Probabilistic Multi-Context Systems

Marco Sotomayor[1], Kewen Wang[1], Yi Dong Shen[2], and John Thornton[1]

[1]Griffith University, Australia
[2]Institute of Software, Chinese Academy of Sciences, China

Abstract

The concept of contexts is widely used in artificial intelligence. Several recent attempts have been made to formalize multi-context systems (MCS) for ontology applications. However, these approaches are unable to handle probabilistic knowledge. This paper introduces a formal framework for representing and reasoning about uncertainty in multi-context systems (called p-MCS). Some important properties of p-MCS are presented and an algorithm for computing the semantics is developed. Examples are also used to demonstrate the suitability of p-MCS.

1 Introduction

The formalization of context is a critical building block towards the achievement of the semantic web vision [5]. In order to deliver accurate and unambiguous information, ontology-driven applications rely significantly on context modeling [6]. An *ontology* is a formal representation of shared terms and their relationships for an application domain. Ontologies have been widely applied in situations where the use and management of shared information are core issues (e.g. in medical applications [12, 14]). According to the vision of the semantic web, information on the internet will also be represented as ontologies [5]. In this setting, explicit models of semantic information are needed in order to support information exchange. Since shared ontologies define a common understanding of terms for an application of interest, the use of ontologies makes it possible to communicate and exchange information between different users and systems on a semantic level. However, ontologies can be used only when a consensus about their contents is reached. Moreover, building and maintaining ontologies can become difficult in a dynamic, open and distributed domain such as the internet. To enhance the use and management of highly distributed ontologies, the framework of *contextual ontologies* has been established and an extension of OWL called *C-OWL* has been introduced in [1]. C-OWL is based on the theory of *multi-context systems* (MCS) [4]. To the best of our knowledge, *the problem of incorporating probabilistic information into multi-context systems is still open*

Multi-Context Systems (MCS), constitute one of the most recognized and mature formalizations of context in AI [15]. MCS are a generalization of Natural Deduction systems, which allow the use of different languages through a mechanism of tagged formulae [15]. This implies that in different languages or contexts, a logical proposition can be interpreted in different ways. However, classical logic cannot express the degree of certainty of premises, nor the degree of certainty in conclusions derived from these premises (Williamson cited by [8]).

Figure 1 illustrates this situation using a typical Magic Box example [4] where Mr1 and Mr2 are unable to distinguish the depth of a ball inside a magic box. It is assumed that Mr1 and Mr2 are both almost blind but have knowledge about the compatibility relation between their different perspectives. So, Mr1 cannot answer with certainty if there is a ball on the right, he can only assume according his knowledge that "there is a ball on the right" with probability p. The same reasoning applies to Mr2, except that the probabilities for Mr2 are calculated in relation to Mr2's context. A probabilistic multi-context system provides a language for representing what Mr1 and Mr2 know about their environment that allows them infer new probabilistic knowledge based on what they already know. Existing semantics of MCS are unable to handle probabilistic knowledge of contexts. For this reason the aim of this paper

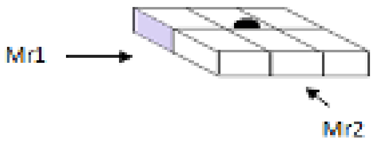

Figure 1: Magic Box

is to introduce a semantic framework for representing and reasoning about probabilistic knowledge in logic-based MCS.

The probabilistic logic approach in this paper is based on the work of [3, 9, 10], which dealt with formalization and semantics for uncertainty in logic programming. An important contribution of the current research is the introduction of probability theory into MCS, which provides more expressive languages in different contexts without losing the original logic of multi-context systems proposed by Giunchiglia [4]. The paper also shows that probabilistic multi-context systems can be reduced to MCS by simply assigning a probability of one to every proposition. This shows that MCS are a particular case of more general probabilistic multi-context systems. In addition, the idea of minimal information in context proposed by [13] is preserved in order to emphasize and contrast the relationship between MCS and p-MCS, and at the same time provide a probabilistic notion of the information entailed logically.

2 Probabilistic Multi-Context Systems

This section introduces the *p-MCS* framework for representing and reasoning about probabilistic information in multi-context systems. The first step in formalizing p-MCS is to specify the type of language used in each context.

Suppose that we have a set of contexts called K and denote context k as an element of K. Each context k is associated with a finite set A_k of labeled atoms of the form $k : a$ (a denotes an atom). Informally, k : a means that atom a belongs to context k. Note that A_k contains only the atomic propositions needed to express the basic knowledge in context k. A propositional language L_k is constructed over A_k in a standard sense. A formula in context k can be expressed as $\phi \in L_k$. The next step is to introduce an uncertainty degree to a formula in L_k.

Let $k : \phi$ be a labeled formula and μ be a point probability between [0, 1], then a formula function in a particular context $(k : \phi)\mu$ is called ***p-labeled formula***. Intuitively, $(k : \phi)\mu$ means that the probability of formula ϕ in context k is μ.

Definition 1. *A p-labeled rule r is of the form $(k : F)\mu \leftarrow (k_1 : F_1)\mu_1, \dots, (k_n : F_n)\mu_n$ where $n \geq 0$, $(k : F)\mu$ and each $(k_i : F_i)\mu_i$ are p-labeled formulas.*

Informally, this rule reads that if the probability of each F_i in context k_i is equal to μ_i for $i = 1, \dots, n$, then the probability of F in context k is equal to μ. The p-labeled formula $(k : F)\mu$ is called the head of the p-labeled rule r, denoted $head(r)$. The set $\{(k_1 : F_1)\mu_1, \dots, (k_n : F_n)\mu_n\}$ of p-labeled formulas is called the body of r, denoted $body(r)$. We remark that in the traditional probabilistic logic programming approaches [10, 9], probabilities are assigned to atoms and the premises and head of the rule belong to the same context.

A local p-rule r for context k is a p-labeled rule without premises or body. Local knowledge of context k is a set of its local p-rules.

A bridge rule r for context k is a p-labeled rule such that (1) $head(r)$ has label k and (2) a p-labeled formula in $body(r)$ belongs to at least one context apart from context k. Such a rule provides a way for inferring new knowledge in context k from other contexts.

Definition 2. *A probabilistic multi-context system (or p-MCS) T is of the form $[(R_1,B_1),\ldots,(R_m,B_m)]$ where R_k is a set of local p-rules for context k and B_k is a set of bridge rules for context k for $k = 1,\ldots,m$.*

In the same way as for ordinary MCS [13], a probabilistic multi-context system is defined as a specification of contextual probabilistic information and inter-contextual probabilistic information flow. Contextual information can be specified through local p-rules (facts) and inter-contextual information flow through bridge rules.

Given a p-MCS $T = [(R_1,B_1),\ldots,(R_m,B_m)]$, each context (R_k,B_k) represents two types of knowledge: (1) the knowledge $\mathscr{B}_k$ directly derived from the local context k and other contexts through bridge rules; and (2) the knowledge η_k inferred from $\mathscr{B}_k$ through probabilistic reasoning [10]. This type of reasoning problem is called probabilistic entailment [10], and is formally expressed as follows:

$$\zeta_k = \mathscr{B}_k \bigcup \eta_k \tag{1}$$

Example 1. *The scenario illustrated in Figure 1 can be formalized as a p-MCS as follows:*
l = "The ball is on the left" r = "The ball is on the right" c = "The ball is in the center".
$T = [(R_1,B_1),(R_2,B_2)]$. The two contexts are defined as $(R_1 = \{r_1\}, B_1 = \{r_3\})$, $(R_2 = \{r_2\}, B_2 = \{r_4\})$.

$$\begin{array}{llll} r_1: & (1:\neg r)0.5 & \leftarrow & \\ r_2: & (2:c)0.5 & \leftarrow & \\ r_3: & (1:l\vee r)0.75 & \leftarrow & (2:l\vee c\vee r)0.875 \\ r_4: & (2:l\vee c\vee r)0.875 & \leftarrow & (1:l\vee r)0.75 \end{array}$$

Then $\mathscr{B}_1 = \{(1:\neg r)0.5, (1:l\vee r)0.75\}$.
To explain how to determine η_1, we need some basics of probabilistic reasoning.

One important concept in probability theory is the notion of worlds or atomic events [11]. Given a propositional language, we define a world as a Herbrand interpretation.

Let N_k be the number of propositions in A_k. Then there are 2^{N_k} possible worlds for context k. For example, if $A_1 = \{1:l, 1:r\}$, then there are four possible worlds W_j $(j = 1,2,3,4)$ and each such world W_j is associated with a probability w_j. We use LM_k to denote a finite set of possible worlds for context k: $LM_k = \{W_1,\ldots,W_m\}$.

Definition 3. *A contextual world probability density function for a context k is defined as a function $WP_k : LM_k \to [0,1]$ satisfying $\Sigma_{W\in LM_K} WP_k(W) = 1$. Denote $WP_k(W_j) = w_j$; $0 \le w_j \le 1$, $1 \le j \le m$ (m denotes possible worlds for context k).*

Because the possible worlds for every context are different, a world probability density function has to be defined for every context.

A *p-local interpretation* for context k is a pair $W:\mu$ of an interpretation W of L_k and an (associated) probability value μ. Let M_k be the set of p-local interpretations for context k : $M_k = \{W_1 : w_1,\ldots,W_m : w_m\}$

Definition 4. *A p-labeled chain is of the form: $c = \{c_1,\ldots,c_m\}$ where $c_k \subseteq M_k$ for $k = 1,\ldots,m$.*

A p-labeled chain describes a world probability density function for every context k.

Every world is mutually exclusive and a proposition is equal to the disjunction of all the worlds where it holds [11]. For example: $l = \{W_1 \bigcup W_2\}$.

The general laws of the probability theory can be deduced through Kolmogorov's axioms [11]. One of Kolmogorov's axioms states that: $p(W_1 \vee W_2) = p(W_1) + p(W_2) - p(W_1 \wedge W_2)$. Because W_1 and W_2 are mutually exclusive, then: $p(l) = p(W_1 \vee W_2) = p(W_1) + p(W_2) - 0 = w_1 + w_2$

Given this demonstration it can be stated that the probability of a proposition is equal to the sum of the probabilities of the worlds where it holds (where a proposition is true).

Definition 5. *A contextual probabilistic interpretation is a mapping from L_k to [0,1] defined as follows: For each $\phi \in L_k$: $\mathscr{I}_{wp_k}(\phi) = \Sigma_{W \models \phi}\ WP_k(W)$*

Given a set of contextual probabilistic interpretations $\mathscr{I}_{wp_k}(\phi)$ in context k, a set of equations is generated under the following constraints:

I. $\Sigma_{W_j \models \phi}\ W_j = \mu$, for all $\mathscr{I}_{wp_k}(\phi) = \mu$.

II. $\Sigma_{j=1}^{m}\ w_j = 1$.

III. $0 \leq w_j \leq 1$, $1 \leq$ j $\leq$ m.

Now the notion of satisfiability for p-MCS can be introduced:

- A p-labeled chain c satisfies a p-labeled formula $(k : \phi)\mu$ iff $\mathscr{I}_{wp_k}(\phi) = \mu$
- A p-labeled chain c satisfies a p-labeled rule, iff whenever c satisfies the body of the rule body(r) then the head of the rule head(r) must be also satisfied.
- A p-labeled chain c satisfies a system T (i.e p-MCS) iff it satisfies every p-labeled rule of the system.

3 Minimal Probabilistic Entailed Chain/Fixpoint

The following section describes the process of constructing the probabilistic solution chain and the minimal probabilistic entailed chain and shows how to test if the contextual probabilistic interpretations are consistent in the set of equations that are generated. Finally, it is shown how to construct the contextual world probability density.

Let $\mathscr{C}$ be the set of all p-labeled chains. It is possible to order the p-labeled chains according to the amount of information that they contain. A p-labeled chain c is less informative than c' ($c \preceq c'$), if for every context k, $c_k \supseteq c_k$' [13].

Definition 6. *A p-labeled solution chain c_p of a p-MCS is a p-labeled chain such that satisfying the p-MCS.*

Based on [13], it can be argued that a minimal solution chain c_s in a non-probabilistic MCS contains all the logical entailments for every context k .

Definition 7. *A minimal probabilistic entailed chain c_e contains all the probabilistic entailments η_k per every context k. $c_e : c_s \to [0,1]; c_e = \{\eta_1, \ldots, \eta_m\}$*

i.e. c_e is the result of a mapping of all Herbrand interpretations of c_s to a probability between [0,1].

Proposition 1. *Let c_p and c_e be a probabilistic solution chain and a minimal probabilistic entailed chain of a p-MCS, respectively. Then $c_p \supseteq c_e$*

A non-probabilistic multi-context system is a particular case of a probabilistic multi-context system where every formula in the system is associated with a probability of one. The minimal solution chain c_s [13] in a non-probabilistic multi-context system discards Herbrand interpretations because these worlds have a probability of zero. However probabilistic multi-context systems have to keep these Herbrand interpretations in the solution chain because these worlds can have probabilities associated to them. For this reason a probabilistic solution chain has to provide information about the minimal solution chain in non-probabilistic multi-context systems and the description of the contextual world probability density function. In cases where the probabilistic interpretations of the system do not comply with the probability theory, a probabilistic solution chain cannot be determined.

Roelofsen and Serafini [13] prove that every non-probabilistic multi-context system S has a unique minimal solution chain c_s. Then, according to proposition 1, c_p contains a unique minimal probabilistic entailed chain c_e.

In order to find the probabilistic solution chain c_p and at the same time the minimal probabilistic entailed chain c_e, the following analysis is conducted.

According to formula 4, $\mathscr{B}_k$ denotes a finite set of sentences or formulas in a context k.

$$\mathscr{B}_k = \{(k:\varphi_1)\mu_1, \ldots, (k:\varphi_m)\mu_m\}$$

Given this set of sentences, it can be inferred that: $\eta_k = \{(k:\varphi_1)\mu_1 \wedge, \ldots, \wedge(k:\varphi_m)\mu_m\} = \{(k:\psi)\mu\}$. For a finite set of contexts K for $k = 1, \ldots, m$, then: $c_e = \{\eta_1, \ldots, \eta_m\} = \{(k_1:\psi_1)\mu_1, \ldots, (k_1:\psi_m)\mu_m\}$.

Because this chain is unique in a system and constitutes the probabilistic entailment for system T, it is necessary to introduce an operator that computes this chain.

We can say that $\mathscr{B}_k$, the set of formulas that constitute the knowledge base in context k, is composed of facts and information obtained through bridge rules. For a chain c, if $c \models body(r)$ then $head(r) \in \mathscr{B}_k$.

Assuming that $\mathscr{B}_1$ (for context 1) contains two formulas φ and ψ: $\mathscr{B}_1 = \{(k_1:\varphi)\mu_1, (k_1:\psi)\mu_2\}$

It can be inferred: $\mathscr{N}_1 = \{(k_1:\varphi)\mu_1 \wedge (k_1:\psi)\mu_2\} = \{(k_1:\upsilon)\mu\}$

In order to address the probabilistic consistency, the following Kolmogorov's axiom can be applied:

$$P(A \wedge B) = P(A) + P(B) - P(A \vee B)$$

This axiom can be expressed as:

$$\mathscr{I}_{wp_1}(\varphi \wedge \psi) = \mathscr{I}_{wp_1}(\varphi) + \mathscr{I}_{wp_1}(\psi) - \mathscr{I}_{wp_1}(\varphi \vee \psi) = \mu_1 + \mu_2 - \mathscr{I}_{wp_1}(\varphi \vee \psi)$$

Then: $\mathscr{N}_1 = (k_1:\upsilon)\mu = (k_1:\upsilon)\mathscr{I}_{wp_1}(\varphi \wedge \psi); \mathscr{I}_{WP_1}(\varphi \wedge \psi) = \sum_{W \models \varphi \wedge \psi} WP_1(W) = \mu; \mu \in [0,1]$

As can be seen, being given the probabilities of two formulas is not sufficient to find the probability of their conjunction. Also, the probabilistic interpretation $\mathscr{I}_{wp_1}(\varphi \wedge \psi)$ provides information of the Herbrand interpretations or worlds where $\varphi \wedge \psi$ holds, which is its logical entailment υ. For example, if φ holds in worlds W_2, W_4 and ψ holds in worlds W_1, W_2, W_3 then $\varphi \wedge \psi$ holds in W_2. This means that the logical entailment υ is equal to W_2.

Every probabilistic interpretation generates an equation. For example, if $m = 4$ (four worlds): $\mathscr{I}_{wp_1}(l) = w_1 + w_2 = 0.5$

The first row in Table 1 represents the default constraint and the second row represents the equation $w_1 + w_2 = 0.5$. Table 2 depicts a system of equations, where the goal is to obtain the probabilistic entailment $\mathscr{N}_1$, given φ and ψ. These equations are processed in such way that the probabilistic consistency in the right column of the table can be verified and $\mathscr{N}_1 = (k_1:\upsilon)\mu$ is obtained in the last row. The logical entailment υ is implicit in the equation ($w_1 + w_2 = 0.5$, Table 2), because Herbrand interpretations that hold correspond to a value of one in the equation and Herbrand interpretations that do not hold correspond to a value of zero in the equation. Let this process be the $\mathscr{K}\mathscr{E}$ function.

Table 1:

w_1	w_2	w_3	w_4	$0 \leq \mu \leq 1$
1	1	1	1	1
1	1	0	0	0.5

Table 2:

$w_1\ w_2 \ldots w_m$	$0 \leq \mu \leq 1$
$\mathscr{I}_{WP_1}(\varphi)$	μ_1
$\mathscr{I}_{WP_1}(\psi)$	μ_2
$\mathscr{I}_{WP_1}(\varphi \vee \psi)$	$\mu_1 + \mu_2 - \mathscr{I}_{WP_1}(\varphi \wedge \psi)$
$\mathscr{I}_{WP_1}(\varphi \wedge \psi)$	μ

Definition 8. *The $\mathscr{KE}$ function for $k = 1, \ldots, m$ is of the form :*
$\mathscr{KE}(\mathscr{I}_{WP_k}(\psi)) = \mathscr{I}_{WP_k}(\mathscr{B}_k \wedge \psi) = \mathscr{N}_k$; if a p-chain $c \models body(r)$ then $head(r) \in \mathscr{B}_k$

$\mathscr{B}_k$ represents the knowledge base in a context k. The $\mathscr{KE}$ function can be applied iteratively for many p-formulas. Also, if the same process is applied for all contexts then c_e can be obtained.

Theorem 1. *$\mathscr{KE}$ is monotonic.*

Proof. Given a default constraint in $\mathscr{LE}_k$ and an interpretation φ in $\mathscr{KE}$:

$$(i)\ \mathscr{KE}(\mathscr{I}_{wp_k}(\varphi)) = \mathscr{I}_{wp_k}(\varphi \wedge 1) = \mathscr{I}_{wp_k}(\varphi)$$

Then adding another interpretation ψ in $\mathscr{KE}$: $(ii)\ \mathscr{KE}(\mathscr{I}_{wp_k}(\psi)) = \mathscr{I}_{wp_k}(\varphi \wedge \psi)$.
If $\mathscr{KE}$ is monotonic: whenever $\mathscr{I}_{WP_k}(\psi) = \mathscr{I}_{WP_k}(\varphi)$, $(iii)\ \mathscr{KE}(\mathscr{I}_{wp_k}(\psi)) \leq \mathscr{KE}(\mathscr{I}_{wp_k}(\varphi))$
Replacing (i) and (ii) in (iii): $\mathscr{I}_{wp_k}(\varphi \wedge \psi) \leq \mathscr{I}_{wp_k}(\varphi)$.
A Kolmogorov's axiom states that:
If ϕ logically implies λ then $P(\phi) \leq P(\lambda)$ because $(\varphi \wedge \psi) \subseteq \varphi$ then $(\varphi \wedge \psi) \rightarrow \varphi$
Then if $(\varphi \wedge \psi)$ implies φ : $\mathscr{I}_{WP_k}(\varphi \wedge \psi) \leq \mathscr{I}_{WP_k}(\varphi)$

Because $\mathscr{KE}$ is *monotonic* and sets of all p-labeled chains ($\mathscr{C}$,$\ll$) form a complete lattice, according to the Knaster-Tarsky theorem (cited by [7]), it can be stated that:

Theorem 2. *$\mathscr{KE}$ has a least fixpoint*

This least fixpoint contains c_e or the set of all the probabilistic entailments $\mathscr{N}_k$.

Example 2. *Continuing Example 1 to find c_p and c_e:*
Step 1:
$r_1 :\quad (1 :\sim r)0.5 \quad \leftarrow$

$c^{\perp}$ *denotes the initial p-labeled chain (containing all p-local models with unknown probabilities). Because body(r_1) is empty then* $c^{\perp} \models \text{body}(r_1)$. *That means that head(r_1)$\in \mathscr{B}_1$*
$\mathscr{I}_{wp_1}(\sim r) = w_2 + w_4 = 0.5$

$$c_1 = \left\{ \begin{array}{c} [\{l,r\} : w_1, \{\mathbf{l}, \sim \mathbf{r}\} : \mathbf{w_2}, \{\sim l, r\} : w_3,\ \{\sim \mathbf{l}, \sim \mathbf{r}\} : \mathbf{w_4}]_1, \\ [\{\mathrm{l,c,r}\} : \mathrm{w}_1, \{\mathrm{l,c}, \sim r\} : w_2, \{l, \sim c, r\} : w_3, \{l, \sim c, \sim r\} : w_4, \{\sim l, c, r\} : w_5, \\ \{\sim l, c, \sim r\} : w_6, \{\sim l, \sim c, r\} : w_7, \{\sim l, \sim c, \sim r\} : w_8]_2 \end{array} \right\}$$

Step 2:
$r_2 :\quad (2 : c)0.5 \quad \leftarrow$

Because body(r_2) is empty then $c_1 \models \text{body}(r_2)$. *That means that head(r_2)$\in \mathscr{B}_2$*
$\mathscr{I}_{wp_2}(c) = w_1 + w_2 + w_5 + w_6 = 0.5$

$$c_2 = \left\{ \begin{array}{c} [\{\mathbf{l},\mathbf{r}\}:\mathbf{w_1}, \{\mathbf{l},\sim\mathbf{r}\}:\mathbf{w_2}, \{\sim l, r\}:w_3,\ \{\sim l,\sim r\}:w_4]_{1,} \\ [\{\mathbf{l},\mathbf{c},\mathbf{r}\}:\mathbf{w_1}, \{\mathbf{l},\mathbf{c},\sim\mathbf{r}\}:\mathbf{w_2}, \{l,\sim c, r\}:w_3, \{l,\sim c,\sim r\}:w_4, \{\sim\mathbf{l},\mathbf{c},\mathbf{r}\}:\mathbf{w_5}, \{\sim\mathbf{l},\mathbf{c},\sim\mathbf{r}\}:\mathbf{w_6}, \\ \{\sim l,\sim c, r\}:w_7, \{\sim l,\sim c,\sim r\}:w_8]_2 \end{array} \right\}$$

Step 3:

$r_3:\quad (1:(l\vee r))0.75 \quad \leftarrow (2:(l\vee c\vee r))0.875$

$\mathscr{I}_{wp_2}(l\vee c\vee r) = w_1+w_2+w_3+w_4+w_5+w_6+w_7 = 0.875$

$w_8 = 1-0.875 = 0.125$

$c_2 \models (2:(l\vee c\vee r))0.875$ *then head*(r_3)$\in \mathscr{B}_1$

$\mathscr{I}_{wp_1}(l\vee r) = w_1+w_2+w_3 = 0.75$

Step 4:

$r_4:\quad (2:(l\vee c\vee r))0.875 \quad \leftarrow (1:(l\vee r))0.75$

$\mathscr{I}_{wp_1}(l\vee r) = w_1+w_2+w_3 = 0.75$

w_1	w_2	w_3	w_4	
1	1	1	1	1
0	1	0	1	0.5
1	1	1	1	1.5-w_2-w_4
0	1	0	1	0.5
1	1	1	0	0.75
1	1	1	1	1.25-w_2
0	1	0	0	0.25
1	1	1	0	0.75
1	1	1	0	1-w_2
0	1	0	0	0.25

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	
1	1	1	1	1	1	1	1	1
1	1	0	0	1	1	0	0	0.5
1	1	1	1	1	1	1	1	1.5-w_1-w_2-w_5-w_6
1	1	0	0	1	1	0	0	0.5
1	1	1	1	1	1	1	0	0.87
1	1	1	1	1	1	1	0	1.37-w_1-w_2-w_5-w_6
1	1	0	0	1	1	0	0	0.5
1	1	1	1	1	1	1	0	0.87
1	1	1	1	1	1	1	0	0.87-w_1-w_2-w_5-w_6
1	1	0	0	1	1	0	0	0.5

$c_3 \models (1:(l\vee r))0.75$ *then head*$(r_4) \in \mathscr{B}_2(\mathbf{2}:(\mathbf{l}\vee\mathbf{c}\vee\mathbf{r}))\mathbf{0.875}$

$\mathscr{I}_{wp_2}(l\vee c\vee r) = w_1+w_2+w_3+w_4+w_5+w_6+w_7 = 0.875$

$$c_4 = \left\{ \begin{array}{c} [\{l,r\}:w_1, \{\mathbf{l},\sim\mathbf{r}\}:\mathbf{0.25}, \{\sim l, r\}:w_3,\ \{\sim l,\sim r\}:0.25]_{1,} \\ [\{\mathbf{l},\mathbf{c},\mathbf{r}\}:\mathbf{w_1}, \{\mathbf{l},\mathbf{c},\sim\mathbf{r}\}:\mathbf{w_2}, \{l,\sim c, r\}:w_3, \{l,\sim c,\sim r\}:w_4, \\ \{\sim\mathbf{l},\mathbf{c},\mathbf{r}\}:\mathbf{w_5}, \{\sim\mathbf{l},\mathbf{c},\sim\mathbf{r}\}:\mathbf{w_6}, \{\sim l,\sim c, r\}:w_7, \{\sim l,\sim c,\sim r\}:0.125]_2 \end{array} \right\}$$

$$c_3 = c_4 \ ; c_p = c_4$$

$$c_e = \left\{ \begin{array}{c} [\{\mathbf{l},\sim\mathbf{r}\}:\mathbf{0.25}]_{1,} \\ [(\{\mathbf{l},\mathbf{c},\mathbf{r}\},\{\mathbf{l},\mathbf{c},\sim\mathbf{r}\},\{\sim\mathbf{l},\mathbf{c},\mathbf{r}\},\{\sim\mathbf{l},\mathbf{c},\sim\mathbf{r}\}):\mathbf{0.5}]_2 \end{array} \right\}$$

Then, c_e can be interpreted as: "The probability that there is a ball on the left and not on the right (relative to Mr1) is 0.25" and "The probability that there is a ball in the center (relative to Mr2) is 0.5".

Although the precise connection between contextual ontologies (e.g C-OWL) and p-MCS is not being developed in this paper, we can extend Example 2, creating a mapping between the probabilistic entailment obtained in context 1 and an ontology, using the following bridge rule:

$$(onto: Ball(ball, left))0.25 \qquad \leftarrow \qquad (1: l\wedge\neg r)0.25$$

4 Conclusion

This paper has proposed a theoretical approach to the introduction of uncertainty in multi-context systems. A more expressive semantics has been presented in order to extend the notion of probability in multi-context systems. This additional expressiveness brings new constraints that have to be harmonious and consistent with probability and logic theory. In order to address this situation, a probabilistic logic semantic approach based on the works of [3, 9, 10] has been extended to MCS. Also, a technique that deals with probabilistic inconsistency and the deduction of a minimal entailed chain has been incorporated to the framework.

There are some additional observations worth making about the characteristics of the framework. Firstly, MCS can be embedded in p-MCS assigning to the propositions a probability of one. This characteristic and the deduction of a minimal entailed chain mean MCS can be incorporated into a more general framework. However, there are practical limitations that need to be addressed in future work. For example, the joint probability function or contextual world probability density function has to be specified explicitly in tabular form, which requires exponentially many parameters. This circumstance could be a limitation in practical applications [2]. Nevertheless, this strategy can be used as a theoretical foundation for different approaches [11]. Finally, Bayesian Networks have been successful the last two decades in reducing the complexity of computation of joint probability functions [2]. This suggests that future work should look at incorporating the notion of conditional probabilities into p-MCS through Bayesian Networks.

References

[1] P. Bouquet, F. Giunchiglia, F. van Harmelen, L. Serafini, and H. Stuckentchmidt. C-OWL: Contextalizing ontologies. *Lecture Notes in Computer Science*, 2003.

[2] F. Cozman, R. Haenni, J. Romeijn, F. Russo, G. Wheeler, and J. Williamson. Combining probability and logic. *Journal of Applied Logic*, 2007.

[3] M. Dekhtyar and A. Dekhtyar. Revisiting the semantics of interval probabilistic logic programs. Proceedings of 8th LPNMR, 2005.

[4] F. Giunchiglia and C. Ghidini. Local model semantics, or contextual reasoning = locality+compatibility. Technical report, Instituto Trentino di Cultura, 1997.

[5] J. Hendler and T. Berners-Lee. From the semantic web to social machines: A research challenge for AI on the world wide web. *Artificial Intelligence*, 2010.

[6] J.Bao, J.Tao, D. McGuinness, and P. Smart. Context representation for the semantic web. In Proceedings of the WebSci10, 2010.

[7] J. Lloyd. *Foundations of logic programming*. Springer-Verlag, 1987.

[8] K. Ng and J. Lloyd. Probabilistic reasoning in a classic logic. *Journal of Applied Logic*, 2009.

[9] R. Ng and V. Subrahmanian. Probabilistic logic programming. *Information and Computation*, 1992.

[10] N. Nilsson. Probabilistic logic. *Artificial Intelligence*, 1986.

[11] P. Norvig and S. Russell. *Artificial intelligence: a modern approach*. Prentice Hall/Pearson Education, 2003.

[12] A. Rector and W. Nowlan. The GALEN project. *Computer Methods and Programs in Biomedicine*, 1994.

[13] F. Roelofsen and L. Serafini. Minimal and absent information in contexts. Proceedings of 19th IJCAI, 2005.

[14] S. Schulz, R. Cornet, and K. Spackman. Consolidating SNOMED CT's ontological commitment. *Applied Ontology*, 2011.

[15] L. Serafini and P. Bouquet. Comparing formal theories of context in AI. *Artificial Intelligence*, 2004.