

# *The Probabilistic Description Logic $BALC$*

LEONARD BOTHA and THOMAS MEYER  
*University of Cape Town and CAIR, South Africa*

RAFAEL PEÑALOZA\*  
*University of Milano-Bicocca, Italy*  
(e-mail: [rafael.penaloza@unimib.it](mailto:rafael.penaloza@unimib.it))

*submitted 18 December 2019; revised 8 October 2020; accepted 10 November 2020*

---

## Abstract

Description logics (DLs) are well-known knowledge representation formalisms focused on the representation of terminological knowledge. Due to their first-order semantics, these languages (in their classical form) are not suitable for representing and handling uncertainty. A probabilistic extension of a light-weight DL was recently proposed for dealing with certain knowledge occurring in uncertain contexts. In this paper, we continue that line of research by introducing the Bayesian extension  $BALC$  of the propositionally closed DL  $ALC$ . We present a tableau-based procedure for deciding consistency and adapt it to solve other probabilistic, contextual, and general inferences in this logic. We also show that all these problems remain EXPTIME-complete, the same as reasoning in the underlying classical  $ALC$ .

**KEYWORDS:** knowledge representation and nonmonotonic reasoning, databases and semantic web reasoning

---

## 1 Introduction

Description logics (DLs) (Baader *et al.* 2007) are a family of logic-based knowledge representation formalisms designed to describe the terminological knowledge of an application domain. DLs have been successfully applied to model several domains, with some particularly successful applications coming from the biomedical sciences. This is due to their clear syntax, formal semantics, the existence of efficient reasoners, and their expressivity. However, in their classical form, these logics are not capable of dealing with uncertainty, which is an unavoidable staple in real-world knowledge. To overcome this limitation, several probabilistic extensions of DLs have been suggested in the literature. The landscape of probabilistic extensions of DLs is too large to be covered in detail in this work. These logics differentiate themselves according to their underlying logical formalism, their interpretation of probabilities, and the kind of uncertainty that they are able to express. For a relevant survey, where all these differences are showcased, we refer the interested reader to Lukasiewicz and Straccia (2008). More recent work where probabilistic DLs are discussed can be found in Ceylan and Lukasiewicz (2018) and Gutiérrez-Basulto *et al.* (2017).

\*Part of this work was carried out while this author was at the Free University of Bozen-Bolzano, Italy.

A related probabilistic DL, called  $\mathcal{BEL}$  (Ceylan and Peñaloza 2017), is the Bayesian extension of the light-weight  $\mathcal{EL}$  (Baader et al. 2005). This logic focuses on modeling certain knowledge that holds only in some contexts, together with uncertainty about the current context.  $\mathcal{BEL}$  is based on a *subjective* interpretation of probabilities or, in Halpern’s terminology, it corresponds to a Type II probabilistic logic (Halpern 1990). One advantage of the formalism underlying  $\mathcal{BEL}$  is that it separates the contextual knowledge, which is *de facto* a classical ontology, from the likelihood of observing this context, which can be influenced by external factors. We present a simple example of the importance of contextual knowledge. Consider the knowledge of construction techniques and materials that vary through time. In the context of a modern house, asbestos and lead pipes are not observable, while in some classes of houses, built during the 1970s, we observe both. However, in all contexts we know that asbestos and lead in drinking water have grave health effects. When confronted with a random house, one might not know to which of these contexts it belongs, and by extension whether it is safe to live in, or drink the water that flows through its pipes. Still, construction data may be used to derive the probabilities of these contexts.

To allow for complex probabilistic relationships between the contexts without needing to result to incompatible independence assumptions, their joint probability distribution (JPD) is encoded via a Bayesian network (BN) (Pearl 1985). This logic is closely related to the probabilistic extension of DL-Lite (Artale et al. 2009) proposed previously in d’Amato et al. (2008) but uses a less restrictive semantics which resembles more the open-world assumption from DLs (for a discussion on the differences between the semantics of these logics, see Ceylan and Peñaloza 2017). Another similar proposal is Probabilistic Datalog<sup>±</sup> (Gottlob et al. 2013), with the difference that uncertainty is represented via a Markov Logic Network, instead of a BN. Since the introduction of  $\mathcal{BEL}$ , the main notions behind it have been generalized to arbitrary ontology languages (Ceylan 2018). However, it has also been shown that efficient and complexity-optimal reasoning methods can only be achieved by studying the properties of each underlying ontology language (Ceylan and Peñaloza 2017). Finally, another family of probabilistic DLs using Type II semantics was proposed in Gutiérrez-Basulto et al. (2017), which we refer to as Prob-DLs. The biggest difference between Bayesian DLs and Prob-DLs is that the latter models uncertain concepts, while the former models uncertain knowledge. For example, Prob-DLs can refer to the class of all individuals having a high probability of a disease infection, but cannot express that an entailment holds with a given probability.

In this paper, we continue with that line of research and study the Bayesian extension of the propositionally closed DL  $\mathcal{ALC}$ . As our main result, we present an algorithm, based on a glass-box modification of the classical tableaux method for reasoning in  $\mathcal{ALC}$ . Our algorithm is able to derive those contexts which are needed to determine the (in)consistency of a  $\mathcal{BALC}$  knowledge base (KB). Using this algorithm, we then describe an effective method for deciding consistency of a  $\mathcal{BALC}$  KB. We also provide a tight EXPTIME complexity bound for this problem.

This is followed by a study of several crisp and probabilistic variants of the standard DL decision problems, namely concept satisfiability, subsumption, and instance checking. Interestingly, our work shows that all our problems can be reduced to some basic computations over a context describing inconsistency, and hence are EXPTIME-complete as well. These complexity bounds are not completely surprising, given the high complexity of the classical  $\mathcal{ALC}$ . However, our tableaux-based algorithm has the potential to

behave better in practical scenarios. This work details and deepens the results that have previously been presented in Botha (2018) and Botha *et al.* (2018; 2019).

## 2 Preliminaries

We start by providing a brief introduction to BNs and the DL  $\mathcal{ALC}$ , which form the basis for the probabilistic DL  $\mathcal{BALC}$ . For a more detailed presentation of these topics, we refer the interested reader to Darwiche (2009) and Baader *et al.* (2017).

### 2.1 Bayesian networks

Bayesian networks (BNs) are graphical models, which are used for representing the JPD of several discrete random variables in a compact manner (Pearl 1985). Before introducing these models formally, we need a few definitions.

Given a random variable  $X$ , let  $\text{val}(X)$  denote the set of values that  $X$  can take. For the scope of this paper, we consider only random variables  $X$  such that  $\text{val}(X)$  is finite. For  $x \in \text{val}(X)$ ,  $X = x$  denotes the *valuation* of  $X$  taking the value  $x$ . This notation is extended to sets of variables in the obvious way. Given a set of random variables  $V$ , a *world*  $\omega$  is a set of valuations containing exactly one valuation for every random variable  $X \in V$ . In other words, a world specifies an exact instantiation of all the variables in  $V$ .

A *V-literal* is an ordered pair of the form  $(X_i, x)$ , where  $X_i \in V$  and  $x \in \text{val}(X_i)$ .  $V$ -literals are similar to valuations, but the syntactic difference is introduced to emphasize their difference in use, as will become clear later.  $V$ -literals are a generalization of Boolean literals, which are typically denoted as  $x$  or  $\neg x$  for the Boolean random variable  $X$ . For simplicity, and following this connection, in this paper we will often use the notation  $X$  for  $(X, T)$  and  $\neg X$  for  $(X, F)$ . A *V-context* is any set of  $V$ -literals. A  $V$ -context is *consistent* if it contains at most one literal for each random variable. We will often also call  $V$ -contexts *primitive contexts*.

A Bayesian network (BN) is defined as a pair  $\mathcal{B} = (G, \Theta)$  where  $G = (V, E)$  is a directed acyclic graph (DAG) and  $\Theta$  is a set of conditional probability distributions for every variable  $X \in V$  given its parents  $\pi(X)$  on the DAG  $G$ ; more precisely, this set has the form  $\Theta = \{P(X = x | \pi(X) = \vec{x}') \mid X \in V\}$ .

The BN  $\mathcal{B}$  specifies a full JPD for the variables in  $V$  by considering independence assumptions depicted in the graph  $G$ , namely every variable is (conditionally) independent of all its *non-descendants* given its parents. Under this assumption, it is easy to see that the JPD of  $V$  can be computed through the chain rule

$$P(\vec{X} = \vec{x}) = \prod_{X_i \in V} P(X_i = x_i \mid \pi(X_i) = \vec{x}_j'),$$

that is, the probability of a world is obtained by multiplying the conditional probabilities of the valuations found in the tables. We let  $P_{\mathcal{B}}$  denote the JPD defined by the BN  $\mathcal{B}$ .

#### Example 1

Figure 1 depicts a BN with four random variables denoting the likelihood of different characteristics of a construction:  $X$  stands for a post-1986 building,  $Y$  for a renovated building,  $Z$  for the presence of lead pipes, and  $W$  for the safety of drinking water. The tables attached to each node are the conditional probability distribution of that node

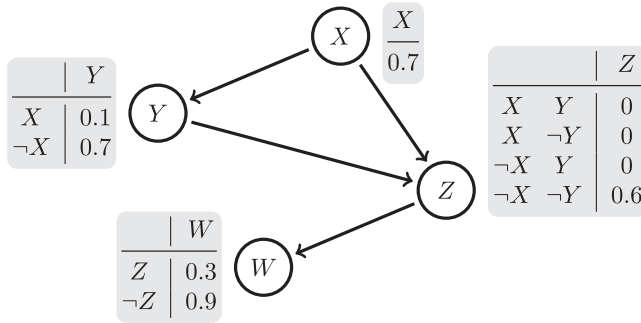


Fig. 1. A Bayesian network with four Boolean variables.

given its parents. Hence, the BN expresses that a post-1986 building has only probability 0.1 of being renovated.

The DAG expresses the conditional independence of  $W$  and  $Y$  given  $Z$ . That is, once that we observe the value of  $Z$  (if we know whether a house has lead pipes or not) then the probability of  $W$  is not affected by the knowledge of the renovation status of a house.

Through the chain rule, we can derive, for example,

$$P(\neg X, \neg Y, Z, \neg W) = 0.3 \cdot 0.3 \cdot 0.6 \cdot 0.7 = 0.0378, \quad \text{and}$$

$$P(X, Y, Z, W) = 0.7 \cdot 0.1 \cdot 0 \cdot 0.3 = 0.$$

That is, it is very unlikely to find an old, non-renovated house, with lead pipes, and unsafe water, and renovated post-1986 houses with safe water cannot have lead pipes. Note that to express the full JPD of these four variables directly, we would need a table with 16 rows.

## 2.2 The DL $\mathcal{ALC}$

$\mathcal{ALC}$  is the smallest propositionally closed DL (Baader et al. 2007; Schmidt-Schauß and Smolka 1991). As with all DLs, its fundamental notions are those of *concepts* which correspond to unary predicates of first-order logic, and *roles* corresponding to binary predicates. Formally, given the mutually disjoint sets  $N_I, N_C$ , and  $N_R$  of *individual*, *concept*, and *role names*, respectively, the class of  $\mathcal{ALC}$  concepts is built through the grammar rule

$$C ::= A \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \exists r.C \mid \forall r.C,$$

where  $A \in N_C$  and  $r \in N_R$ .

The knowledge from an application domain is represented through a set of axioms, which restrict the possible interpretations of concepts and roles. In  $\mathcal{ALC}$ , *axioms* are either *general concept inclusions* (GCIs) of the form  $C \sqsubseteq D$ , *concept assertions*  $C(a)$ , or *role assertions*  $r(a, b)$  where  $a, b \in N_I, r \in N_R$ , and  $C, D$  are concepts. An *ontology* is a finite set of axioms. As customary in DLs, we sometimes partition an ontology into the *TBox*  $\mathcal{T}$  composed exclusively of GCIs, and the *ABox*  $\mathcal{A}$  containing all concept and role assertions, when it is relevant which kind of axiom is being used.

The semantics of  $\mathcal{ALC}$  is defined by interpretations akin to the first-order semantics. An *interpretation* is a pair of the form  $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$  where  $\Delta^{\mathcal{I}}$  is a non-empty set called

the domain and  $\cdot^{\mathcal{I}}$  is the interpretation function that maps every  $a \in N_I$  to an element  $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$ , every  $A \in N_C$  to a set  $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$  and every  $r \in N_R$  to a binary relation  $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ . This interpretation function is inductively extended to arbitrary concepts by defining for any two concepts  $C, D$ :

- $(\neg C)^{\mathcal{I}} := \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$ ,
- $(C \sqcap D)^{\mathcal{I}} := C^{\mathcal{I}} \cap D^{\mathcal{I}}$ ,
- $(C \sqcup D)^{\mathcal{I}} := C^{\mathcal{I}} \cup D^{\mathcal{I}}$ ,
- $(\exists r.C)^{\mathcal{I}} := \{\delta \in \Delta^{\mathcal{I}} \mid \exists \eta \in C^{\mathcal{I}}. (\delta, \eta) \in r^{\mathcal{I}}\}$ , and
- $(\forall r.C)^{\mathcal{I}} := \{\delta \in \Delta^{\mathcal{I}} \mid \forall \eta \in \Delta^{\mathcal{I}}. (\delta, \eta) \in r^{\mathcal{I}} \Rightarrow \eta \in C^{\mathcal{I}}\}$ .

The interpretation  $\mathcal{I}$  satisfies the GCI  $C \sqsubseteq D$  iff  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ ; the concept assertion  $C(a)$  iff  $a^{\mathcal{I}} \in C^{\mathcal{I}}$ ; and the role assertion  $r(a, b)$  iff  $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}}$ . We denote as  $\mathcal{I} \models \alpha$  if  $\mathcal{I}$  satisfies the axiom  $\alpha$ .  $\mathcal{I}$  is a model of the ontology  $\mathcal{O}$  (denoted by  $\mathcal{I} \models \mathcal{O}$ ) iff it satisfies all axioms in  $\mathcal{O}$ .

An important abbreviation in  $\mathcal{ALC}$  is the bottom concept  $\perp := A \sqcap \neg A$ , where  $A$  is any concept name. Clearly, this concept stands for a contradiction and, for any interpretation  $\mathcal{I}$ ,  $\perp^{\mathcal{I}} = \emptyset$ . Similarly, the top concept  $\top := A \sqcup \neg A$  stands for a tautology and  $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$  holds for every interpretation  $\mathcal{I}$ .

The basic reasoning task in  $\mathcal{ALC}$  consists in deciding whether a given ontology  $\mathcal{O}$  is consistent, that is, whether there exists a model  $\mathcal{I}$  of  $\mathcal{O}$ . This problem is known to be EXPTIME-complete (Schild 1991; Donini and Massacci 2000). Other reasoning problems, such as subsumption and instance checking, can be polynomially reduced to this one and hence preserve the same complexity.

*Example 2*

In  $\mathcal{ALC}$ , it is possible to express the notion that water pipes do not contain lead through the GCI  $\text{WaterPipe} \sqsubseteq \forall \text{contains}. \neg \text{Lead}$ . In addition, we can express that the object `pipe1` is a water pipe ( $\text{WaterPipe}(\text{pipe1})$ ), that `pipe1` contains `substance1` by the assertion  $\text{contains}(\text{pipe1}, \text{substance1})$ , and that `substance1` is in fact lead ( $\text{Lead}(\text{substance1})$ ). Note that the ontology containing all four axioms is inconsistent.

### 3 $\mathcal{BALC}$

We now introduce the probabilistic DL  $\mathcal{BALC}$ , which combines BNs, to compactly express JPDs, and  $\mathcal{ALC}$  to express background knowledge. In addition,  $\mathcal{BALC}$  can express logical (as opposed to probabilistic) dependencies between axioms. More precisely,  $\mathcal{BALC}$  axioms are required to hold only in some (possibly uncertain) contexts, which are expressed through annotations. The uncertainty of these contexts is expressed by the BN. These ideas are formalized next.

*Definition 3 (KB)*

Let  $V$  be a finite set of discrete random variables. A  $V$ -restricted axiom ( $V$ -axiom) is an expression of the form  $\alpha^\kappa$ , where  $\alpha$  is an  $\mathcal{ALC}$  axiom and  $\kappa$  is a  $V$ -context. A  $V$ -ontology is a finite set of  $V$ -axioms. A  $\mathcal{BALC}$  KB over  $V$  is a pair  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  where  $\mathcal{B}$  is a BN over  $V$ , and  $\mathcal{O}$  is a  $V$ -ontology.

Note that the contexts labeling the axioms, and the variables of the BN from a  $\mathcal{BALC}$  KB both come from the same set of discrete random variables  $V$ .

To define the semantics of  $\mathcal{BALC}$ , we extend the notion of an interpretation from  $\mathcal{ALC}$  to also take contexts into account. The probabilities expressed by the BN are associated with the axioms following the multiple-world approach.

*Definition 4 (V-interpretation)*

Let  $V$  be a finite set of discrete random variables. A  $V$ -interpretation is a tuple of the form  $\mathcal{V} = (\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}}, v^{\mathcal{V}})$ , where  $(\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}})$  is an  $\mathcal{ALC}$  interpretation and  $v^{\mathcal{V}}$  is a valuation function  $v^{\mathcal{V}} : V \rightarrow \cup_{X \in V} \text{val}(X)$  which maps every  $X \in V$  to some  $v^{\mathcal{V}}(X) \in \text{val}(X)$ .

Given a valuation function  $v^{\mathcal{V}}$ , a Bayesian world  $\omega$ , and a context  $\kappa$ , we use the notation  $v^{\mathcal{V}} = \omega$  to express that  $v^{\mathcal{V}}$  assigns to each random variable the same value as it has in  $\omega$ ;  $v^{\mathcal{V}} \models \kappa$  to state that  $v^{\mathcal{V}}(X) = x$  for all  $(X, x) \in \kappa$ ; and  $\omega \models \kappa$  when there is  $\omega = v^{\mathcal{V}}$  such that  $v^{\mathcal{V}} \models \kappa$ .

*Definition 5 (Model)*

The  $V$ -interpretation  $\mathcal{V} = (\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}}, v^{\mathcal{V}})$  is a *model* of the  $V$ -axiom  $\alpha^{\kappa}$  (denoted as  $\mathcal{V} \models \alpha^{\kappa}$ ) iff (i)  $v^{\mathcal{V}} \not\models \kappa$ , or (ii)  $(\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}})$  satisfies  $\alpha$ .  $\mathcal{V}$  is a model of the ontology  $\mathcal{O}$  iff it is a model of all axioms in  $\mathcal{O}$ . In this case, we denote it by  $\mathcal{V} \models \mathcal{O}$ .

It is important to notice at this point that  $\mathcal{BALC}$  is a generalization of the classical DL  $\mathcal{ALC}$ . Recall that a  $V$ -context is a set of  $V$ -literals, that is, of pairs  $(X, x)$ . In particular, the empty set of literals  $\emptyset$  is also a  $V$ -context. By definition, every valuation function  $v^{\mathcal{V}}$  is such that  $v^{\mathcal{V}} \models \emptyset$ . This means that the  $\mathcal{ALC}$  axiom  $\alpha$  is equivalent to the  $V$ -axiom  $\alpha^{\emptyset}$ . As a whole, every  $\mathcal{ALC}$  ontology can be seen as a  $V$ -ontology by simply associating the empty context with all its axioms. The original  $\mathcal{ALC}$  ontology and the  $V$ -ontology constructed this way have the same class of models, except that the latter allows for any arbitrary valuation function in its third component. For brevity, and the reasons just described, for the rest of this paper we will abbreviate axioms of the form  $\alpha^{\emptyset}$  simply as  $\alpha$ . When it is clear from the context, we will also omit the  $V$  prefix and refer only to, for example, contexts, GCIs, or ontologies.

A  $V$ -interpretation – more precisely, its valuation function – refers to only one possible world of the random variables in  $V$ .  $\mathcal{BALC}$  KBs, on the other hand, have information about the uncertainty of being in one world or another. In the multiple-world semantics, probabilistic interpretations combine multiple  $V$ -interpretations and the probability distribution described by the BN to give information about the uncertainty of the axioms, and their consequences.

*Definition 6 (Probabilistic model)*

A *probabilistic interpretation* is a pair of the form  $\mathcal{P} = (\mathcal{J}, \mathcal{P}_{\mathcal{J}})$ , where  $\mathcal{J}$  is a finite set of  $V$ -interpretations and  $\mathcal{P}_{\mathcal{J}}$  is a probability distribution over  $\mathcal{J}$  such that  $\mathcal{P}_{\mathcal{J}}(\mathcal{V}) > 0$  for all  $\mathcal{V} \in \mathcal{J}$ . The probabilistic interpretation  $\mathcal{P}$  is a *model* of the axiom  $\alpha^{\kappa}$  (denoted by  $\mathcal{P} \models \alpha^{\kappa}$ ) iff every  $\mathcal{V} \in \mathcal{J}$  is a model of  $\alpha^{\kappa}$ .  $\mathcal{P}$  is a *model* of the ontology  $\mathcal{O}$  iff every  $\mathcal{V} \in \mathcal{J}$  is a model of  $\mathcal{O}$ .

We say that the probability distribution  $\mathcal{P}_{\mathcal{J}}$  is *consistent* with the BN  $\mathcal{B}$  if for every possible world  $\omega$  of the variables in  $V$  it holds that

$$\sum_{\mathcal{V} \in \mathcal{J}, v^{\mathcal{V}} = \omega} P_{\mathcal{J}}(\mathcal{V}) = P_{\mathcal{B}}(\omega).$$

Recall that  $P_{\mathcal{B}}$  denotes the JPD defined by the BN  $\mathcal{B}$ . The probabilistic interpretation  $\mathcal{P}$  is a *model* of the KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  iff it is a (probabilistic) model of  $\mathcal{O}$ , and is consistent with  $\mathcal{B}$ .

In other words, a probabilistic interpretation describes a class of possible worlds, in which the knowledge of  $\mathcal{O}$  is interpreted, and each world is associated with a probability. For this interpretation to be a model of the KB, it has to satisfy the constraints it requires, that is, the ontological knowledge must be satisfied in each world, and the probabilities should be coherent with the distribution from the BN.

*Example 7*

Consider the KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  where  $\mathcal{B}$  is the BN depicted in Figure 1, and  $\mathcal{O}$  is the ontology

$$\mathcal{O} := \{ \begin{array}{ll} \text{Pipe} \sqcap \exists \text{contains.Lead} \sqsubseteq \text{LeadPipe}^{\emptyset}, & \\ \text{Pipe} \sqsubseteq \forall \text{contains.}\neg \text{Lead}^X, & \text{Pipe} \sqsubseteq \forall \text{contains.}\neg \text{Lead}^Y, \\ \text{Pipe} \sqsubseteq \exists \text{contains.Lead}^Z, & \text{Water} \sqcap \exists \text{hasAlkalinity.Low} \sqsubseteq \neg \text{Drinkable}^Z, \\ \text{Water} \sqsubseteq \text{Drinkable}^W, & \text{Water} \sqsubseteq \neg \text{Drinkable}^{\neg W} \}. \end{array}$$

The first axiom provides a partial definition, stating that every pipe which contains lead is a lead pipe. Note that it is labeled with the empty context, which means that it always holds. The two axioms in the second row express that pipes in post-1986 (i.e., within the context  $X$ ) and in renovated buildings (in context  $Y$ ) do not contain lead.

The axioms in the third row refer exclusively to the context of lead pipes ( $Z$ ). In this case, our knowledge is that pipes do contain lead, and that water with low alkalinity is not drinkable, as it absorbs the lead from the pipes it travels on. Notice that the axioms in the second row contradict  $\text{Pipe} \sqsubseteq \exists \text{contains.Lead}$  appearing in the third row. This is not a problem because they are required to hold in different contexts. Indeed, we can observe from Figure 1 that any context which makes  $Z$ , and either  $X$  or  $Y$  true must have probability 0. This means that no probabilistic model can contain a world that satisfies these conditions.

Finally, the last two axioms state that water at a building is drinkable iff we are in the context of drinkable water ( $W$ ). Note that it is important to provide both axioms in order to guarantee an equivalence. Suppose that we had not included the last axiom (i.e.,  $\text{Water} \sqsubseteq \neg \text{Drinkable}^{\neg W}$ ). Then, by our semantics, we could still produce a  $V$ -interpretation  $\mathcal{V}$  whose valuation satisfies  $\neg W$  but such that  $\text{Water}^{\mathcal{V}} \cap \text{Drinkable}^{\mathcal{V}} \neq \emptyset$ . This can only be avoided by the introduction of the last axiom.

It is a simple exercise to verify that this ontology has a model. The smallest probabilistic model of the KB  $\mathcal{K}$  contains 10  $V$ -interpretations, one for each world with positive probability.

Recall that  $V$ -contexts are also called primitive contexts. This is to emphasize the difference with what we will, from now on, call complex contexts. Formally, a *complex context*  $\phi$  is a finite non-empty set of primitive contexts. Note that primitive contexts can be seen as complex ones, for example, the primitive context  $\kappa$  corresponds to the complex context  $\{\kappa\}$ , and we will use them interchangeably when there is no risk of ambiguity.



Given a valuation function  $v^{\mathcal{V}}$  and a complex context  $\phi = \{\alpha_1, \dots, \alpha_n\}$ , we say that  $v^{\mathcal{V}}$  satisfies  $\phi$  (written as  $v^{\mathcal{V}} \models \phi$ ) iff  $v^{\mathcal{V}}$  satisfies *at least one*  $\alpha_i \in \phi$ , in particular  $v^{\mathcal{V}} \models \kappa$  iff  $v^{\mathcal{V}} \models \{\kappa\}$ . Thus, for the rest of the paper we assume that all contexts are in complex form unless explicitly stated otherwise. We emphasize here that for a valuation function to satisfy a complex context, it suffices that it satisfies one of the simple contexts in it. Intuitively, a simple context is a conjunction of valuations, while a complex context is a disjunction of simple contexts. However, one should not forget that the variables appearing in these contexts can potentially take more than two different values.

*Definition 8*

Given complex contexts  $\phi = \{\kappa_1, \dots, \kappa_n\}$  and  $\psi = \{\lambda_1, \dots, \lambda_m\}$ , we define the operations

$$\begin{aligned} \phi \vee \psi &:= \phi \cup \psi, & \text{and} \\ \phi \wedge \psi &:= \bigcup_{\kappa \in \phi, \lambda \in \psi} \{\kappa \cup \lambda\} = \{\kappa \cup \lambda \mid \kappa \in \phi, \lambda \in \psi\}. \end{aligned}$$

These operations generalize propositional disjunction ( $\vee$ ) and propositional conjunction ( $\wedge$ ), where disjunction has the property that either one of the two contexts holds and conjunction requires that both hold.

Given two complex contexts  $\phi$  and  $\psi$ , we say that  $\phi$  entails  $\psi$  (denoted by  $\phi \models \psi$ ) iff for all  $v^{\mathcal{V}}$  such that  $v^{\mathcal{V}} \models \phi$  it follows that  $v^{\mathcal{V}} \models \psi$ . Alternatively  $\phi \models \psi$  holds iff for all Bayesian worlds  $\omega$  such that  $\omega \models \phi$  it follows that  $\omega \models \psi$ . It is easy to see that for all worlds  $\omega$  and complex contexts  $\phi, \psi$  it holds that (i)  $\omega \models \phi \vee \psi$  iff  $\omega \models \phi$  or  $\omega \models \psi$ , and (ii)  $\omega \models \phi \wedge \psi$  iff  $\omega \models \phi$  and  $\omega \models \psi$ . Two important special complex contexts are top ( $\top$ ) and bottom ( $\perp$ ), which are satisfied by all or no world, respectively. If there are  $n$  consistent primitive contexts  $\kappa_1, \dots, \kappa_n$  and  $\kappa$  is an inconsistent context, these are defined as  $\top := \{\kappa_1, \dots, \kappa_n\}$  and  $\perp := \kappa$ .

This concludes the definition of the relevant components of the logic *BALC*. In the next section, we study the problem of consistency of a *BALC* KB, and its relation to other reasoning problems.

## 4 Consistency

As for *ALC*, the most basic decision problem one can consider in *BALC* is *consistency*. Formally, this problem consists of deciding whether a given *BALC* KB  $\mathcal{K}$  has a probabilistic model or not. To deal with this problem, it is convenient to consider the classical *ALC* ontologies that should hold at each specific world, which we call the restriction.

*Definition 9 (restriction)*

Given the *BALC* KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  and the world  $\omega$ , the *restriction* of  $\mathcal{O}$  to  $\omega$  is the *ALC* ontology defined by

$$\mathcal{O}_\omega := \{\alpha \mid \alpha^\kappa \in \mathcal{O}, \omega \models \kappa\}.$$

Recall that a probabilistic model  $\mathcal{P} = (\mathcal{J}, \mathcal{P}_{\mathcal{J}})$  of  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  is a class of classical interpretations associated with worlds  $(\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}}, \omega)$ , where each is a model of the *V*-ontology  $\mathcal{O}$ . For such an interpretation to be a model of  $\mathcal{O}$ , it must hold that, for each axiom  $\alpha^\kappa \in \mathcal{O}$  such that  $\omega \models \kappa$ ,  $(\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}}) \models \alpha$  (see Definition 5). This condition is equivalent to stating that  $(\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}})$  is a classical model of  $\mathcal{O}_\omega$ .



*Example 10*

Consider the KB  $\mathcal{K}$  from Example 7, and the two worlds  $\omega_1 = \{X, \neg Y, \neg Z, W\}$  and  $\omega_2 = \{X, \neg Y, Z, W\}$ . We then have

$$\begin{aligned} \mathcal{O}_{\omega_1} := \{ & \text{Pipe} \sqcap \exists \text{contains.Lead} \sqsubseteq \text{LeadPipe}, \text{Pipe} \sqsubseteq \forall \text{contains.}\neg \text{Lead}, \\ & \text{Water} \sqsubseteq \text{Drinkable}\}, \end{aligned} \quad \text{and}$$

$$\begin{aligned} \mathcal{O}_{\omega_2} := \{ & \text{Pipe} \sqcap \exists \text{contains.Lead} \sqsubseteq \text{LeadPipe}, \text{Pipe} \sqsubseteq \forall \text{contains.}\neg \text{Lead}, \\ & \text{Pipe} \sqsubseteq \exists \text{contains.Lead}, \text{Water} \sqcap \exists \text{hasAlkalinity.Low} \sqsubseteq \neg \text{Drinkable}, \\ & \text{Water} \sqsubseteq \text{Drinkable}\}. \end{aligned}$$

Note that both ontologies are consistent, but  $\mathcal{O}_{\omega_2}$  can only be satisfied by interpretations  $\mathcal{I}$  such that  $\text{Pipe}^{\mathcal{I}} = \emptyset$ .

In addition to satisfying the restrictions, the probability distribution  $\mathcal{P}_{\mathcal{J}}$  of a model must be consistent with the BN  $\mathcal{B}$ . This means that the probabilities of the interpretations associated with the world  $\omega$  must add to  $P_{\mathcal{B}}(\omega)$ . Using this insight, we obtain the following result.

*Theorem 11*

The  $\mathcal{BALC}$  KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  is consistent iff for every world  $\omega$  with  $P_{\mathcal{B}}(\omega) > 0$  it holds that  $\mathcal{O}_{\omega}$  is consistent.

*Proof*

Suppose first that for every world  $\omega$  such that  $P_{\mathcal{B}}(\omega) > 0$  the ontology  $\mathcal{O}_{\omega}$  is consistent. We build a probabilistic model  $\mathcal{P}$  of  $\mathcal{K}$  as follows. Given a world  $\omega$  with  $P_{\mathcal{B}}(\omega) > 0$ , let  $\mathcal{I}_{\omega} = (\Delta^{\mathcal{I}_{\omega}}, \cdot^{\mathcal{I}_{\omega}})$  be an arbitrary (classical) model of  $\mathcal{O}_{\omega}$ . We construct the  $V$ -interpretation  $\mathcal{V}_{\omega} = (\Delta^{\mathcal{I}_{\omega}}, \cdot^{\mathcal{I}_{\omega}}, \omega)$ , that is, the same interpretation  $\mathcal{I}_{\omega}$ , but associated with the world  $\omega$ . We claim that  $\mathcal{V}_{\omega} \models \mathcal{O}$ . Indeed, take an arbitrary  $\alpha^{\kappa} \in \mathcal{O}$ . If  $\omega \not\models \kappa$ , then  $\mathcal{V}_{\omega} \models \alpha^{\kappa}$  trivially. Otherwise,  $\alpha \in \mathcal{O}_{\omega}$  and hence (since  $\mathcal{I}_{\omega} \models \mathcal{O}_{\omega}$ )  $(\Delta^{\mathcal{I}_{\omega}}, \cdot^{\mathcal{I}_{\omega}}) \models \alpha$ . Let now  $\Omega$  be the set of all worlds  $\omega$  such that  $P_{\mathcal{B}}(\omega) > 0$ . We define the set of  $V$ -interpretations  $\mathcal{J}_{\Omega} := \{\mathcal{V}_{\omega} \mid \omega \in \Omega\}$ , and the probability distribution  $\mathcal{P}_{\mathcal{J}_{\Omega}}$  that sets  $\mathcal{P}_{\mathcal{J}_{\Omega}}(\mathcal{V}_{\omega}) = P_{\mathcal{B}}(\omega)$  for every  $\omega \in \Omega$ . The probabilistic interpretation  $\mathcal{P} = (\mathcal{J}_{\Omega}, \mathcal{P}_{\mathcal{J}_{\Omega}})$  is a model of  $\mathcal{K}$ .

Conversely, suppose that  $\mathcal{K}$  is consistent, and let  $\mathcal{P} = (\mathcal{J}, \mathcal{P}_{\mathcal{J}})$  be a model of  $\mathcal{K}$ . Given a world  $\omega$  such that  $P_{\mathcal{B}}(\omega) > 0$ , we know that  $\sum_{\mathcal{V} \in \mathcal{J}, v^{\mathcal{V}} = \omega} P_{\mathcal{J}}(\mathcal{V}) = P_{\mathcal{B}}(\omega)$ . In particular, this means that there must exist some  $\mathcal{V} = (\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}}, v^{\mathcal{V}}) \in \mathcal{J}$  such that  $v^{\mathcal{V}} = \omega$ .  $\mathcal{P} \models \mathcal{K}$  means that  $\mathcal{V} \models \mathcal{O}$  and hence  $(\Delta^{\mathcal{V}}, \cdot^{\mathcal{V}}) \models \mathcal{O}_{\omega}$ . Thus,  $\mathcal{O}_{\omega}$  is consistent.  $\square$

Based on this result, we can derive a process for deciding consistency that provides a tight complexity bound for this problem.

*Corollary 12*

$\mathcal{BALC}$  KB consistency is EXPTIME-complete.

*Proof*

Recall that consistency checking in  $\mathcal{ALC}$  is EXPTIME-complete, and that  $\mathcal{BALC}$  is a generalization of  $\mathcal{ALC}$ . This yields EXPTIME-hardness for  $\mathcal{BALC}$  consistency. To obtain the upper bound, note that there are exponentially many worlds  $\omega$ . For each of them, we

can check (classical) consistency of  $\mathcal{O}_\omega$  (in exponential time) and that  $P_{\mathcal{B}}(\omega) > 0$ , which is linear in the size of  $\mathcal{B}$  by the chain rule.  $\square$

The algorithm described in the proof of this corollary is optimal in terms of *worst-case* complexity, but it also runs in exponential time in the *best case*. Indeed, it enumerates all the (exponentially many) Bayesian worlds, before being able to make any decision. In practice, it is infeasible to use an algorithm that requires exponential time on every instance. For that reason, we present a new algorithm based on the tableau method originally developed for  $\mathcal{ALC}$ . To describe this algorithm, we need to introduce some additional notation.

We denote the context that describes all worlds  $\omega$  such that  $\mathcal{O}_\omega$  is inconsistent as  $\phi_{\mathcal{K}}^\perp$ . That is, the context  $\phi_{\mathcal{K}}^\perp$  is such that  $\omega \models \phi_{\mathcal{K}}^\perp$  iff  $\mathcal{O}_\omega$  is inconsistent. Moreover,  $\phi_{\mathcal{B}}$  is a context such that  $\omega \models \phi_{\mathcal{B}}$  iff  $P(\omega) = 0$ . Theorem 11 states that  $\mathcal{K}$  is inconsistent whenever there is a world that models both  $\phi_{\mathcal{K}}^\perp$  and  $\neg\phi_{\mathcal{B}}$ . This is formalized in the following result.

*Theorem 13*

Given the  $\mathcal{BALC}$  KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$ , let  $\phi_{\mathcal{K}}^\perp$  and  $\phi_{\mathcal{B}}$  be the contexts described above. Then  $\mathcal{K}$  is inconsistent iff  $\phi_{\mathcal{K}}^\perp \wedge \neg\phi_{\mathcal{B}}$  is satisfiable.

*Proof*

$\mathcal{K}$  is inconsistent iff there is a world  $\omega$  such that  $P_{\mathcal{B}}(\omega) > 0$  and  $\mathcal{O}_\omega$  is inconsistent (Theorem 11) iff  $\omega \not\models \phi_{\mathcal{B}}$  and  $\omega \models \phi_{\mathcal{K}}^\perp$  iff  $\omega \models \neg\phi_{\mathcal{B}} \wedge \phi_{\mathcal{K}}^\perp$  iff  $\neg\phi_{\mathcal{B}} \wedge \phi_{\mathcal{K}}^\perp$  is satisfiable.  $\square$

*Example 14*

Consider once again the KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  from Example 7, and notice that  $\mathcal{O}_\omega$  is consistent for all worlds  $\omega$ . Hence  $\phi_{\mathcal{K}}^\perp = \perp$ . In particular, this means that  $\phi_{\mathcal{K}}^\perp \wedge \neg\phi_{\mathcal{B}}$  is unsatisfiable, and hence  $\mathcal{K}$  is consistent.

Let now  $\mathcal{K}' = (\mathcal{O}', \mathcal{B})$ , where  $\mathcal{O}' = \mathcal{O} \cup \{\text{Pipe}(\text{pipe1})^\emptyset\}$ . We can see that for any world  $\omega$  which entails any of the contexts  $X, Z$  or  $Y, Z$ ,  $\mathcal{O}'_\omega$  is inconsistent. Hence, we have that  $\phi_{\mathcal{K}'}^\perp = \{\{X, Z\}, \{Y, Z\}\}$ . Looking at the probability tables from  $\mathcal{B}$ , we also see that  $\phi_{\mathcal{B}} = \{\{X, Z\}, \{Y, Z\}\}$ , that is, any world that satisfies  $\{X, Z\}$  or  $\{Y, Z\}$  must have probability 0. Thus,  $\phi_{\mathcal{K}'}^\perp \wedge \neg\phi_{\mathcal{B}} = \perp$  and  $\mathcal{K}'$  is consistent as well.

To decide consistency, it then suffices to find a method which is capable of deriving the contexts  $\phi_{\mathcal{K}}^\perp$  and  $\phi_{\mathcal{B}}$ . For the former, we present a variant of the glass-box approach for so-called axiom pinpointing (Lee et al. 2006; Meyer et al. 2006; Baader and Peñaloza 2010), which is originally based on the ideas from (Baader and Hollunder 1995). This approach modifies the standard tableaux algorithm for  $\mathcal{ALC}$  – which tries to construct a model by decomposing complex constraints into simpler ones – to additionally keep track of the contexts in which the derived elements from the tableau hold. In a nutshell, tableau algorithms apply different decomposition rules to make implicit constraints explicit. For axiom pinpointing, whenever a rule application requires the use of an axiom from the ontology, this fact is registered as part of a propositional formula. In our case, we need a context, rather than a propositional formula, to take care of the multiple values that the random variables take. Although the underlying idea remains essentially unchanged, the handling of these multiple values requires additional technicalities. Before we look at this algorithm in detail, it is important to keep in mind that the goal is to detect all

$\sqcap$ -rule	if	$(C_1 \sqcap C_2)(a)^\phi \in \mathcal{A}$ , and either $C_1(a)^\phi$ or $C_2(a)^\phi$ is $\mathcal{A}$ -insertable
	then	$\mathcal{A}' := (\mathcal{A} \oplus C_1(a)^\phi) \oplus C_2(a)^\phi$
$\sqcup$ -rule	if	$(C_1 \sqcup C_2)(x)^\phi \in \mathcal{A}$ , and both $C_1(a)^\phi$ and $C_2(a)^\phi$ are $\mathcal{A}$ -insertable
	then	$\mathcal{A}' := \mathcal{A} \oplus C_1(a)^\phi, \mathcal{A}'' := \mathcal{A} \oplus C_2(a)^\phi$
$\exists$ -rule	if	$(\exists R.C)(a)^\phi \in \mathcal{A}$ , there is no $c$ such that neither $R(a,c)^\phi$ nor $C(c)^\phi$ is $\mathcal{A}$ -insertable, and $a$ is not blocked
	then	$\mathcal{A}' := (\mathcal{A} \oplus R(a,b)^\phi) \oplus C(b)^\phi$ , where $b$ is a new individual name
$\forall$ -rule	if	$\{(\forall R.C)(a)^\phi, R(a,b)^\psi\} \subseteq \mathcal{A}$ , and $C(b)^{\phi \wedge \psi}$ is $\mathcal{A}$ -insertable
	then	$\mathcal{A}' := \mathcal{A} \oplus C(b)^{\phi \wedge \psi}$
$\sqsubseteq$ -rule	if	$(C \sqsubseteq D)^\phi \in \mathcal{O}$ , $a$ appears in $\mathcal{A}$ , and $(\neg C \sqcup D)(a)^\phi$ is $\mathcal{A}$ -insertable
	then	$\mathcal{A}' := \mathcal{A} \oplus (\neg C \sqcup D)(a)^\phi$

Fig. 2. Expansion rules for constructing  $\phi_K^\perp$ .

the worlds whose restriction is an inconsistent  $\mathcal{ALC}$  ontology; hence, we are interested in detecting and highlighting contradictions.

For this algorithm, it will be important to distinguish the terminological portion of the ontology (the TBox) from the assertional knowledge (the ABox). From a very high-level view, the algorithm starts by creating a partial model that satisfies all the constraints from the ABox, and then uses rules to decompose complex concepts until an actual valuation (based exclusively on concept names and role names) is reached. Throughout the process, rules are also used to guarantee that all the constraints from the TBox are satisfied by the individuals in this model.

The algorithm starts with the ABox  $\mathcal{A}$  from  $\mathcal{O}$ . If  $\mathcal{O}$  contains no assertions, we assume implicitly that the ABox is  $\{\top(a)\}$  for an arbitrary individual name  $a$ . Recall that all the axioms in  $\mathcal{A}$  are labeled with a context, and we preserve the context at this initialization step. From this point, the algorithm deals with a set of ABoxes  $\mathfrak{A}$ , which are created and modified following the rules from Figure 2. Throughout the execution of the algorithm, each assertion in the ABoxes is labeled with a context which, as usual, is a propositional formula over the context variables; the algorithm modifies these labels accordingly to store all the relevant information. As a pre-requisite for the execution of the algorithm, we assume that all concepts appearing in the ontology are in *negation normal form* (NNF), that is, only concept names can appear in the scope of a negation operator. This assumption is made w.l.o.g. because every concept can be transformed into NNF in linear time by applying the De Morgan laws, the duality of the quantifiers, and eliminating double negations. Each rule application chooses an ABox  $\mathcal{A} \in \mathfrak{A}$  and replaces it by one or two new ABoxes that expand  $\mathcal{A}$  with new assertions. We explain the details of these rule applications next.

An assertion  $\alpha^\phi$  is  *$\mathcal{A}$ -insertable* iff for all  $\psi$  such that  $\alpha^\psi \in \mathcal{A}$ ,  $\phi \not\equiv \psi$ . In the expansion rules, the symbol  $\oplus$  is used as shorthand for the operation

$$\mathcal{A} \oplus \alpha^\phi := \begin{cases} (\mathcal{A} \setminus \{\alpha^\psi\}) \cup \{\alpha^{\phi \vee \psi}\} & \alpha^\psi \in \mathcal{A} \\ \mathcal{A} \cup \{\alpha^\phi\} & \text{otherwise.} \end{cases}$$

The main idea of this operator is to update the contexts in which an assertion is required to hold. Either the assertion is not already present in the ABox, and hence it is included with the context obtained by conjoining all the contexts triggering the rule application or, if the assertion is already present, there are multiple ways to derive the same assertion,

whose contextual causes are all disjoint. Rather than keeping two different assertions in the ABox, we simply update its label.

Within an ABox  $\mathcal{A}$ , we say that the individual  $x$  is an *ancestor* of  $y$  if there is a chain of role assertions connecting  $x$  to  $y$ ; more formally, if there exist role names  $r_1, \dots, r_n$  and individuals  $a_0, \dots, a_n$  with  $n \geq 1$  such that  $\{r_i(a_{i-1}, a_i) \mid 1 \leq i \leq n\} \subseteq \mathcal{A}$ ,  $a_0 = x$ , and  $a_n = y$ . The node  $x$  *blocks*  $y$  iff  $x$  is an ancestor of  $y$  and for every  $C(y)^\psi \in \mathcal{A}$ , there is a  $\phi$  such that  $C(x)^\phi \in \mathcal{A}$  and  $\psi \models \phi$ ;  $y$  is *blocked* if there is a node that blocks it.

The algorithm applies the expansion rules from Figure 2 until  $\mathfrak{A}$  is *saturated*, that is, until no rule is applicable to any  $\mathcal{A} \in \mathfrak{A}$ . We say that the ABox  $\mathcal{A}$  contains a *clash* if  $\{A(x)^\phi, \neg A(x)^\psi\} \subseteq \mathcal{A}$  for some individual  $x$  and concept name  $A$ . Informally, a clash means that the ABox has an obvious contradiction, as it requires to interpret the element  $x$  to belong to the concept  $A$  and its negation. We define the context

$$\phi_{\mathcal{A}} := \bigvee_{A(a)^\phi, \neg A(a)^\psi \in \mathcal{A}} \phi \wedge \psi,$$

which intuitively describes the contexts of all the clashes that appear in  $\mathcal{A}$ . When  $\mathfrak{A}$  is saturated, we return the context  $\phi_{\mathcal{K}}^\perp = \bigwedge_{\mathcal{A} \in \mathfrak{A}} \phi_{\mathcal{A}}$  expressing the need of having clashes in every ABox  $\mathcal{A}$  for inconsistency to follow. This is because each ABox in  $\mathfrak{A}$  is a potential model, where the multiplicity arises from the non-deterministic choice when decomposing disjunctions  $\sqcup$ . If all potential choices are obviously contradictory, then no model can exist. It is important for the correctness of the algorithm to notice that the definition of a clash does not impose any constraints on the contexts  $\phi$  and  $\psi$  labeling the assertions  $A(x)$  and  $\neg A(x)$ , respectively. Indeed,  $A(x)$  and  $\neg A(x)$  could hold in contradictory contexts. In that case, the conjunction appearing in  $\phi_{\mathcal{A}}$  would not be affected, that is, this clash will not provide any new information about inconsistency. For example, if  $\mathcal{A} = \{\text{WaterPipe}(\text{pipe1})^X, \neg \text{WaterPipe}(\text{pipe1})^{\neg X}\}$ , then the context  $\phi_{\mathcal{A}}$  is  $X \wedge \neg X$ . This context cannot be satisfied by any world, meaning that  $\mathcal{A}$  is not inconsistent.

Informally, the context  $\phi_{\mathcal{K}}^\perp$  corresponds to the *clash formula* (also called pinpointing formula) for explaining inconsistency of an  $\mathcal{ALC}$  ontology (Lee et al. 2006; Baader and Peñaloza 2010). The main differences are that the variables appearing in a context are not necessarily Boolean but multi-valued, and that the axioms in  $\mathcal{O}$  are not labeled with unique variables but rather with contexts. The correctness of the approach follows from the fact that each assertion  $\alpha$  is labeled with a formula representing the contexts in which  $\alpha$  is necessarily true. Thus, every inconsistent context will be represented in  $\phi_{\mathcal{K}}^\perp$ . Conversely, saturation of the expansion rules guarantees that no other derivations of inconsistency as possible; otherwise, some rule would still be applicable.

Notice that the expansion rules in Figure 2 generalize the expansion rules for  $\mathcal{ALC}$ , but may require new rule applications to guarantee that all possible derivations of a clash are detected. Indeed, while the classical algorithm never applies a rule where the derived assertion already exists in the ABox, our modified algorithm may apply the rule if the label associated with the assertion does not already imply that which will be included. As observed in Baader and Peñaloza (2007; 2010), one has to be careful with claims of termination of the modified method, since the additional rule applications may in fact lead to an infinite execution. Specifically, the modified algorithm obtained from a terminating tableau is not guaranteed to terminate. However, Baader and Peñaloza (2010) and Peñaloza (2009) devised some general conditions, which are sufficient to guarantee

termination of a method modified in this way. Importantly, those conditions are met whenever the original algorithm deals only with unary and binary assertions, as is the case of the  $\mathcal{ALC}$  algorithm presented here. Hence, our approach is also guaranteed to terminate. This yields the following result.

*Theorem 15*

The modified tableau algorithm terminates, and the context  $\phi_{\mathcal{K}}^{\perp}$  is such that for every world  $\omega$ ,  $\omega \models \phi_{\mathcal{K}}^{\perp}$  iff  $\mathcal{K}_{\omega}$  is inconsistent.

The proof of this theorem is analogous to the proofs of correctness and termination of pinpointing extensions for tableau algorithms from Baader and Peñaloza (2010) only taking the differences between propositional formulas and contexts into account.

We now turn our attention to the computation of the formula  $\phi_{\mathcal{B}}$ . Recall that in a BN, the JPD is the product of the conditional probabilities of each variable given its parents. Hence, a world  $\omega$  can only have probability 0 if it evaluates some variable in  $X \in V$  and its parents  $\pi(X)$  to values  $x$  and  $\vec{x}$ , respectively, such that  $P(X = x \mid \pi(X) = \vec{x}) = 0$ . Thus, to compute  $\phi_{\mathcal{B}}$  it suffices to find out the cells in the conditional probability tables in  $\Theta$  with value 0.

*Theorem 16*

Let  $\mathcal{B} = (V, \Theta)$  be a BN, and define

$$\phi_{\mathcal{B}} := \bigvee_{P(X=x \mid \pi(X)=\vec{x})=0} \left( (X, x) \wedge \bigwedge_{Y \in \pi(X), Y=y} (Y, y) \right).$$

Then, for every world  $\omega$ ,  $\omega \models \phi_{\mathcal{B}}$  iff  $P_{\mathcal{B}}(\omega) = 0$ .

*Proof*

If  $\omega \models \phi_{\mathcal{B}}$ , then there is a variable  $X$  such that  $\omega \models (X, x) \wedge \bigwedge_{Y \in \pi(X), Y=y} (Y, y)$  and  $P(X = x \mid \pi(X) = \vec{x}) = 0$ . By the chain rule, it follows that  $P_{\mathcal{B}}(\omega) = 0$ . Conversely, if  $0 = P_{\mathcal{B}}(\omega) = \prod_{X \in V} P(X \mid \pi(X))$ , it must be the case that at least one factor in the product is 0, but then that factor satisfies a disjunct in  $\phi_{\mathcal{B}}$ .  $\square$

Notice that, in general, the context  $\phi_{\mathcal{B}}$  can be computed faster than simply enumerating all possible worlds. In particular, if the conditional probability tables in  $\Theta$  contain no 0-valued cell, then  $\phi_{\mathcal{B}} = \perp$ , that is, it is satisfied by no world.

To summarize, in this section we have shown that deciding inconsistency of a  $\mathcal{BALC}$  KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$  can be reduced to checking satisfiability of the context  $\phi_{\mathcal{K}}^{\perp} \wedge \neg\phi_{\mathcal{B}}$ , where  $\phi_{\mathcal{K}}^{\perp}$  is the context representing all the worlds  $\omega$  where  $\mathcal{O}_{\omega}$  is inconsistent, and  $\phi_{\mathcal{B}}$  is the context representing all worlds  $\omega$  with  $P_{\mathcal{B}}(\omega) = 0$ . We then provided a tableaux-like method for computing  $\phi_{\mathcal{K}}^{\perp}$ , and an enumeration approach for finding  $\phi_{\mathcal{B}}$ . We note that, just like in classical  $\mathcal{ALC}$ , the tableau algorithm is not optimal in terms of computational complexity. In fact, although consistency of  $\mathcal{ALC}$  ontologies is in ExpTime (Donini and Massacci 2000), the tableau algorithm we use as the basis of our approach runs in double exponential time; the main reason for this is the  $\sqcup$ -rule which may duplicate the number of ABoxes under consideration. We conjecture that it can be implemented in a goal-directed manner which behaves well in practice, but leave this for future work.

Although consistency is a very important problem to be studied, we are interested also in other reasoning tasks, which can arise through the use of probabilistic ontologies. In particular, we should also take into account the contexts and the probabilities provided by the BN beyond the question of whether they are positive or not. In the next section, we study variants of satisfiability and subsumption problems, before turning our attention to instance checking.

## 5 Satisfiability and subsumption

In this section, we focus on two problems – satisfiability and subsumption – which depend only on the TBox part of an ontology, and hence assume for the sake of simplicity that the ABox is empty. Thus, we will often write a  $\mathcal{BALC}$  KB as a pair  $(\mathcal{T}, \mathcal{B})$ , where  $\mathcal{T}$  is a TBox and  $\mathcal{B}$  is a BN. We are in general interested in understanding the properties and relationships of concepts.

### Definition 17

Given two concepts  $C, D$  and a  $\mathcal{BALC}$  KB  $\mathcal{K}$ , we say that  $C$  is *satisfiable* w.r.t.  $\mathcal{K}$  iff there exists a probabilistic model  $\mathcal{P} = (\mathcal{J}, \mathcal{P}_{\mathcal{J}})$  of  $\mathcal{K}$  s.t.  $C^{\mathcal{V}} \neq \emptyset$  for all  $\mathcal{V} \in \mathcal{J}$ . We say that  $C$  is *subsumed* by  $D$  w.r.t.  $\mathcal{K}$  iff for all models  $\mathcal{P} = (\mathcal{J}, \mathcal{P}_{\mathcal{J}})$  of  $\mathcal{K}$  and all  $\mathcal{V} \in \mathcal{J}$   $C^{\mathcal{V}} \subseteq D^{\mathcal{V}}$ .

It is possible to adapt the well-known reductions from the classical case to show that these two problems are EXPTIME-complete.

### Theorem 18

Satisfiability and subsumption w.r.t.  $\mathcal{BALC}$  KBs are EXPTIME-complete.

### Proof

Let  $\mathcal{K} = (\mathcal{T}, \mathcal{B})$  and  $C, D$  be concepts. It is easy to see that  $C$  is subsumed by  $D$  w.r.t.  $\mathcal{K}$  iff the KB  $\mathcal{K}' = (\mathcal{T} \cup \{(C \sqcap \neg D)(a)^{\emptyset}\}, \mathcal{B})$  is inconsistent, where  $a$  is an arbitrary individual name. Similarly,  $C$  is satisfiable iff  $\mathcal{K}'' = (\mathcal{T} \cup \{C(a)^{\emptyset}\}, \mathcal{B})$  is consistent.  $\square$

In the following, we study variants of these problems which take contexts and probabilities into account. For a more concise presentation, we will present only the cases for subsumption. Analogous results hold for satisfiability based on the fact that for every  $\mathcal{ALC}$  interpretation  $\mathcal{I}$ , it holds that  $C^{\mathcal{I}} = \emptyset$  iff  $C^{\mathcal{I}} \subseteq \perp^{\mathcal{I}}$ ; hence, an approach that is similar to the one used to solve subsumption can be used to solve satisfiability. First, we consider additional information about contexts; afterwards we compute the probability of an entailment, and finally the combination of both kinds of information.

### Definition 19 (contextual subsumption)

Let  $\mathcal{K} = (\mathcal{T}, \mathcal{B})$  be a  $\mathcal{BALC}$  KB,  $C, D$  concepts, and  $\kappa$  a context.  $C$  is *subsumed by  $D$  in context  $\kappa$*  w.r.t.  $\mathcal{K}$ , denoted as  $\mathcal{K} \models (C \sqsubseteq D)^{\kappa}$ , if every probabilistic model of  $\mathcal{K}$  is also a model of  $(C \sqsubseteq D)^{\kappa}$ .

This definition introduces the natural extension of subsumption that considers also the contexts. Note that in our setting, contexts provide a means to express and reason with probabilities.

*Definition 20 (subsumption probability)*

Let  $\mathcal{P} = (\mathcal{J}, P_{\mathcal{J}})$  be a probabilistic model of the KB  $\mathcal{K}$ ,  $\kappa$  a context, and  $C, D$  two concepts. The *probability of  $(C \sqsubseteq D)^\kappa$  w.r.t.  $\mathcal{P}$*  is

$$P_{\mathcal{P}}((C \sqsubseteq D)^\kappa) = \sum_{\mathcal{V} \in \mathcal{J}, \mathcal{V} \models (C \sqsubseteq D)^\kappa} P_{\mathcal{J}}(\mathcal{V}).$$

The *probability of  $(C \sqsubseteq D)^\kappa$  w.r.t.  $\mathcal{K}$*  is

$$P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) = \inf_{\mathcal{P} \models \mathcal{K}} P_{\mathcal{P}}((C \sqsubseteq D)^\kappa).$$

We say that  $C$  is *positively subsumed* by  $D$  in  $\kappa$  iff  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) > 0$ ; it is  *$p$ -subsumed* iff  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) \geq p$ ; it is *exactly  $p$ -subsumed* iff  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) = p$ ; and it is *almost certainly subsumed* iff  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) = 1$ .

The definition above allows us to consider the problem of probabilistic reasoning. That is, the probability of a subsumption in a specific model is the sum of the probabilities of the worlds in which  $C$  is subsumed by  $D$  in context  $\kappa$ ; notice that this trivially includes all worlds where  $\kappa$  does not hold, by construction (see Definition 5). In the case where  $\mathcal{K}$  is inconsistent, we define the probability of all subsumptions as 1 to ensure our definition is consistent with general probability theory (recall that  $\inf(\emptyset) = \infty$  in general).

When we speak about the probability of a subsumption w.r.t. a *BALC* KB, we consider the infimum over the probabilities provided by all possible models of  $\mathcal{K}$ . The reason for this is that we want to find the tightest constraint, taking into account the open-world semantics. The infimum is, to a degree, the natural interpretation of the information that holds in all possible models. When reasoning about probabilities it is not always important to compute a precise value but rather find some relevant bounds. Thus,  $p$ -subsumption considers a lower bound, while positive subsumption only cares about the probability being zero or not (dually, almost-certainty refers to the probability being one or not). All these decision problems have different applications and properties.

Contextual subsumption is related to subsumption probability in the obvious way. Namely, a KB  $\mathcal{K}$  entails a contextual subsumption iff the probability of the subsumption in  $\mathcal{K}$  is 1.

*Theorem 21*

Given a KB  $\mathcal{K}$ , concepts  $C$  and  $D$ , and a context  $\kappa$ , it holds that:

$$\mathcal{K} \models (C \sqsubseteq D)^\kappa \text{ iff } P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) = 1.$$

*Proof*

$\mathcal{K} \models (C \sqsubseteq D)^\kappa$  iff for every probabilistic interpretation  $\mathcal{P}$ , if  $\mathcal{P} \models \mathcal{K}$ , then  $\mathcal{P} \models (C \sqsubseteq D)^\kappa$  iff for every probabilistic model  $\mathcal{P} = (\mathcal{I}, P_{\mathcal{I}})$  of  $\mathcal{K}$  and every  $\mathcal{V} \in \mathcal{I}$ ,  $\mathcal{V} \models (C \sqsubseteq D)^\kappa$  iff for every model  $\mathcal{P}$  of  $\mathcal{K}$   $P_{\mathcal{P}}((C \sqsubseteq D)^\kappa) = 1$  iff  $\inf_{\mathcal{P} \models \mathcal{K}} P_{\mathcal{P}}((C \sqsubseteq D)^\kappa) = 1$ .  $\square$

This is convenient as it provides a method of reusing our results from Section 4, which finds a context describing inconsistency, to compute subsumption probabilities.



Table 1. Worlds  $\omega$  such that  $\omega \models W$  and their probability w.r.t the BN from Figure 1

World $\omega$				$P_{\mathcal{B}}(\omega)$
$X$	$Y$	$Z$	$W$	0
$X$	$Y$	$\neg Z$	$W$	0.063
$X$	$\neg Y$	$Z$	$W$	0
$X$	$\neg Y$	$\neg Z$	$W$	0.567
$\neg X$	$Y$	$Z$	$W$	0
$\neg X$	$Y$	$\neg Z$	$W$	0.189
$\neg X$	$\neg Y$	$Z$	$W$	0.0162
$\neg X$	$\neg Y$	$\neg Z$	$W$	0.0108

*Theorem 22*

Let  $\mathcal{K} = (\mathcal{T}, \mathcal{B})$  be a consistent  $\mathcal{BALC}$  KB,  $C, D$  two concepts, and  $\kappa$  a context. For the KB  $\mathcal{K}' = (\mathcal{T} \cup \{C(a)^\kappa, \neg D(a)^\kappa\}, \mathcal{B})$  it holds that

$$P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) = \sum_{\omega \models \phi_{\mathcal{K}'}} P_{\mathcal{B}}(\omega) + 1 - P_{\mathcal{B}}(\kappa).$$

*Proof*

Let  $\mathcal{O} = (\mathcal{T} \cup \{C(a)^\kappa, \neg D(a)^\kappa\})$ , so  $\mathcal{K}' = (\mathcal{O}, \mathcal{B})$ . For an arbitrary but fixed model  $\mathcal{P} = (\mathcal{J}, P_{\mathcal{J}})$  of  $\mathcal{K}$ , we have that

$$\begin{aligned} P_{\mathcal{P}}((C \sqsubseteq D)^\kappa) &= \sum_{\mathcal{V} \in \mathcal{J}, \mathcal{V} \models (C \sqsubseteq D)^\kappa} P_{\mathcal{J}}(\mathcal{V}) \\ &= \sum_{\mathcal{V} \not\models \kappa} P_{\mathcal{J}}(\mathcal{V}) + \sum_{\mathcal{V} \models \kappa, \mathcal{V} \models C \sqsubseteq D} P_{\mathcal{J}}(\mathcal{V}) \\ &= 1 - P_{\mathcal{B}}(\kappa) + \sum_{\mathcal{V} \not\models \kappa'} P_{\mathcal{J}}(\mathcal{V}) \\ &\leq 1 - P_{\mathcal{B}}(\kappa) + \sum_{\{\mathcal{V} \mid \mathcal{O}_{\mathcal{V}} \text{ inconsistent}\}} P_{\mathcal{J}}(\mathcal{V}). \end{aligned}$$

Hence,  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) \leq \sum_{\omega \models \phi_{\mathcal{K}'}} P_{\mathcal{B}}(\omega) + 1 - P_{\mathcal{B}}(\kappa)$ . For the lower bound, it suffices to build a minimal model of  $\mathcal{K}$ , where for each world  $\omega$  we have an interpretation  $\mathcal{V}_\omega$  such that  $\mathcal{V}_\omega \models C \sqsubseteq D$  iff  $\mathcal{O}_\omega \models C \sqsubseteq D$  as in the proof of Theorem 11.  $\square$

*Example 23*

Returning to our running example with the KB  $\mathcal{K}$  from Example 7, suppose that we are interested in finding the probability of water being drinkable. That is, we want to compute  $P_{\mathcal{K}}((\text{Water} \sqsubseteq \text{Drinkable})^\emptyset)$ . Since  $P_{\mathcal{B}}(\emptyset) = 1$ , Theorem 22 states that it suffices to find the probabilities of all worlds  $\omega$  such that the restriction of  $\mathcal{O} \cup \{\text{Water}(a), \neg \text{Drinkable}(a)\}$  to  $\omega$  is inconsistent. We note that the only such worlds are those which satisfy  $W$ . Hence,

$$P_{\mathcal{K}}((\text{Water} \sqsubseteq \text{Drinkable})^\emptyset) = \sum_{\omega \models W} P_{\mathcal{B}}(\omega).$$

Those worlds, and their probabilities, are shown in Table 1. Adding all these values, we get that the probability is 0.8460. In other words, according to our model, the probability of a building having drinkable water (in the absence of any other information) is very high.

Notice that the formula  $\phi_{\mathcal{K}}^\perp$ , requires at most exponential space on the size of  $\mathcal{T}$  to be encoded. For each of the exponentially many worlds, computing  $P_{\mathcal{B}}(\omega)$  requires polynomial time due to the chain rule. Hence, overall, the computation of the subsumption probabilities requires exponential time. Importantly, this bound does not depend on how  $\phi_{\mathcal{K}}^\perp$  was computed; it could have, for example, been computed through the process described in Corollary 12. This provides an exponential upper bound for computing the probability of a subsumption.

*Corollary 24*

The probability of a subsumption w.r.t. a KB can be computed in exponential time on the size of the KB.

Obviously, an exponential-time upper bound for computing the exact probability of a subsumption relation immediately yields an EXPTIME upper bound for deciding the other problems introduced in Definition 20 as well. All these problems are also generalizations of the subsumption problem in  $\mathcal{ALC}$ . More precisely, given an  $\mathcal{ALC}$  TBox  $\mathcal{T}$ , we can create the  $\mathcal{BALC}$  KB  $\mathcal{K} = (\mathcal{T}', \mathcal{B})$  where  $\mathcal{T}'$  contains all the axioms in  $\mathcal{T}$  labeled with the context  $x$  and  $\mathcal{B}$  contains only one Boolean node  $x$  that holds with probability 1. Given two concepts  $C, D$   $\mathcal{T} \models C \sqsubseteq D$  iff  $C$  is almost certainly subsumed by  $D$  in context  $x$ . Since subsumption in  $\mathcal{ALC}$  is already EXPTIME-hard, we get that all these problems are EXPTIME-complete.

In practice, however, it may be too expensive to compute the exact probability when we are only interested in determining lower bounds, or the possibility of observing an entailment, for instance, when considering positive subsumption. Recall once again that, according to our semantics, a contextual GCI  $(C \sqsubseteq D)^\kappa$  will hold in any world  $\omega$  such that  $\omega \not\models \kappa$ . Thus, if the probability of this world is positive ( $P_{\mathcal{B}}(\omega) > 0$ ), we can immediately guarantee that  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) > 0$ . Thus, positive subsumption can be decided without any ontological reasoning for any context that is not almost certain. In all other cases, the problem can still be reduced to inconsistency.

*Theorem 25*

Let  $\mathcal{K} = (\mathcal{T}, \mathcal{B})$  be a consistent KB. The concept  $C$  is positively subsumed by  $D$  in context  $\kappa$  w.r.t.  $\mathcal{K}$  iff  $\mathcal{K}' = (\mathcal{T} \cup \{C(a)^\kappa, \neg D(a)^\kappa\}, \mathcal{B})$  is inconsistent or  $P_{\mathcal{B}}(\kappa) < 1$ .

*Proof*

Assuming that  $\mathcal{K}$  is consistent, if  $P_{\mathcal{B}}(\kappa) < 1$ , the result is an immediate consequence of Theorem 22. Otherwise,  $P_{\mathcal{B}}(\kappa) = 1$ , and since  $\mathcal{K}$  is consistent, inconsistency of  $\mathcal{K}'$  can only arise from the two new assertions, which means that  $\kappa \models \phi_{\mathcal{K}'}^\perp$ , and in particular  $\sum_{\omega \models \phi_{\mathcal{K}'}^\perp} P_{\mathcal{B}}(\omega) = 1$ . □

This theorem requires the context  $\kappa$  to be specified to decide positive subsumption. If  $\kappa$  is not known before hand, it is also possible to leverage the inconsistency decision process, which is the most expensive part of this method; recall that otherwise we only need  $\phi_{\mathcal{B}}$ , which can be easily computed from  $\mathcal{B}$ . Let  $\mathcal{K}_\emptyset := (\mathcal{T} \cup \{C(a)^\emptyset, \neg D(a)^\emptyset\}, \mathcal{B})$ . That is,

$\mathcal{K}_\emptyset$  extends  $\mathcal{K}$  with assertions negating the subsumption relation, which should hold in all contexts. From Theorem 15, we conclude that  $\phi_{\mathcal{K}_\emptyset}^\perp$  encodes all the contexts in which  $C \sqsubseteq D$  must hold. Notice that the computation of  $\phi_{\mathcal{K}_\emptyset}^\perp$  does not depend on the context  $\kappa$  but can be used to decide positive subsumption for any context.

*Corollary 26*

The concept  $C$  is positively subsumed by  $D$  in context  $\kappa$  w.r.t.  $\mathcal{K} = (\mathcal{T}, \mathcal{B})$  iff  $\kappa$  entails  $\phi_{\mathcal{K}_\emptyset}^\perp$  or  $P_{\mathcal{B}}(\kappa) < 1$ .

Considering the probabilities of contextual subsumption relations may lead to unexpected or unintuitive results arising from the contextual semantics. Indeed, it always holds (see Theorem 25) that  $P_{\mathcal{K}}((C \sqsubseteq D)^\kappa) \geq 1 - P_{\mathcal{B}}(\kappa)$ . In other words, the probability of a subsumption in a very unlikely context will always be very high, regardless of the KB and concepts used. In essence, this is caused because the semantics of contextual subsumption impose a variation of a material implication – if  $\kappa$  holds, then the subsumption holds – which is vacuously true if the premise (if  $\kappa$  holds) is false. In some cases, it may be more meaningful to consider a *conditional* probability under the assumption that the context  $\kappa$  holds, as defined next.

*Definition 27 (conditional subsumption)*

Let  $\mathcal{P} = (\mathcal{J}, P_{\mathcal{J}})$  be a probabilistic model of the KB  $\mathcal{K}$ ,  $\kappa, \lambda$  two contexts with  $P_{\mathcal{B}}(\lambda) > 0$ , and  $C, D$  two concepts. The *conditional probability of  $(C \sqsubseteq D)^\kappa$  given  $\lambda$*  w.r.t.  $\mathcal{P}$  is

$$P_{\mathcal{P}}((C \sqsubseteq D)^\kappa \mid \lambda) = \frac{\sum_{\mathcal{V} \in \mathcal{J}, \mathcal{V} \models \lambda, \mathcal{V} \models (C \sqsubseteq D)^\kappa} P_{\mathcal{J}}(\mathcal{V})}{P_{\mathcal{B}}(\lambda)}.$$

The *conditional probability of  $(C \sqsubseteq D)^\kappa$  given  $\lambda$*  w.r.t.  $\mathcal{K}$  is

$$P_{\mathcal{K}}((C \sqsubseteq D)^\kappa \mid \lambda) = \inf_{\mathcal{P} \models \mathcal{K}} P_{\mathcal{P}}((C \sqsubseteq D)^\kappa \mid \lambda).$$

This definition follows the same principles of conditioning in probability theory, but extended to the open-world interpretation provided by our model-based semantics. Notice that, in addition to the scaling factor  $P_{\mathcal{B}}(\lambda)$  in the denominator, the nominator is also differentiated from Definition 20 by considering only the worlds that satisfy the context  $\lambda$  already.

Notice that the conditional probabilities in Definition 27 are pretty similar to those in Definition 20, except that the sum restricts to only the worlds that satisfy the context  $\lambda$ . Thus, the numerator can be obtained from the contextual probability in the context  $\kappa \wedge \lambda$  excluding the worlds that violate  $\lambda$ . More formally, we have that

$$\begin{aligned} \sum_{\mathcal{V} \in \mathcal{J}, \mathcal{V} \models \lambda, \mathcal{V} \models (C \sqsubseteq D)^\kappa} P_{\mathcal{J}}(\mathcal{V}) &= \sum_{\mathcal{V} \models \lambda, \mathcal{V} \not\models \kappa} P_{\mathcal{J}}(\mathcal{V}) + \sum_{\mathcal{V} \models \lambda \wedge \kappa, \mathcal{V} \models C \sqsubseteq D} P_{\mathcal{J}}(\mathcal{V}) \\ &= P_{\mathcal{B}}(\lambda) - P_{\mathcal{B}}(\lambda \wedge \kappa) + P_{\mathcal{P}}((C \sqsubseteq D)^{\lambda \wedge \kappa}) - 1 + P_{\mathcal{B}}(\lambda \wedge \kappa) \\ &= P_{\mathcal{P}}((C \sqsubseteq D)^{\lambda \wedge \kappa}) + P_{\mathcal{B}}(\lambda) - 1. \end{aligned}$$

Thus, we get the following result.

*Theorem 28*

$$P_{\mathcal{K}}((C \sqsubseteq D)^\kappa \mid \lambda) = \frac{P_{\mathcal{P}}((C \sqsubseteq D)^{\lambda \wedge \kappa}) + P_{\mathcal{B}}(\lambda) - 1}{P_{\mathcal{B}}(\lambda)}.$$

In particular, this means that also conditional probabilities can be computed through contextual probabilities, with a small overhead of computing the probability (on the BN  $\mathcal{B}$ ) of the conditioning context  $\lambda$ .

As in the contextual case, if one is only interested in knowing that the subsumption is possible (i.e., that it has a positive probability), then one can exploit the complex context describing the inconsistent contexts which, as mentioned before, can be precompiled to obtain the contexts in which a subsumption relation must be satisfied. However, in this case, entailment between contexts is not sufficient; one must still compute the probability of the contextual subsumption.

*Corollary 29*

Let  $\mathcal{K} = (\mathcal{T}, \mathcal{B})$  be a consistent KB,  $C, D$  concepts, and  $\kappa, \lambda$  contexts s.t.  $P_{\mathcal{B}}(\lambda) > 0$ .  $P_{\mathcal{K}}((C \sqsubseteq D)^{\kappa} \mid \lambda) > 0$  iff  $P_{\mathcal{P}}((C \sqsubseteq D)^{\lambda \wedge \kappa}) > 1 - P_{\mathcal{B}}(\lambda)$ .

In other words,  $P((C \sqsubseteq D)^{\kappa} \mid \lambda) > 0$  iff  $C$  is  $p$ -subsumed by  $D$  in  $\kappa \wedge \lambda$  with  $p = 1 - P_{\mathcal{B}}(\lambda)$ .

*Example 30*

Continuing Example 23, suppose that we know that the context  $\neg X$  holds, that is, the house was built before 1986. Following Theorem 28, we can compute

$$P_{\mathcal{K}}((\text{Water} \sqsubseteq \text{Drinkable})^{\emptyset} \mid \neg X) = \frac{P_{\mathcal{K}}((\text{Water} \sqsubseteq \text{Drinkable})^{\neg X}) + P_{\mathcal{B}}(\neg X) - 1}{P_{\mathcal{B}}(\neg X)}.$$

The denominator can be read directly from the probability distribution in Figure 1, that is,  $P_{\mathcal{B}}(\neg X) = 0.3$ . From Theorem 22, the numerator becomes the sum of the probabilities of all worlds that become inconsistent with the addition of the assertions  $\{\text{Water}(a)^{\neg X}, \neg \text{Drinkable}(a)^{\neg X}\}$ . These are the same appearing in the last four rows of Table 1. Hence,

$$P_{\mathcal{K}}((\text{Water} \sqsubseteq \text{Drinkable})^{\emptyset} \mid \neg X) = \frac{0.216}{0.3} = 0.72.$$

As expected, knowing that the house is pre-1986 increases the probability of having lead tubing, which means a decrease in the probability of drinkable water.

Analogously to Definitions 19, 20, and 27, it is possible to define the notions of consistency of a concept  $C$  to hold in only some contexts, and based on it, the (potentially conditional) probability of such a contextual consistency problem. As mentioned already, it is well known that for every  $\mathcal{ALC}$  interpretation  $\mathcal{I}$  it holds that  $C^{\mathcal{I}} = \emptyset$  iff  $C^{\mathcal{I}} \sqsubseteq \perp^{\mathcal{I}}$ . Hence, all these problems can be solved through a direct reduction to their related subsumption problem.

We now turn our attention to the problem of instance checking. In this problem, the ABox also plays a role. Hence, we consider once again ontologies  $\mathcal{O}$  that can have in addition to GCIs, concept, and role assertions.

### 6 Instance checking

We consider a probabilistic extension to the classical instance checking problem. In  $\mathcal{BALC}$ , we call this problem probabilistic instance checking and we define both a decision problem and probability calculation for it next, following the same pattern as in the previous section.

*Definition 31 (Instance)*

Given a KB  $\mathcal{K}$  and a context  $\kappa$ , the individual name  $a$  is an *instance* of the concept  $C$  in  $\kappa$  w.r.t.  $\mathcal{K}$ , written  $\mathcal{K} \models C(a)^\kappa$ , iff for all probabilistic models  $\mathcal{P} = (\mathcal{J}, P_{\mathcal{J}})$  of  $\mathcal{K}$  and for all  $\mathcal{V} \in \mathcal{J}$  it holds that  $\mathcal{V} \models C(a)^\kappa$ .

That is,  $a$  is an instance of  $C$  in  $\kappa$  if every interpretation in each model  $\mathcal{P}$  satisfies the assertion  $C(a)^\kappa$ .

Note that as before, instance checking in  $\mathcal{ALC}$  is a special case of this definition that can be obtained by considering a BN with only one variable that is true with probability 1. Notice that, contrary to the case of satisfiability studied at the end of the previous section, it is not possible to reduce instance checking to subsumption since an instance may be – and in fact usually is – caused by ABox assertions, which are ignored for subsumption tests. The latter observation is not of great consequence though, since both instance checking and subsumption can be reduced to consistency checking.

*Theorem 32*

Given an individual name  $a \in N_I$ , a concept  $C$ , a context  $\kappa$ , and a KB  $\mathcal{K} = (\mathcal{O}, \mathcal{B})$ ,  $\mathcal{K} \models C(a)^\kappa$  iff the KB  $\mathcal{K}' = (\mathcal{O} \cup \{(-C(a))^\kappa\}, \mathcal{B})$  is inconsistent.

In particular, this means that instance checking is at most as hard as deciding consistency of a KB. As mentioned already, it is also at least as hard as instance checking in the classical  $\mathcal{ALC}$ . Hence, we get the following result.

*Corollary 33*

Instance checking w.r.t.  $\mathcal{BALC}$  KBs is EXPTIME-complete.

Let us now consider the probabilistic entailments related to instance checking.

*Definition 34 (instance probability)*

The *probability of an instance* in a probabilistic model  $\mathcal{P} = (\mathcal{J}, P_{\mathcal{J}})$  of the KB  $\mathcal{K}$  is

$$P_{\mathcal{P}}(C(x)^\kappa) = \sum_{\mathcal{V} \in \mathcal{J}, \mathcal{V} \models C(x)^\kappa} P_{\mathcal{J}}(\mathcal{V}).$$

The *instance probability* w.r.t. a KB  $\mathcal{K}$  is

$$P_{\mathcal{K}}(C(x)^\kappa) = \inf_{\mathcal{P} \models \mathcal{K}} P_{\mathcal{P}}(C(x)^\kappa).$$

The conditional probability of an instance in a particular probabilistic model  $\mathcal{P} = (\mathcal{J}, P_{\mathcal{J}})$ , where  $P_{\mathcal{B}}(\lambda) > 0$ , is

$$P_{\mathcal{P}}(C(x)^\kappa \mid \lambda) = \frac{\sum_{\mathcal{V} \in \mathcal{J}, \mathcal{V} \models \lambda, \mathcal{V} \models C(x)^\kappa} P_{\mathcal{J}}(\mathcal{V})}{P_{\mathcal{B}}(\lambda)},$$

The probability of the conditional instance in  $\mathcal{K}$  is

$$P_{\mathcal{K}}(C(x)^\kappa \mid \lambda) = \inf_{\mathcal{P} \models \mathcal{K}} P_{\mathcal{P}}(C(x)^\kappa \mid \lambda).$$

The probability of all instance checks for an inconsistent KB is always 1 to keep our definitions consistent with probability theory, just as we did for the probability of subsumption over an inconsistent KB in the previous section.

As we did for subsumption, we can exploit the reasoning techniques that were developed for deciding inconsistency of a  $\mathcal{BALC}$  KB to find out also the contextual and

conditional probabilities of an instance. Moreover, the method can be further optimized in the cases where we are only interested in probabilistic bounds. In particular, we can adapt Theorem 28 to this case, yielding the following result.

*Theorem 35*

$$P_{\mathcal{K}}(C(x)^{\kappa} \mid \lambda) = \frac{P_{\mathcal{K}}(C(x)^{\kappa \wedge \lambda}) + P_{\mathcal{B}}(\lambda) - 1}{P(\lambda)}.$$

*Example 36*

Suppose that we extend the ontology from Example 7 with the assertions  $\text{Pipe}(p)^{\neg X}$ ,  $\text{Lead}(m)^Z$ , and  $\text{contains}(p, m)^{\neg Y}$ . To compute  $P_{\mathcal{K}}(\text{LeadPipe}(p)^{\emptyset})$ , we must find out what  $P_{\mathcal{B}}(\neg X, \neg Y, Z)$  is. Following the chain rule from the BN, this is computed as  $P(\neg X)P(\neg Y \mid \neg X)P(Z \mid \neg X, \neg Y) = 0.3 \cdot 0.3 \cdot 0.6 = 0.054$ . That is, even though we have (probabilistic) information about a pipe  $p$ , which may contain a material which may be lead, the probability that  $p$  is a lead pipe is very low.

With Theorem 35, we conclude the technical analysis of contextual and probabilistic reasoning tasks over  $\mathcal{BALC}$  KBs. As it can be readily seen, the general approach is based on being able to detect the contexts in which the conditions imposed by the ontology are inconsistent. The probabilistic computations rely on cross-checking those contexts with the distribution described by the BN.

## 7 Conclusions

We have presented a new probabilistic extension of the DL  $\mathcal{ALC}$ , which is based on the ideas of Bayesian ontology languages. In this setting, KBs contain certain knowledge, which is dependent on the uncertain context where it holds. Our work extends the results on  $\mathcal{BEL}$  (Ceylan and Peñaloza 2014) to a propositionally closed ontology language. The main notions follow the basic ideas of general Bayesian ontology languages as presented in Ceylan and Peñaloza (2017); however, by focusing on a specific logic – rather than on a general notion of ontology language – we are able to produce a tableaux-based decision algorithm for KB consistency, in contrast to the generic black-box algorithms which exist in the literature. Our algorithm extends the classical tableau algorithm for  $\mathcal{ALC}$  with techniques originally developed for axiom pinpointing. The main differences in our approach are the use of multi-valued variables in the definition of the contexts, and the possibility of having complex contexts (instead of only unique variables) labeling individual axioms. In general, we have shown that adding context-based uncertainty to an ontology does not increase the complexity of reasoning in this logic: all (probabilistic) reasoning problems can still be solved in exponential time. Of course, this tight complexity bound is not extremely surprising given the EXPTIME-hardness of the underlying classical DL.

Theorems 13–16 yield an effective decision method for  $\mathcal{BALC}$  KB consistency, through the computation and handling of two complex contexts, which describe the logical and probabilistic properties of the worlds and the restricted ontologies they define. Notice that  $\phi_{\mathcal{B}}$  can be computed in linear time on the size of  $\mathcal{B}$ , and satisfiability of the context  $\phi_{\mathcal{K}}^{\perp} \wedge \phi_{\mathcal{B}}$  can be checked in non-deterministic polynomial time in the size of this context, following techniques akin to those from propositional satisfiability (Biere *et al.* 2009). However, the tableau algorithm for computing  $\phi_{\mathcal{K}}^{\perp}$  is not optimal w.r.t. worst-case complexity. In fact,

this algorithm requires double exponential time in the worst case, although the context it computes is only of exponential size. The benefit of this method, as in the classical case, is that it provides a better behavior in the average case; indeed, depending on the input ontology, it may not need to verify all the possible contexts but simply construct a model. As a simple example, if the TBox is empty, the process stops in exponential time. To further improve the efficiency of our approach, one can think of adapting the methods from Zese et al. (2018) to construct a compact representation of the context – akin to a binary decision diagram (Lee 1959; Brace et al. 1990) for multi-valued variables – allowing for efficient weighted model counting. A task for future work is to exploit these data structures for practical development of our methods.

Recall that the most expensive part of our approach is the computation of the context  $\phi_{\mathcal{K}}^{\perp}$ . By slightly modifying the KB, we have shown that one computation of this context suffices to solve different problems of interest, in particular contextual and conditional entailments – being subsumption, satisfiability, or instance checking – can be solved using  $\phi_{\mathcal{K}'}^{\perp}$ , for an adequately constructed  $\mathcal{K}'$ , regardless of the contexts under consideration.

An important next step will be to implement the methods described here and compare the efficiency of our system to other probabilistic DL reasoners based on similar semantics. In particular, we would like to compare against the tools from Zese et al. (2018). Even though this latter system is also based on an extension of the tableaux algorithm for DLs and uses multiple-world semantics closely related to ours, a direct comparison would be unfair. Indeed, Zese et al. (2018) make use of stronger independence assumptions than ours. However, a well-designed experiment can shed light on the advantages and disadvantages of each method.

Another interesting problem for future work is to extend the query language beyond instance queries. To maintain some efficiency, this may require some additional restrictions on the language or the probabilistic structure. A more detailed study of this issue is needed.

## References

- ARTALE, A., CALVANESE, D., KONTCHAKOV, R. AND ZAKHARYASCHEV, M. 2009. The DL-Lite family and relations. *Journal of Artificial Intelligence Research* 36, 1–69.
- BAADER, F., BRANDT, S. AND LUTZ, C. 2005. Pushing the EL envelope. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI 2005)*, L. P. Kaelbling and A. Saffiotti, Eds. Professional Book Center, 364–369.
- BAADER, F., CALVANESE, D., MCGUINNESS, D., NARDI, D., AND PATEL-SCHNEIDER, P., Eds. 2007. *The Description Logic Handbook: Theory, Implementation, and Applications*, 2nd ed. Cambridge University Press.
- BAADER, F. AND HOLLUNDER, B. 1995. Embedding defaults into terminological knowledge representation formalisms. *Journal of Automated Reasoning* 14, 1, 149–180.
- BAADER, F., HORROCKS, I., LUTZ, C. AND SATTLER, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.
- BAADER, F. AND PEÑALOZA, R. 2007. Axiom pinpointing in general tableaux. In *Proceedings of the 16th International Conference on Analytic Tableaux and Related Methods (TABLEAUX 2007)*, N. Olivetti, Ed. Lecture Notes in Artificial Intelligence, vol. 4548. Springer-Verlag, Aix-en-Provence, France, 11–27.



- BAADER, F. AND PEÑALOZA, R. 2010. Axiom pinpointing in general tableaux. *Journal of Logic and Computation* 20, 1 (February), 5–34. Special Issue: Tableaux and Analytic Proof Methods.
- BIERE, A., HEULE, M., VAN MAAREN, H. AND WALSH, T. 2009. *Handbook of Satisfiability: Volume 185 Frontiers in Artificial Intelligence and Applications*. IOS Press, Amsterdam, The Netherlands.
- BOTHA, L. 2018. The Bayesian description logic ALC. M.S. thesis, University of Cape Town, South Africa.
- BOTHA, L., MEYER, T. AND PEÑALOZA, R. 2018. The Bayesian description logic BALC. In *Proceedings of the 31st International Workshop on Description Logics (DL 2018)*, M. Ortiz and T. Schneider, Eds. CEUR Workshop Proceedings, vol. 2211. CEUR-WS.org.
- BOTHA, L., MEYER, T. AND PEÑALOZA, R. 2019. A Bayesian extension of the description logic ALC. In *Proceedings of the 16th European Conference on Logics in Artificial Intelligence (JELIA 2019)*, F. Calimeri, N. Leone, and M. Manna, Eds. Lecture Notes in Computer Science, vol. 11468. Springer, 339–354.
- BRACE, K. S., RUDELL, R. L. AND BRYANT, R. E. 1990. Efficient implementation of a BDD package. In *Proceedings of the 27th ACM/IEEE Design Automation Conference, DAC 1990*. ACM, New York, NY, USA, 40–45.
- CEYLAN, İ. İ. 2018. Query answering in probabilistic data and knowledge bases. Ph.D. thesis, Dresden University of Technology, Germany.
- CEYLAN, İ. İ. AND LUKASIEWICZ, T. 2018. A tutorial on query answering and reasoning over probabilistic knowledge bases. In *14th International Summer School on Reasoning Web*, C. d’Amato and M. Theobald, Eds. Lecture Notes in Computer Science, vol. 11078. Springer, 35–77.
- CEYLAN, İ. İ. AND PEÑALOZA, R. 2014. The Bayesian description logic BEL. In *Proceedings of the 7th International Joint Conference on Automated Reasoning (IJCAR 2014)*, S. Demri, D. Kapur, and C. Weidenbach, Eds. Lecture Notes in Computer Science, vol. 8562. Springer, 480–494.
- CEYLAN, İ. İ. AND PEÑALOZA, R. 2014. Tight complexity bounds for reasoning in the description logic BEL. In *Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA 2014)*, E. Fermé and J. Leite, Eds. Lecture Notes in Computer Science, vol. 8761. Springer, 77–91.
- CEYLAN, İ. İ. AND PEÑALOZA, R. 2017. The Bayesian ontology language BEL. *Journal of Automated Reasoning* 58, 1, 67–95.
- D’AMATO, C., FANIZZI, N. AND LUKASIEWICZ, T. 2008. Tractable reasoning with Bayesian description logics. In *Proceedings of the Second International Conference on Scalable Uncertainty Management*, S. Greco and T. Lukasiewicz, Eds. Lecture Notes in Computer Science, vol. 5291. Springer, 146–159.
- DARWICHE, A. 2009. *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press.
- DONINI, F. M. AND MASSACCI, F. 2000. Exptime tableaux for ALC. *Artificial Intelligence* 124, 1, 87–138.
- GOTTLOB, G., LUKASIEWICZ, T., MARTINEZ, M. V. AND SIMARI, G. I. 2013. Query answering under probabilistic uncertainty in datalog+/- ontologies. *Annals of Mathematics and Artificial Intelligence* 69, 1, 37–72.
- GUTIÉRREZ-BASULTO, V., JUNG, J. C., LUTZ, C., AND SCHRÖDER, L. 2017. Probabilistic description logics for subjective uncertainty. *Journal of Artificial Intelligence Research* 58, 1–66.
- HALPERN, J. Y. 1990. An analysis of first-order logics of probability. *Artificial Intelligence* 46, 3, 311–350.
- LEE, C. Y. 1959. Representation of switching circuits by binary-decision programs. *The Bell System Technical Journal* 38, 985–999.

- LEE, K., MEYER, T. A., PAN, J. Z. AND BOOTH, R. 2006. Computing maximally satisfiable terminologies for the description logic ALC with cyclic definitions. In *Proceedings of the 2006 International Workshop on Description Logics (DL2006)*, B. Parsia, U. Sattler, and D. Toman, Eds. CEUR Workshop Proceedings, vol. 189. CEUR-WS.org.
- LUKASIEWICZ, T. AND STRACCIA, U. 2008. Managing uncertainty and vagueness in description logics for the semantic web. *Journal of Web Semantics* 6, 4, 291–308.
- MEYER, T. A., LEE, K., BOOTH, R. AND PAN, J. Z. 2006. Finding maximally satisfiable terminologies for the description logic ALC. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*. AAAI Press, 269–274.
- PEARL, J. 1985. Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings of Cognitive Science Society (CSS-7)*, 329–334.
- PEÑALOZA, R. 2009. Axiom-pinpointing in description logics and beyond. Ph.D. thesis, Dresden University of Technology, Germany.
- SCHILD, K. 1991. A correspondence theory for terminological logics: Preliminary report. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI 1991)*, J. Mylopoulos and R. Reiter, Eds. Morgan Kaufmann, 466–471.
- SCHMIDT-SCHAUSS, M. AND SMOLKA, G. 1991. Attributive concept descriptions with complements. *Artificial Intelligence* 48, 1, 1–26.
- ZESE, R., BELLODI, E., RIGUZZI, F., COTA, G. AND LAMMA, E. 2018. Tableau reasoning for description logics and its extension to probabilities. *Annals of Mathematics and Artificial Intelligence* 82, 1–3, 101–130.