency level (from n to $n - k + 1$) even though we assumed that the formulas ϕ_i are highly compatible $\left(\mu(\phi_i | \phi_j) > 1 - \frac{1}{n}\right)$. The reason is, as can be seen from the proof of Theorem 4.3, that replacing any two formulas from $\{\phi_1, \ldots, \phi_k\}$ by their conjunction, reduces the bound for conditional probabilities from $1 - \frac{1}{n}$ to $1 - \frac{1}{n-1}$ and results in n-probability being replaced by $(n - 1)$-probability.

5. Strong *n*-probability

In this section, we try a different approach to "conditioning" by using a non-standard probability measure. Namely, instead of "high" conditional probability we use a much stronger notion of conditional probability infinitesimally close to 1. This will allow simpler definitions and stronger theorems.

Let R^* be a non-standard elementary extension of the standard real numbers (see [22]). An element ε of R^* is an infinitesimal if $|\varepsilon| < \frac{1}{n}$ for every positive integer n. For non-standard real numbers r and s we write $r \approx s$ to denote the fact that $|r - s|$ is an infinitesimal. We may define non-standard probability measures as in Definition 3.1. Furthermore, all definitions we introduced and all previous results hold for non-standard probability measures as well.

Definition 5.1. For a non-standard probability measure μ we say that ψ is a μ-consequence of ϕ, if $\mu(\psi | \phi) \approx 1$. We also say that the set $T \subseteq For_{\mathcal{P}}$ is a set of μ-consequences of the set $\Phi \subseteq For_{\mathcal{P}}$, if for every $\psi \in T$ there exist $\phi \in \Phi$ such that ψ is a μ-consequence of ϕ.

Example 5.1. Let $\phi, \theta \in For_{\mathcal{P}}$, $\mu(\phi) = \frac{1}{2}$, $\mu(\theta) \approx 0$ and $\mu(\theta) \neq 0$. It follows that $\mu(\phi | \phi \vee \theta) \approx 1$, $\mu(\neg\phi \mid \neg\phi \vee \theta) \approx 1$ and $\mu((\phi \vee \theta) \wedge (\neg\phi \vee \theta)) = \mu(\theta) \neq 0$. So, the set $\Phi = \{\phi \vee \theta, \neg\phi \vee \theta\}$ is consistent, but the set of μ-consequences $T = \{\phi, \neg\phi\}$ is inconsistent.

Our aim is to define a new notion of n-probability (relative to probability measure μ) which will be preserved under μ-consequence.

Definition 5.2. Let n be a positive integer. The theory T is strongly n-probable for μ, if for each n formulas $\psi_1, \ldots, \psi_n \in T$ the number $\mu(\psi_1 \wedge \cdots \wedge \psi_n)$ is not infinitesimal.

The theory T is strongly probable for μ, if it is strongly n-probable for μ for all $n \in \omega$.

Note that, for a fixed probability measure μ, strong n-probability (for μ) is the notion strictly stronger then n-consistency. Furthermore, strong probability for μ implies consistency; the converse implication does not hold.

Theorem 5.1. *Let Φ be a nonempty set, strongly n-probable for μ and let T be the set of μ-consequences of Φ.Then, T is also strongly n-probable for μ.*

Proof. Let $T = \{\psi_i | i \in \omega\}$ and suppose that, for all $i \in \omega, \phi_i$ is an element of Φ such that $\mu(\psi_i | \phi_i) \approx 1$ holds (note that for $i \neq j$ ϕ_i is not necessarily distinct from ϕ_j). Without the loss of generality, it is sufficient to prove that $\mu(\psi_1 \wedge \cdots \wedge \psi_n) \approx 0$ does not hold. Let $\psi \in \{\psi_1, \ldots, \psi_n\}$ and for the sake of simplicity suppose that $\psi = \psi_1$. As in the proof of Theorem 4.3, it holds that

$$\mu(\psi \mid \phi_1 \wedge \ldots \wedge \phi_n) \geqslant 1 - \frac{1 - \mu(\psi_1 | \phi_1)}{\mu(\phi_2 \wedge \cdots \wedge \phi_n | \phi_1)}.$$

Since $\mu(\psi_1 | \phi_1) \approx 1$ and $\mu(\phi_2 \wedge \cdots \wedge \phi_n | \phi_1) \geqslant \mu(\phi_1 \wedge \cdots \wedge \phi_n)$ is not an infinitesimal, by assumption, it follows that $\mu(\psi | \phi_1 \wedge \cdots \wedge \phi_n) \approx 1$, for all $\psi \in \{\psi_1, \ldots, \psi_n\}$.

Applying inequality (1) to measure $\mu(\cdot | \phi_1 \wedge \cdots \wedge \phi_n)$, we get

$$\mu(\psi_1 \wedge \ldots \wedge \psi_n | \phi_1 \wedge \ldots \wedge \phi_n) \geqslant \mu(\psi_1 | \phi_1 \wedge \cdots \wedge \phi_n) + \cdots + \mu(\psi_n | \phi_1 \wedge \cdots \wedge \phi_n) - (n-1) \approx 1.$$

In other words,

$$\frac{\mu(\psi_1 \wedge \cdots \wedge \psi_n \wedge \phi_1 \wedge \cdots \wedge \phi_n)}{\mu(\phi_1 \wedge \cdots \wedge \phi_n)} \approx 1.$$

or, equivalently,

$$\mu(\psi_1 \wedge \cdots \wedge \psi_n \wedge \phi_1 \wedge \cdots \wedge \phi_n) \approx \mu(\phi_1 \wedge \cdots \wedge \phi_n).$$

Since $\mu(\phi_1 \wedge \cdots \wedge \phi_n)$ is not infinitesimal, by the assumption, using Lemma 3.1 we conclude that $\mu(\psi_1 \wedge \cdots \wedge \psi_n)$ is not infinitesimal either. It follows that T is strongly n-probable for μ. □

This theorem is similar to Theorem 4.3 in the sense that we replaced the condition $\mu(\psi|\phi_i) > 1 - \frac{1}{n}$ by a much stronger condition $\mu(\psi - \phi_i) \approx 1$, which allowed us to drop the condition on "mutual support" of elements of Φ. Note that here Φ may be infinite.

Corollary 5.2. *Let Φ be a nonempty set, strongly probable for μ, and let T be the set of μ-consequences of Φ. Then, T is also strongly probable for μ.*

By this corollary, strong probability of the theory Φ is sufficient to provide strong probability of the set of its "consequences" T. If we are only interested in consistency of the set T, this assumption is not necessary:

Example 5.2. Suppose that $\Phi = \{\phi_i | i \in \omega\}$ is a consistent theory and $T = \{\psi_i | i \in \omega\}$ is a set of formulas. Let μ be a probability measure which satisfies the following conditions:

- For all $i \in \omega$ there exists an infinitesimal number ε_i such that $\mu(\psi_i|\phi_i) = 1 - \varepsilon_i$ holds,
- $\mu(\phi_1 \wedge \cdots \wedge \phi_n) > n \max\{\varepsilon_i | i = 1, \ldots, n\}$, for all $n \in \omega$ (this assumption is weaker then strong consistency of Φ for μ).

As in the proof of Theorem 4.3, we can derive the inequality (for $i \in \omega$)

$$\mu(\psi_i|\phi_1 \wedge \cdots \wedge \phi_n) \geqslant 1 - \frac{1 - \mu(\psi_i|\phi_i)}{\mu(\phi_1 \wedge \cdots \phi_{i-1} \wedge \phi_{i+1} \wedge \cdots \wedge \phi_n|\phi_i)}.$$

Hence, by assumption, we have

$$\mu(\psi_i|\phi_1 \wedge \cdots \wedge \phi_n) \geqslant 1 - \frac{\varepsilon_i}{\mu(\phi_1 \wedge \cdots \wedge \phi_n)} \mu(\phi_i) \geqslant 1 - \frac{1}{n}.$$

It follows that the set $\{\psi_1, \ldots, \psi_n\}$ is n-probable and, therefore, consistent. Consistency of T is a consequence of Compactness Theorem.

Since the notion of μ-consequence corresponds to defaults, as will be discussed in the next section, Theorem 5.1 may be understood as saying that the property of strong n-probability is preserved under default derivation. Corollary 5.2 says the same for the property of strong probability.

6. Application to defaults

Defaults are, roughly, rules with exceptions, like "birds fly", which allow deriving some conclusion in the absence of complete information. It has long been proposed that they should be interpreted as conditional probability infinitesimally close to 1 ([23,24]). Alternative proposals for interpreting defaults using standard probability were given in [25–27]. The paper [28] contains a proposal based on belief functions, while [29] uses confidence and possibility relations. Kraus et al. defined in [8] a system *P* of formal derivation rules which is generally recognized as capturing the core of default reasoning. Lehmann and Magidor proved in [9] that any consequence relation satisfying the system *P* and the rule of Rational Monotonicity (*RM*) may be interpreted by some non-standard probability measure (in the sense that $\alpha \mid\!\sim \beta$ iff ($\mu(\beta|\alpha) \approx 1$ or $\mu(\alpha) = 0$)), and conversely, that every relation defined in this way, starting from some non-standard probability measure μ, will satisfy the system *P* and *RM*.

In this sense, the relation of μ-consequence introduced in Definition 5.1 may be regarded as defining a default relation and, as mentioned at the end of the Section 5, Theorem 5.1 may be interpreted as speaking of preservation of strong n-probability under default derivation. Now we define a syntactic analogue of strong n-probability.

First we need a notion from [9]. We say that a formula ϕ is exceptional (for a default relation $\mid\!\sim$) iff $\top \mid\!\sim \neg\phi$.

Definition 6.1. The theory Φ is strongly n-consistent for a default relation $\mid\!\sim$, if for any n formulas $\phi_1, \ldots, \phi_n \in \Phi$ the formula $\phi_1 \wedge \cdots \wedge \phi_n$ is not exceptional.

The theory Φ is strongly consistent for $\mid\!\sim$, if it is strongly n-consistent for $\mid\!\sim$ for all $n \in \omega$.

Using Theorem 5.1 and the correspondence between default relations and non-standard probability measures from [9] we immediately obtain the following theorem:

Theorem 6.1. *Let Φ be nonempty set, strongly n-consistent for $\mid\!\sim$ and let Ψ be the set such that for every $\psi \in \Psi$, $\phi \mid\!\sim \psi$ holds for some $\phi \in \Phi$. Then, Ψ is also strongly n-consistent for $\mid\!\sim$.*

The meaning of this theorem may be illustrated by the familiar example "Nixon Diamond". If we denote the default "quakers are pacifists" by $q \mid\!\sim p$ and "republicans are not pacifists" by $r \mid\!\sim \neg p$, we see that from a consistent set $\{r, q\}$ we may obtain, by default derivation, an inconsistent set $\{p, \neg p\}$. Theorem 6.1s shows that an inconsistent set of consequences results from an "exceptional" set of assumptions.

We turn now to another application of our results to non-monotonic reasoning. It is clear that the people applying defaults to practical problems would prefer having a conditional probability finitely close to 1 rather than the one infinitely close to 1, i.e., they would rather work with the standard probability. The problem with such approach is that using standard conditional probability, with each step of deduction the probability decreases, so after a number of steps we may get a useless conclusions (e.g., that the probability is greater than 0).

We try now, in the style of Nilsson [30], or more recently Paris et al. [31], to define some device which would allow using standard measure but with controlled decrease of the probability throughout the derivation. We may notice that the conditions in the definition of "n-probable modulo $\{\phi_1,\ldots,\phi_k\}$" resemble one of the rules of the system P, so called rule of Cautious Monotonicity:

$$\frac{\alpha \mid\!\sim \beta, \alpha \mid\!\sim \gamma}{\alpha \wedge \beta \mid\!\sim \gamma}.$$

Restricting, in Definition 4.2, T to a single formula ψ and assuming $k = 2$ we obtain that ψ is n-probable modulo $\{\psi_1, \psi_2\}$ if (for some μ) $\mu(\psi|\phi_1) > 1 - \frac{1}{n}$ and $\mu(\phi_2|\phi_1) > 1 - \frac{1}{n}$ (plus $\mu(\phi_1 \wedge \phi_2) > 0$). Applying Theorem 4.3 we get that ψ will be at least $(n-1)$-probable modulo $(\phi_1 \wedge \phi_2)$. This gave us the idea to try to formulate "finite approximations" of rules of P in such a way that we obtain precise estimates of how much the probability decreases in each step of a deduction. In the sequel we define a series of non-monotonic consequence relations indexed by n, where n may be regarded as a degree of belief in entailment.

Definition 6.2. Let μ be a neat probability measure over $For_{\mathcal{P}}$. For every $n \in \omega$ we define binary relation $\mid\!\sim_n^\mu$ on $For_{\mathcal{P}}$, by $\alpha \mid\!\sim_n^\mu \beta$ iff $\mu(\beta|\alpha) > 1 - \frac{1}{n}$.

In the sequel we will omit μ from $\mid\!\sim_n^\mu$ whenever it is determined by the context. The following theorem expresses the P-like properties of this relation.

Theorem 6.2. *Let μ be a neat probability measure over $For_{\mathcal{P}}$. If the binary relations $\mid\!\sim_n$ are defined as above, then the following rules hold:*

$$REF_n : \frac{}{\alpha \mid\!\sim_n \alpha}; \quad LLE_n : \frac{\vdash \alpha \leftrightarrow \beta, \alpha \mid\!\sim_n \gamma}{\beta \mid\!\sim_n \gamma};$$

$$RW_n : \frac{\vdash \alpha \rightarrow \beta, \gamma \mid\!\sim_n \alpha}{\gamma \mid\!\sim_n \beta}; \quad AND_n : \frac{\alpha \mid\!\sim_{2n} \beta, \alpha \mid\!\sim_{2n} \gamma}{\alpha \mid\!\sim_n \beta \wedge \gamma};$$

$$OR_n : \frac{\alpha \mid\!\sim_{2n} \gamma, \beta \mid\!\sim_{2n} \gamma}{\alpha \vee \beta \mid\!\sim_n \gamma}; \quad CM_n : \frac{\alpha \mid\!\sim_n \beta, \alpha \mid\!\sim_n \gamma}{\alpha \wedge \beta \mid\!\sim_{n-1} \gamma}.$$

Proof. The proofs of REF_n, LLE_n and RW_n are trivial.

CM_n is already discussed.

AND_n: Let $\mu(\beta|\alpha) > 1 - \frac{1}{2n}$ and $\mu(\gamma|\alpha) > 1 - \frac{1}{2n}$; using CM_{2n} we obtain $\mu(\gamma|\alpha \wedge \beta) > 1 - \frac{1}{2n-1}$. From the equality

$$\mu(\beta \wedge \gamma|\alpha) = \mu(\beta|\alpha)\mu(\gamma|\alpha \wedge \beta)$$

it follows that $\mu(\beta \wedge \gamma|\alpha) > \left(1 - \frac{1}{2n}\right)\left(1 - \frac{1}{2n-1}\right) = 1 - \frac{1}{n}$ and hence $\alpha \mid\!\sim_n \beta \wedge \gamma$.

OR_n: Let $\mu(\gamma|\alpha) > 1 - \frac{1}{2n}$ and $\mu(\gamma|\beta) > 1 - \frac{1}{2n}$. We will use abbreviations α_0 and β_0 for $\alpha \wedge \neg\gamma$ and $\beta \wedge \neg\gamma$, respectively.

$$\mu(\gamma|\alpha \vee \beta) = \frac{\mu((\alpha \wedge \gamma) \vee (\beta \wedge \gamma))}{\mu(\alpha \vee \beta)} = \frac{\mu((\alpha \vee \beta) \wedge \neg(\alpha_0 \vee \beta_0))}{\mu(\alpha \vee \beta)} \geqslant \frac{\mu(\alpha \vee \beta) - \mu(\alpha_0 \vee \beta_0)}{\mu(\alpha \vee \beta)} = 1 - \frac{\mu(\alpha_0 \vee \beta_0)}{\mu(\alpha \vee \beta)}$$

By assumption, $\mu(\alpha_0) < \frac{\mu(\alpha)}{2n} \leqslant \frac{\mu(\alpha \vee \beta)}{2n}$ and $\mu(\beta_0) \leqslant \frac{\mu(\alpha \vee \beta)}{2n}$, which implies $\mu(\alpha_0 \vee \beta_0) < \frac{\mu(\alpha \vee \beta)}{n}$.

Finally, $\mu(\gamma|\alpha \vee \beta) > 1 - \frac{1}{n}$, or, equivalently, $\alpha \vee \beta \mid\!\sim_n \gamma$. □

Example 6.1. This example demonstrates how the relations $\mid\!\sim_n$ may be used to overcome the so called "inheritance blocking" problem in default reasoning. Other proposals for solving this problem were given in [28,32,33].

Let the default base consist of the following rules:

The Swedes are, generally, blond ($s \mid\!\sim b$).
The Swedes are, generally, tall ($s \mid\!\sim t$).

The problem is that in most default systems we cannot derive the intuitively plausible conclusion that Swedes which are not tall, generally are also blond ($s \wedge \neg t \mid\!\sim b$). We propose two solutions, based on adding different types of additional assumptions.

In the first approach, along the lines proposed in [34] (see also [35]), our belief in blondness of Swedes is much higher than our belief in their tallness. Say $s \mid\!\sim_{n^2} b$ and $s \mid\!\sim_{n-1} t$ but not $s \mid\!\sim_n t$. Then we obtain (as in the proof of Theorem 4.3)

$$\mu(b|s \wedge \neg t) > 1 - \frac{1 - \mu(b|s)}{\mu(\neg t|s)}.$$

As $1 - \mu(b|s) < \frac{1}{n^2}$ and $\mu(\neg t|s) \geqslant \frac{1}{n}$ we conclude

$$\mu(b|s \wedge \neg t) > 1 - \frac{1}{n}.$$

Therefore $s \wedge \neg t \mid\!\sim_n b$.

Note, however, that in this approach we cannot obtain $s \wedge \neg b \mid\!\sim_n t$!

In another approach, we note that the desirability of the conclusion $s \wedge \neg t \mid\!\sim b$ results from our intuition that height and hair color are independent features. However, independence of two variables on a set does not imply their independence on every subset. Obvious solution seems to be to add assumption that on the "set of Swedes", blondness and tallness are independent:

$$\mu(b \wedge t|s) = \mu(b|s)\mu(t|s).$$

Now, assuming $s \mid\!\sim_m b$ and $s \mid\!\sim_n t$ (where m and n are any two "large" numbers) we get

$$s \wedge \neg t \mid\!\sim_m b,$$
$$s \wedge \neg b \mid\!\sim_n t.$$

This is seen from the condition of independence:

$$\frac{\mu(b \wedge s \wedge \neg t)}{\mu(s)} = \frac{\mu(b \wedge s)}{\mu(s)} \frac{\mu(\neg t \wedge s)}{\mu(s)} \tag{5}$$

(note that independence of b and t implies also the independence of b and $\neg t$, and $\neg b$ and t).

Finally, from

$$\mu(b|s \wedge \neg t) = \frac{\mu(b \wedge s \wedge \neg t)}{\mu(s \wedge \neg t)}$$

and equality (5) we obtain

$$\mu(b|s \wedge \neg t) = \frac{\mu(b \wedge s)\mu(\neg t \wedge s)}{\mu(\neg t \wedge s)\mu(s)} = \mu(b|s) > 1 - \frac{1}{m}.$$

Remark 6.1. Note that the converse of Theorem 6.2 does not hold. Namely, there are obvious properties of the probability measure which are not captured by the given rules. For example,

$$\frac{\alpha \mid\!\sim_n^\mu \beta}{\alpha \mid\!\sim_m^\mu \beta}$$

will be interpreted as true for any μ and $m < n$, but it is not a consequence of the above rules. Therefore, we could have some relation which satisfies all the rules from Theorem 6.2, but not this one. It would not be possible to find a probability which would interpret this relation.

We have shown that applying non-monotonic rules to consequence relations $\mid\!\sim_n$, the strength of the conclusion is, in the worst case, twice weaker then the strength of assumptions.

Next, we turn to the following question: if we use both rational relation $\mid\!\sim$ and $\mid\!\sim_n$ in one of the non-monotonic rules, does the strength (n) transfer to the conclusion? Consider the following example, which is essentially a modification of examples from [25].

Example 6.2. Suppose that the statistical knowledge "more than 95% of birds fly" is available, and that we accept the default rule "generally, birds have wings". (The former can be expressed in our terminology with $b \mid\!\sim_{20} f$, while the later is expressible by default rule $b \mid\!\sim w$, usually interpreted as "conditional probability of w knowing b is approximately 1".)

What can we say about the birds with the wings? Intuitively, the conclusion that they fly with the probability greater than 95% is quite acceptable. On the other hand, the best we can calculate is that the probability is either greater than or infinitely close to 95%.

We overcome the above difficulty by slightly changing the notion of $\mid\!\sim_n$.

Theorem 6.3. *Let $\mid\!\sim$ be a rational relation on $For_{\mathcal{P}}$, and let μ be a corresponding neat non-standard probability measure. If the binary relations $\mid\!\sim_n$ on $For_{\mathcal{P}}$ are defined by $\alpha \mid\!\sim_n \beta$ iff $\mu(\beta|\alpha) > 1 - \frac{1}{n}$ or $\mu(\beta|\alpha) \approx 1 - \frac{1}{n}$, then the following rules hold:*

$$LLE_n^{\approx}: \frac{\vdash \alpha \leftrightarrow \beta, \alpha \mid\!\sim_n \gamma}{\beta \mid\!\sim_n \gamma}; \quad RW_n^{\approx}: \frac{\vdash \alpha \to \beta, \gamma \mid\!\sim_n \alpha}{\gamma \mid\!\sim_n \beta};$$
$$OR_n^{\approx}: \frac{\alpha \mid\!\sim \gamma, \beta \mid\!\sim_n \gamma}{\alpha \vee \beta \mid\!\sim_n \gamma}; \quad AND_n^{\approx}: \frac{\alpha \mid\!\sim \beta, \alpha \mid\!\sim_n \gamma}{\alpha \mid\!\sim_n \beta \wedge \gamma};$$
$$CM1_n^{\approx}: \frac{\alpha \mid\!\sim_n \beta, \alpha \mid\!\sim \gamma}{\alpha \wedge \beta \mid\!\sim_n \gamma}; \quad CM2_n^{\approx}: \frac{\alpha \mid\!\sim \beta, \alpha \mid\!\sim_n \gamma}{\alpha \wedge \beta \mid\!\sim_n \gamma}.$$

Proof. The proof is an easy modification of the proof of Theorem 6.2. We give just one example. For the proof of $OR_n^{\approx}$ let us suppose $\mu(\gamma|\alpha) = 1 - \varepsilon_1$ and $\mu(\gamma|\beta) > 1 - \frac{1}{n} - \varepsilon_2$ (ε_1,$\varepsilon_2 \approx 0$). As in the proof of OR_n, we can obtain $\mu(\gamma|\alpha \vee \beta) \geqslant 1 - \frac{\mu(\alpha_0 \vee \beta_0)}{\mu(\alpha \vee \beta)}$, where α_0 is $\alpha \wedge \neg\gamma$ and β_0 is $\beta \wedge \neg\gamma$. From $\mu(\alpha_0) \leqslant \varepsilon_1 \ \mu(\alpha \vee \beta)$ and $\mu(\beta_0) < (\varepsilon_2 + \frac{1}{n})\mu(\alpha \vee \beta)$ we conclude $\mu(\alpha_0 \vee \beta_0) \leqslant (\varepsilon_1 + \varepsilon_2 + \frac{1}{n})\mu(\alpha \vee \beta)$. Thus, $\mu(\gamma|\alpha \vee \beta) \geqslant 1 - \frac{1}{n} - \varepsilon_1 + \varepsilon_2 \approx 1 - \frac{1}{n}$. □

The above statement is in spirit of [25, Theorem 5.1] where some rules with similar combinations of default knowledge and probabilistic knowledge are presented.

The LPP^S-logic introduced in [34] is a suitable syntactic framework for modeling default reasoning. The logic enriches propositional calculus with probabilistic operators which are applied to propositional formulas: $CP_{\geqslant s}(\alpha,\beta)$, $CP_{\leqslant s}(\alpha,\beta)$ and $CP_{\approx s}(\alpha,\beta)$, with the intended meaning "the conditional probability of α given β is at least s", "at most s" and "approximately s", respectively. The corresponding range of probabilistic functions is chosen to be the unit interval of a recursive nonarchimedean field, making it possible to express formulas of the form $CP_{\approx 1}(\alpha,\beta)$ that may be used to model defaults $\beta \mid\!\sim \alpha$. For example, the above rule $OR_n^{\approx}$ can be written as

$$CP_{\approx 1}(\gamma, \alpha) \wedge (CP_{\geqslant 1-\frac{1}{n}}(\gamma, \beta) \vee CP_{\approx 1-\frac{1}{n}}(\gamma, \beta)) \to (CP_{\geqslant 1-\frac{1}{n}}(\gamma, \alpha \vee \beta) \vee CP_{\approx 1-\frac{1}{n}}(\gamma, \alpha \vee \beta)).$$

The paper [34] provides a consistent and strongly complete axiomatization of the logic, so that (the LPP^S-translations) of all P-rules are theorems of the logic. The same holds for (the LPP^S-translations) of all rules from the Theorems 6.2 and 6.3. Additionally, [34] proves decidability of the LPP^S-logic. It follows easily that if a formula (representing the LPP^S-translation of defaults and/or approximate defaults) is not an LPP^S-theorem, then it is not a consequence of (the LPP^S-translations) of the above rules.

7. Conclusion

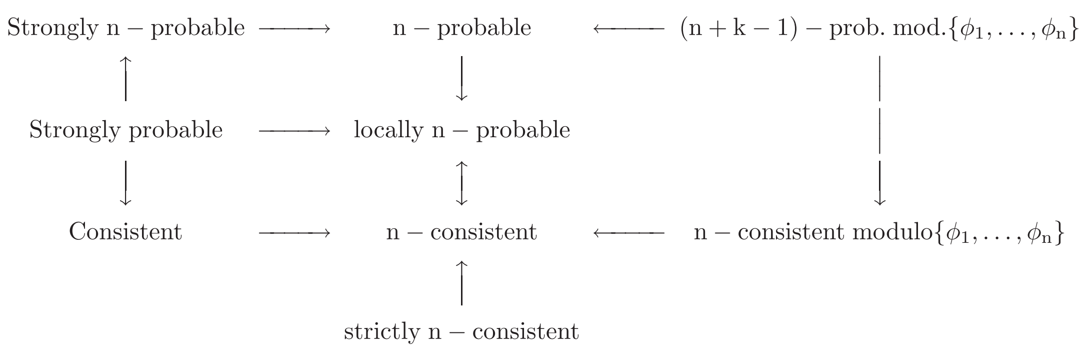
Attempting to differentiate inconsistent theories we introduced a syntactic notion of n-consistency (based on the number of formulas needed to derive a contradiction) and a semantic notion of n-probability (based on the existence of a probability measure which assigns the probability greater than $1 - \frac{1}{n}$ to every formula of the theory). Our notion of n-probability is similar to Knight's notion of η-consistency from [20] (as discussed in detail in the Introduction and Remark 3.1).

It turned out that the notion of n-probability is stronger and we introduced the weaker notion of local n-probability which is equivalent to n-consistency. This seems quite reasonable because both notions refer to $(n + 1)$-element subsets of the theory, unlike the notion of n-probability.

Using conditional probabilities we introduced the notion of n-probability modulo a formula ϕ and n-probability modulo a set of formulas, and showed the connections between these two properties and the property of n-consistency. Similar, but more elegantly formulated results are obtained by switching to non-standard probability measure.

Finally, we apply these results to default systems. Using the connection between rational default relations and non-standard probabilities ([9]), we introduced the notion of strong n-consistency (for a default relation $\mid\!\sim$) and showed that it is preserved under default derivation. Another application tries to define finite approximations of defaults in the style of Nilsson ([30]). Using our notion "ψ is n-probable modulo ϕ" we defined relations $\mid\!\sim_n$ which satisfy a weak version of system P from [8]. If we interpret $\phi \mid\!\sim_n \psi$ as saying that we believe with degree n that ψ follows from ϕ, the new rules determine by how much the degree of belief decreases in each step of possible deduction. We also show that if we combine $\mid\!\sim_n$ with the usual defaults in such way that in the rules of the system P with two premises, one premiss is with $\mid\!\sim$ and the other with $\mid\!\sim_n$ (similar to [25]), the probability of the conclusion does not decrease.

We may sum up the relations between the introduced notions in the following diagram:



One possible line of continuing this research would be to investigate if there are any relations of measures of inconsistency introduced in this paper with other syntactic measures of inconsistency which are inspired by paraconsistent logic like ones defined by Schotch and Jennings [36] or Brown [37]. Also, since the converse of Theorem 6.2 does not hold, it would be interesting to find additional rules that would make the system complete for the probability measure semantics.

Acknowledgments

This work was supported by the Ministarstvo za nauku i tehnolo śkirazvoj Republike Srbije, through Matematički Institut [Project 144013].

The authors thank the anonymous referees whose comments helped clarify a number of points and improve the text.

References

[1] A. Hunter, S. Konieczny, Approaches to measuring inconsistent information, in: Inconsistency Tolerance, Lecture Notes in Computer Science, vol. 3300, Springer, 2006, pp. 189–234.
[2] D.M. Gabbay, A. Hunter, Making inconsistency respectable (Part 1), in: Fundamentals of Artificial Intelligence Research, Lecture Notes in Artificial Intelligence, vol. 535, Springer-Verlag, Berlin, 1991, pp. 19–32.
[3] D.M. Gabbay, A. Hunter, Making inconsistency respectable (Part 2), in: Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Lecture Notes in Computer Science, vol. 747, Springer-Verlag, Berlin, 1993, pp. 129–144.
[4] D. Dubois, J. Lang, H. Prade, Possibilistic logic, in: D.M. Gabbay et al. (Eds.), Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3, Oxford University Press, 1994, pp. 439–513.
[5] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and *n*-person games, Artificial Intelligence 77 (1995) 321–357.
[6] P. Gardenfors, H. Rott, Belief revision, Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 4, Oxford University Press, 1995, pp. 35–132.
[7] A. Hunter, Paraconsistent logics, in: D.M. Gabbay, Ph. Smets (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 2, Kluwer, 1998, pp. 11–36.
[8] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artificial Intelligence 44 (1990) 167–207.
[9] D. Lehmann, M. Magidor, What does a conditional knowledge base entail?, Artificial Intelligence 55 (1992) 1–60
[10] G. Priest, Paraconsistent logics, Handbook of Philosophical Logic, vol. 6, Kluwer, 2002.
[11] D. Makinson, General patterns in nonmonotonic reasoning, Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3, Clarendon Press, Oxford, 1994, pp. 35–110.
[12] L. Amgoud, S. Kaci, An argumentation framework for merging conflicting knowledge bases, International Journal of Approximate Reasoning 45 (2007) 321–340.
[13] A. Hunter, S. Konieczny, Measuring inconsistency through minimal inconsistent sets, in: Proceedings of the 11th International Conference on Knowledge Representation, AAAI Press, 2008, pp. 358–366.
[14] J. Grant, A. Hunter, Measuring inconsistency in knowledgebases, Journal of Intelligent Information Systems 27 (2006) 159–184.
[15] A. Hunter, S. Konieczny, Shapely inconsistency values, in: Proceedings of the 10th International Conference on Knowledge Representation, 2006, pp. 249–259.
[16] S. Konieczny, J. Lang, P. Marquis, Quantifying information and contradiction in propositional logic through epistemic actions, in: Proceedings of the 18th International Joint Conference on Artificial Intelligence, 2003, pp. 106–111.
[17] R. Sorensen, Blindspots, Claredon Press, Oxford, 1988.
[18] S. Konieczny, J. Lang, P. Marquis, Reasoning under inconsistency: the forgotten connective, in: Proceedings of the 19th International Joint Conference on Artificial Intelligence, 2005, pp. 484–489.
[19] N. Rescher, R. Manor, On inference from inconsistent premises, Theory and Decision 1 (1970) 179–219.
[20] K. Knight, Measuring inconsistency, Journal of Philosophical Logic 31 (2001) 77–91.
[22] K.D. Stroyan, W.A.J. Luxemburg, Introduction to the Theory of Infinitesimals, Academic Press, 1976.
[23] E.W. Adams, The Logic of Conditionals, Reidel, Dordrecht, 1975.
[24] J. Pearl, System Z: a natural ordering of defaults with tractable applications to default reasoning, in: Proceedings of the 3rd Conference on the Theoretical Aspects of Reasoning about Knowledge, Morgan Kaufmann, 1990, pp. 121–135.
[25] T. Lukasiewicz, Weak nonmonotonic probabilistic logic, Artificial Intelligence 168 (1–2) (2005) 119–161.
[26] V. Biazzo, A. Gilio, T. Lukasiewicz, G. Sanfilippo, Probabilistic logic under coherence, model theoretic probabilistic logic, and default reasoning in system P, Journal of Applied Non-Classical Logics 12 (2) (2002) 189–213.
[27] P. Snow, Diverse confidence levels in a probabilistic semantics for conditional logics, Artificial Intelligence 113 (1999) 269–279.
[28] S. Benferhat, A. Saffiotti, P. Smets, Belief functions and default reasoning, Artificial Intelligence 122 (2000) 1–69.
[29] D. Dubois, H. Fargier, H. Prade, Ordinal and probabilistic representations of acceptance, Journal of Artificial Intelligence Research 22 (2004) 23–56.
[30] N. Nilsson, Probabilistic logic, Artificial Intelligence 28 (1986) 71–87.
[31] J. Paris, D. Picado Muino, M. Rosefield, Inconsistency as qualified truth: a probability logic approach, International Journal of Approximate Reasoning 50 (2008) 1151–1163.
[32] S.Tan, J. Pearl, Specificity and inheritance in default reasoning, in: Proceedings of International Joint Conference on Artificial Intelligence, 1995, pp. 1480–1486.
[33] E. Weydert, System JLZ – rational default reasoning by minimal ranking constructions, Journal of Applied Logic 1 (3–4) (2003) 273–308.
[34] M. Rašković, Z. Marković, Z. Ognjanović, A logic with approximate conditional probabilities that can model default reasoning, International Journal of Approximate Reasoning 49 (2008) 52–66.
[35] M. Rašković, Z. Ognjanović, Z. Marković, A logic with conditional probabilities, in: Logics in Artificial Intelligence, Lecture Notes in Artificial Intelligence, vol. 3229, Springer-Verlag, 2004, pp. 226–238.
[36] P. Schotch, R. Jennings, On detonating, in: Paraconsistent Logic: Essays on the Inconsistent, Philosophia Verlag, 1989. pp. 306–327.
[37] B. Brown, Yes, Virginia, there really are paraconsistent logics, Journal of Philosophical Logic 28 (1999) 137–148.