Typicalities and probabilities of exceptions in nonmotonic Description Logics

Published version:

DOI:10.1016/j.ijar.2019.02.003

Typicalities and Probabilities of Exceptions in Nonmotonic Description Logics[☆]

Gian Luca Pozzato[a,*]

[a] *Dipartimento di Informatica, Università degli Studi di Torino, corso Svizzera 185, 10149 Torino, Italy*

Abstract

We introduce a nonmonotonic procedure for preferential Description Logics in order to reason about typicality by taking probabilities of exceptions into account. We consider an extension, called $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, of the logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ by inclusions of the form $\mathbf{T}(C) \sqsubseteq_p D$ with probability p, whose intuitive meaning is that "all the typical Cs are Ds, and the probability that a C is not a D is $1 - p$". We consider a notion of extension of an ABox containing only some typicality assertions, then we equip each extension with a probability. We then restrict entailment of a query F to those extensions whose probabilities belong to a given and fixed range. We propose a decision procedure for reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ and we exploit it to show that entailment is ExpTime-complete as for the underlying $\mathcal{ALC}$.

Keywords: Description Logics, Typicality, Nonmonotonic Reasoning, Probabilities of exceptions

1. Introduction

Description Logics [1], for short: DLs, represent one of the most important formalisms of knowledge representation and are at the base of the languages for building ontologies in the Semantic Web such as OWL. Their success is essentially motivated by two key advantages: on the one hand, DLs have a well-defined semantics based on first-order logic; on the other hand, they provide a good trade-off between the expressivity of the language and the computational complexity of their reasoning services. Description Logics are useful in practice in several application domains.

According to Description Logics, a knowledge base contains two components:

- a TBox, containing inclusion relations among concepts: for instance, we would need to formalize the fact that cats are mammals, and this is represented by $Cat \sqsubseteq Mammal$;

[☆] This work is partially supported by the project "ExceptionOWL", Università di Torino and Compagnia di San Paolo, call 2014 "Excellent (young) PI", project ID: Torino_call2014_L1_111.

[*] Corresponding author

Email address: gianluca.pozzato@unito.it (Gian Luca Pozzato)

- an ABox containing facts about the domain, for instance we would need to formalize that Tom is a cat, and this is represented by $Cat(tom)$.

Standard Description Logics – and, as a consequence, existing ontologies – are not able to represent prototypical properties and to reason about defeasible inheritance. Recalling a well known example coming from the literature of nonmonotonic reasoning, we can have a TBox representing that birds fly ($\mathit{Bird} \sqsubseteq \mathit{Fly}$), but that penguins are birds that do not fly ($\mathit{Penguin} \sqsubseteq \mathit{Bird}$ and $\mathit{Penguin} \sqsubseteq \neg \mathit{Fly}$). This knowledge base is consistent only if there are no penguins. In order to tackle this problem, nonmonotonic extensions of Description Logics have been actively investigated since the early 90s [2, 3, 4, 5, 6, 7, 8], allowing one to represent *prototypical* properties of classes and to reason about *defasible* inheritance.

A simple but powerful nonmonotonic extension of DLs is proposed in [9]: in this approach "typical" or "normal" properties can be directly specified by means of a "typicality" operator $\mathbf{T}$ enriching the underlying DL, and a TBox can contain inclusions of the form $\mathbf{T}(C) \sqsubseteq D$ to represent that "typical Cs are also Ds" or "normally, Cs have the property D". The Description Logic so obtained is called $\mathcal{ALC} + \mathbf{T_R}$ and, as a difference with standard DLs, one can consistently express exceptions and reason about defeasible inheritance as well. For instance, a knowledge base can consistently express that "normally, referees award penalty kicks", whereas "Italian referees usually do not award penalty kicks" (since in Italian "serie A" a video assistant referee "VAR" often intervenes in order to change the official referee's decisions) as follows:

$$\mathbf{T}(\mathit{Referee}) \sqsubseteq \exists awards.\mathit{PenaltyKick}$$
$$\mathbf{T}(\mathit{Referee} \sqcap \mathit{Italian}) \sqsubseteq \neg \exists awards.\mathit{PenaltyKick}$$

The semantics of the $\mathbf{T}$ operator is characterized by a set of postulates that are essentially a restatement of axioms and rules of *rational entailment* as introduced in [10], recognized as the core properties of nonmonotonic reasoning. As a consequence, $\mathbf{T}$ inherits well-established properties like *specificity*: in the example, if one knows that Daniele is a typical Italian referee, then the logic $\mathcal{ALC} + \mathbf{T_R}$ allows us to infer that he usually does not award penalty kicks, giving preference to the most specific information.

The logic $\mathcal{ALC} + \mathbf{T_R}$ itself is too weak in several application domains. Indeed, although the operator $\mathbf{T}$ is nonmonotonic ($\mathbf{T}(C) \sqsubseteq E$ does not imply $\mathbf{T}(C \sqcap D) \sqsubseteq E$), the logic $\mathcal{ALC} + \mathbf{T_R}$ is monotonic, in the sense that if the fact F follows from a given knowledge base KB, then F also follows from any KB' $\supseteq$ KB. As a consequence, unless a KB contains explicit assumptions about typicality of individuals, there is no way of inferring defeasible properties about them: in the above example, if KB contains the fact that Mark is a referee, i.e.

$$\mathit{Referee}(mark)$$

belongs to KB, it is not possible to infer that he awards penalty kicks

$$\exists awards.\mathit{PenaltyKick}(mark).$$

This would be possible only if the stronger information that Mark is a *typical* referee

$$\mathbf{T}(Referee)(mark)$$

belongs to (or can be inferred from) KB. In order to overwhelm this limit and perform useful inferences, in [11, 12] the authors have introduced a nonmonotonic extension of the logic $\mathcal{ALC} + \mathbf{T_R}$ based on a minimal model semantics, corresponding to a notion of *rational closure* as defined in [10] for propositional logic. Intuitively, the idea is to restrict our consideration to (canonical) models that maximize typical instances of a concept when consistent with the knowledge base. The resulting logic, call it $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, supports typicality assumptions, so that if one knows that Mark is a referee, one can nonmonotonically assume that he is also a *typical* referee if this is consistent, and therefore that he awards penalty kicks. From a semantic point of view, the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ is based on a preference relation among $\mathcal{ALC} + \mathbf{T_R}$ models and a notion of *minimal entailment* restricted to models that are minimal with respect to such preference relation.

The logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{RaCl}$ imposes to consider *all* typicality assumptions that are consistent with a given KB. Let us consider another example, where a Description Logic knowledge base expresses that typical students are young persons that, normally, make use of social networks, as well as that, normally, Italians love spaghetti. Furthermore, the knowledge base states that a typical young person goes to parties. Moreover, we have that Mario, Fabrizio, Pietro, Ruggero, Patrizia, Roberta and Donatella are Italian students. If it is consistent to assume that they are typical ones, then the logic imposes that they are *all* social network users, that they *all* love spaghetti and that they *all* go to parties. We have seven different students, and the logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{RaCl}$ assumes that each one of them corresponds to a prototypical one. This would also happen in case we had hundreds of students, leading to the assumption that, in absence of explicit information (for instance, in case we discover that Donatella does not like spaghetti), there are no exceptions: this seems to be too strong in several application domains. It could be useful to reason about scenarios with *exceptional individuals*, or one could need to assign different *probabilities* to typicality inclusions. In the example, one could need to represent that the properties of being young and being part of the social media ecosphere are all typical properties of students: however, it could be needed to also describe that the probability of finding exceptional students not being young is lower than the one of finding exceptional students not using social networks.

In this work we introduce a new Description Logic called $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, which extends $\mathcal{ALC}$ by means of typicality inclusions equipped by *probabilities of exceptionality* of the form

$$\mathbf{T}(C) \sqsubseteq_p D,$$

where $p \in (0, 1)$. The intuitive meaning is that:

"normally, Cs are Ds and the probability of having exceptional Cs – not being Ds – is $1 - p$".

In other words, all the typical instances of the concept C are also instances of the concept D, and the probability that a C element is not also a D element, i.e. it is an exceptional C element, is $1 - p$. For instance, we can have

$\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.6} \mathit{SportLover}$
$\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.9} \mathit{SocialNetworkUser}$

whose intuitive meaning is that being sport lovers and social network users are both typical properties of students, however the probability of having exceptional students not loving sport is higher than the one of finding students not using social networks, in particular we have the evidence that the probability of having exceptions is 40% and 10%, respectively.

It is worth noticing that the probability p equipping a typicality inclusion $\mathbf{T}(C) \sqsubseteq_p D$ could be wrongly interpreted as "typical Cs are also Ds with probability p", stating that if an individual is a typical instance of the concept C, then there is the probability p that such an instance is also a D element, then that there are some typical Cs that are not Ds: this is not the case, since in the semantics of the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$ underlying our logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, that we will recall in Section 2, all typical Cs are Ds and, as mentioned here above, p is used to represent the probability of (not) finding exceptional Cs not being Ds.

As a difference with DLs under the distributed semantics introduced in [13, 14], where probabilistic axioms of the form $p\ :: C \sqsubseteq D$ are used to capture uncertainty in order to represent that Cs are Ds with probability p, in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ we are able to ascribe typical properties to concepts and to reason about probabilities of exceptions to those typicalities. We define different extensions of an ABox containing only some of the "plausible" typicality assertions: each extension represents a scenario having a specific probability. Then, we provide a notion of nonmonotonic entailment restricted to extensions whose probabilities belong to a given and fixed range, in order to reason about scenarios that are not necessarily the most probable. We introduce a decision procedure for checking entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ and we exploit it in order to show that reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ with probabilities of exceptions is EXPTIME complete, therefore we retain the same complexity of the underlying standard $\mathcal{ALC}$.

The plan of the paper, which extends and revises a preliminary version appeared in [15], is as follows. In Section 2 we recall the basic concepts of Description Logics extended with the typicality operator. In Section 3 we extend such logics in order to deal with probabilities of exceptions introducing the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, whereas in Section 4 we introduce a decision procedure for reasoning in the proposed logic and we study its complexity. We conclude with a discussion and a comparison with related approaches in Section 5.

2. Preferential Description Logics

Let us first recall the main notions about the Description Logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ introduced in [9, 16, 11]. The logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ is obtained by adding to standard $\mathcal{ALC}$ the typicality operator $\mathbf{T}$ [9]. The intuitive idea is that $\mathbf{T}(C)$ selects the *typical* instances of a concept C. We can therefore distinguish between the properties that hold for all instances of concept C ($C \sqsubseteq D$), and those that only hold for the normal or typical instances of C ($\mathbf{T}(C) \sqsubseteq D$).

The semantics of the $\mathbf{T}$ operator can be given by means of a set of postulates that are a reformulation of axioms and rules of nonmonotonic entailment in rational logic **R**

[10]: in this respect an assertion of the form $\mathbf{T}(C) \sqsubseteq D$ is equivalent to the conditional assertion $C \mid\!\sim D$ in **R**. A model $\mathcal{M}$ is a triple $\langle \Delta^{\mathcal{I}}, f_{\mathbf{T}}, .^{\mathcal{I}} \rangle$: given a domain $\Delta^{\mathcal{I}}$ and an evaluation function $.^{\mathcal{I}}$, one can define a function $f_{\mathbf{T}} : Pow(\Delta^{\mathcal{I}}) \longmapsto Pow(\Delta^{\mathcal{I}})$ that selects the *typical* instances of any $S \subseteq \Delta^{\mathcal{I}}$; in case $S = C^{\mathcal{I}}$ for a concept C, the selection function selects the typical instances of C, namely:

$$(\mathbf{T}(C))^{\mathcal{I}} = f_{\mathbf{T}}(C^{\mathcal{I}}).$$

$f_{\mathbf{T}}$ has the following properties for all subsets S of $\Delta^{\mathcal{I}}$, that are essentially a restatement of the properties characterizing rational logic **R**:

$(f_{\mathbf{T}} - 1)$ $f_{\mathbf{T}}(S) \subseteq S$
$(f_{\mathbf{T}} - 2)$ if $S \neq \emptyset$, then also $f_{\mathbf{T}}(S) \neq \emptyset$
$(f_{\mathbf{T}} - 3)$ if $f_{\mathbf{T}}(S) \subseteq R$, then $f_{\mathbf{T}}(S) = f_{\mathbf{T}}(S \cap R)$
$(f_{\mathbf{T}} - 4)$ $f_{\mathbf{T}}(\bigcup S_i) \subseteq \bigcup f_{\mathbf{T}}(S_i)$
$(f_{\mathbf{T}} - 5)$ $\bigcap f_{\mathbf{T}}(S_i) \subseteq f_{\mathbf{T}}(\bigcup S_i)$
$(f_{\mathbf{T}} - 6)$ if $f_{\mathbf{T}}(S) \cap R \neq \emptyset$, then $f_{\mathbf{T}}(S \cap R) \subseteq f_{\mathbf{T}}(S)$

The semantics of the $\mathbf{T}$ operator can be equivalently formulated in terms of *rational models* [11]: a model $\mathcal{M}$ is any structure $\langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ where $\Delta^{\mathcal{I}}$ is the domain, $<$ is an irreflexive, transitive, well-founded and modular (for all x, y, z in $\Delta^{\mathcal{I}}$, if $x < y$ then either $x < z$ or $z < y$) relation over $\Delta^{\mathcal{I}}$. In this respect, $x < y$ means that x is "more normal" than y, and that the typical members of a concept C are the minimal elements of C with respect to this relation. An element $x \in \Delta^{\mathcal{I}}$ is a *typical instance* of some concept C if $x \in C^{\mathcal{I}}$ and there is no C-element in $\Delta^{\mathcal{I}}$ *more typical* than x. In detail, $.^{\mathcal{I}}$ is the extension function that maps each concept C to $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and each role R to $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. For concepts of $\mathcal{ALC}$, $C^{\mathcal{I}}$ is defined as usual. For the $\mathbf{T}$ operator, we have

$$(\mathbf{T}(C))^{\mathcal{I}} = Min_{<}(C^{\mathcal{I}}),$$

where $Min_{<}(C^{\mathcal{I}}) = \{x \in C^{\mathcal{I}} \mid \nexists y \in C^{\mathcal{I}} \text{ s.t. } y < x\}$. We have obtained a result similar to those ones obtained in [17, 18, 19] for the propositional case:

Theorem 1 (Representation theorem in [11]). *A knowledge base is satisfiable in a model $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ as above if and only if it is satisfiable in a model $\mathcal{M} = \langle \Delta, f_{\mathbf{T}}, .^{\mathcal{I}} \rangle$ where $f_{\mathbf{T}}$ satisfies $(f_{\mathbf{T}} - 1) - (f_{\mathbf{T}} - 6)$, and $(\mathbf{T}(C))^{\mathcal{I}} = f_{\mathbf{T}}(C^{\mathcal{I}})$.*

A model $\mathcal{M}$ can be equivalently defined by postulating the existence of a function $k_{\mathcal{M}} : \Delta^{\mathcal{I}} \longmapsto \mathbb{N}$, where $k_{\mathcal{M}}$ assigns a finite rank to each domain element: the rank function $k_{\mathcal{M}}$ and $<$ can be defined from each other by letting $x < y$ if and only if $k_{\mathcal{M}}(x) < k_{\mathcal{M}}(y)$.

Given standard definitions of satisfiability of a KB in a model, we define a notion of entailment in $\mathcal{ALC} + \mathbf{T_R}$. Given a query F (either an inclusion $C \sqsubseteq D$ or an assertion $C(a)$ or an assertion of the form $R(a, b)$), we say that F is entailed from a KB, written KB $\models_{\mathcal{ALC}+\mathbf{T_R}} F$, if F holds in all $\mathcal{ALC} + \mathbf{T_R}$ models satisfying KB.

Even if the typicality operator $\mathbf{T}$ itself is nonmonotonic (i.e. $\mathbf{T}(C) \sqsubseteq E$ does not imply $\mathbf{T}(C \sqcap D) \sqsubseteq E$), what is inferred from a KB can still be inferred from any KB' with KB $\subseteq$ KB', i.e. the logic $\mathcal{ALC} + \mathbf{T_R}$ is monotonic. In order to perform useful nonmonotonic inferences, in [11] the authors have strengthened the above semantics by restricting entailment to a class of minimal models. Intuitively, the idea is to restrict entailment to models that *minimize the untypical instances of a concept*. The resulting logic is called $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ and it corresponds to a notion of *rational closure* on top of $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$. Such a notion is a natural extension of the rational closure construction provided in [10] for the propositional logic.

The nonmonotonic semantics of $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ relies on minimal rational models that minimize the *rank of domain elements*. Informally, given two models of KB, one in which a given domain element x has rank 2 (because for instance $z < y < x$), and another in which it has rank 1 (because only $y < x$), we prefer the latter, as in this model the element x is assumed to be "more typical" than in the former. Query entailment is then restricted to minimal *canonical models*. The intuition is that a canonical model contains all the individuals that enjoy properties that are consistent with KB. A model $\mathcal{M}$ is a minimal canonical model of KB if it satisfies KB, it is minimal and it is canonical[1]. A query F is minimally entailed from a KB, written KB $\models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{RaCl}} F$, if it holds in all minimal canonical models of KB.

In order to ascribe typical properties to individuals, the notion of rational closure is extended to the ABox: in particular, the typicality of an individual is maximized by minimizing its rank. In general, it is not possible to separately assign a unique minimal rank to each individual, then alternative minimal ranks must be considered. The idea is that of considering all the possible minimal consistent assignments of ranks to the individuals explicitly named in the ABox. Each assignment adds some properties to named individuals which can be used to infer new conclusions. A skeptical view of considering only those conclusions which hold for all assignments is then adopted. More formally, the idea is that an individual a_i can have a given rank $k_j(a_i)$ just in case it is compatible with all the inclusions of the TBox that do not contain the $\mathbf{T}$ operator or that have a $\mathbf{T}(C)$ on the left-hand side with C's rank which is at least $k_j(a_i)$. The minimal possible rank assignment k_j for all a_i is computed as follows: μ_i^j computes all the concepts that a_i would need to satisfy in case it had the rank $k_j(a_i)$. The algorithm verifies whether μ_i^j is compatible with the rational closure of the TBox and whether it is minimal. All constants are considered simultaneously, since the possible ranks of different individual constants depend on each other. The union of all μ_i^j (for all a_i) takes into account the ranks attributed to all individual constants.

In [11] it is shown that query entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ is in EXPTIME. The construction of the rational closure and the correspondence between semantics and construction is recalled in Section 4.

3. Dealing with Probabilities of Exceptions

In this section we define an alternative semantics that allows us to equip a typicality inclusion with the probability of *not* having exceptions for that, and then to reason

[1]In Theorem 10 in [11] the authors have shown that for any consistent KB there exists a finite minimal canonical model of KB.

about such inclusions. In the resulting Description Logic, called $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, a typicality inclusion has the form

$$\mathbf{T}(C) \sqsubseteq_p D,$$

and its intuitive meaning is "typical Cs are also Ds, and the probability of having exceptional Cs not being Ds is $1-p$". We then define a nonmonotonic procedure whose aim is to describe alternative completions of the ABox obtained by assuming typicality assertions about the individuals explicitly named in the ABox: the basic idea is similar to the one proposed in [9], where a completion of an $\mathcal{ALC}$+$\mathbf{T}$ ABox is proposed in order to assume that every individual constant of the ABox is a typical element of the most specific concept he belongs to, if this is consistent with the knowledge base. An analogous approach is proposed in [20], where different extensions of the ABox are introduced in order to define plausible but *surprising* scenarios. Here we propose a similar, algorithmic construction in order to compute only *some* assumptions of typicality of individual constants, in order to describe alternative scenarios having different probabilities: different extensions/scenarios are obtained by considering different sets of typicality assumptions of the form $\mathbf{T}(C)(a)$, where a occurs in the ABox.

Definition 1. *We consider an alphabet of concept names $\mathcal{C}$, of role names $\mathcal{R}$, and of individual constants $\mathcal{O}$. Given $A \in \mathcal{C}$ and $R \in \mathcal{R}$, we define:*

$$C := A \mid \top \mid \bot \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \forall R.C \mid \exists R.C$$

An $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ knowledge base is a pair $(\mathcal{T}, \mathcal{A})$. $\mathcal{T}$ contains axioms of the form either

- $C \sqsubseteq C$ *or*
- $\mathbf{T}(C) \sqsubseteq_p C$

where $p \in \mathbb{R}, p \in (0,1)$.

$\mathcal{A}$ contains assertions of the form either

- $C(a)$ *or*
- $R(a,b)$

where $a, b \in \mathcal{O}$.

Given an inclusion $\mathbf{T}(C) \sqsubseteq_p D$, the higher the probability p the more the inclusion is "exceptions-free" or, equivalently, the less is the probability of having exceptional Cs not being also Ds. In this respect, the probability p is a real number included in the open interval $(0,1)$: the probability 1 is not allowed, in the sense that an inclusion $\mathbf{T}(C) \sqsubseteq_1 D$ (the probability of having exceptional Cs not being Ds is 0) corresponds to a *strict* inclusion $C \sqsubseteq D$ (all Cs are Ds). Given another inclusion $\mathbf{T}(C') \sqsubseteq_{p'} D'$, with $p' < p$, we assume that this inclusion is less "strict" than the other one, i.e. the probability of having exceptional C's is higher than the one of having exceptional Cs with respect to properties D' and D, respectively. Recalling the example of the Introduction, where KB contains $\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.9} \mathit{SocialNetworkUser}$ and $\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.6} \mathit{SportLover}$, we have that typical students make use of social networks, and that normally they also love sport; however, the second inclusion is less probable with respect to the first one: both are properties of a prototypical student, however there are more exceptions of students not loving sport with respect to those not being active on social networks.

Before introducing formal definitions, we provide an example inspired to Example 1 in [20] in order to give an intuitive idea of what we mean for reasoning in $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ with probabilities of exceptions. We will complete it with part 2 in Example 5.

Example 1 (Reasoning in $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ part 1). *We aim at providing a formalization of some information about illnesses and symptoms. Let KB = ($\mathcal{T}, \mathcal{A}$) where $\mathcal{T}$ is as follows:*

$$\begin{array}{lr} \mathit{Bipolar} \sqsubseteq \mathit{Depressed} & (1) \\ \mathbf{T}(\mathit{Depressed}) \sqsubseteq_{0.85} \neg \exists \mathit{hasSymptom}.\mathit{MoodReactivity} & (2) \\ \mathbf{T}(\mathit{Bipolar}) \sqsubseteq_{0.7} \exists \mathit{hasSymptom}.\mathit{MoodReactivity} & (3) \\ \mathbf{T}(\mathit{ProstateCancerPatient}) \sqsubseteq_{0.6} \exists \mathit{hasSymptom}.\mathit{MoodReactivity} & (4) \\ \mathbf{T}(\mathit{ProstateCancerPatient}) \sqsubseteq_{0.8} \exists \mathit{hasSymptom}.\mathit{Nocturia} & (5) \end{array}$$

The above TBox $\mathcal{T}$ represents that (2)*, normally,* depressed *people do not have* mood reactivity*, namely the ability to feel better temporarily in response to positive life events. On the contrary,* (3) *states that this is a typical symptom of the bipolar disorder, a subtype of depression with atypical features that shares many of the typical symptoms of depression but is characterized by improved mood in response to positive events. Inclusion* (1) *intuitively represents that the bipolar disorder is a kind of depression. Mood reactivity, as well as* nocturia*, are also typical symptoms of* prostatic cancer *(inclusions* (4) *and* (5)*, respectively): more in detail,* (4) *says that we have a probability of* 40% *of having exceptional prostatic cancer patients with no mood swings, whereas* (5) *says that the probability of having exceptional prostatic cancer patients without nocturia is* 20%*.*

Concerning TBox reasoning, as a first example we have that in the logic $\mathcal{ALC}+\mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$, from the above knowledge base we can infer[2] *that, normally, depression in patients is not classified as bipolar disorder:*

$$\mathbf{T}(\mathit{Depressed}) \sqsubseteq \neg \mathit{Bipolar},$$

and this is a wanted inference. As another example of reasoning about the TBox, we have that

$$(6)\ \mathbf{T}(\mathit{Depressed} \sqcap \mathit{Spleenless}) \sqsubseteq \neg \exists \mathit{hasSymptom}.\mathit{MoodReactivity}$$

follows from KB, and this is also a wanted inference, since undergoing spleen removal is irrelevant with respect to mood reactivity as far as we know. This is a nonmonotonic

[2] As mentioned, at this point of the presentation we only want to give an intuition of inferences characterizing $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$. Technical details and definitions will be provided in Definitions 5 and 6.

inference that does no longer follow if it is discovered that typical depressed people without their spleen are subject to mood reactivity: given

$$\mathcal{T}' = \mathcal{T} \cup \{\mathbf{T}(\mathit{Depressed} \sqcap \mathit{Spleenless}) \sqsubseteq \exists \mathit{hasSymptom}.\mathit{MoodReactivity}\},$$

we have that the inclusion (6) *does no longer follow from KB with* $\mathcal{T}'$ *in the logic* $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$.

As for rational closure, the set of inclusions that are entailed from a $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ *KB is closed under the property known as* rational monotonicity*: for instance, from KB and the fact that the inclusion representing that, normally, depressed people are not elder*

$$\mathbf{T}(\mathit{Depressed}) \sqsubseteq \neg \mathit{Elder}$$

is not entailed from KB in $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$*, it follows that we can infer the inclusion*

$$\mathbf{T}(\mathit{Depressed} \sqcap \mathit{Elder}) \sqsubseteq \neg \exists \mathit{hasSymptom}.\mathit{MoodReactivity},$$

namely, a typical depressed and elder patient has not mood reactivity (the subconcept $\mathit{Depressed} \sqcap \mathit{Elder}$ *inherits the typical properties of the concept* $\mathit{Depressed}$*).*

Concerning ABox reasoning, if we know that Jim is depressed:

$$\mathcal{A} = \{\mathit{Depressed}(\mathit{jim})\},$$

then we can infer that Jim has not mood swings with a probability of 85%*, since* $\mathbf{T}(\mathit{Depressed})(\mathit{jim})$ *is minimally entailed from KB in* $\mathcal{ALC} + \mathbf{T}^{\mathit{RaCl}}_{\mathbf{R}}$ *and the inclusion* (2) *is equipped by a probability of* 0.85*. If we discover that Jim is affected by a bipolar disorder, then* $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ *allows us to retract such inference, whereas the fact that Jim has mood swings (*$\exists \mathit{hasSymptom}.\mathit{MoodReactivity}(\mathit{jim})$*) is entailed and evaluated having probability of* 70%*. The same conclusions are also entailed in case we discover that Jim is elder, i.e.* $\mathit{Elder}(\mathit{jim})$ *is added to the ABox, in detail:*

- *from* $(\mathcal{T}, \{\mathit{Depressed}(\mathit{jim}), \mathit{Elder}(\mathit{jim})\})$*, the logic* $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ *allows us to infer* $\neg \exists \mathit{hasSymptom}.\mathit{MoodReactivity}(\mathit{jim})$ *with probability of* 85%*;*

- *from* $(\mathcal{T}, \{\mathit{Bipolar}(\mathit{jim}), \mathit{Elder}(\mathit{jim})\})$*, the logic* $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ *allows us to infer* $\exists \mathit{hasSymptom}.\mathit{MoodReactivity}(\mathit{jim})$ *with probability of* 70%.

It is worth noticing that it is possible to have knowledge bases containing inclusions of the form $\mathbf{T}(C) \sqsubseteq_p D$, where $p \leq 0.5$ that, if wrongly interpreted, could be considered as counter intuitive. For instance, the inclusion $\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.3} \mathit{Young}$ could be wrongly interpreted as "normally, students are not young people". However, probabilities in $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$ are not intended to express neither degrees of belief of the inclusions they equip nor a notion of proportion of exceptions. In the example, even if its corresponding probability of exceptionality is low, the right interpretation of $\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.3} \mathit{Young}$ is that being a young person is *anyway* a property of a prototypical student: as a difference with $\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.9} \neg \mathit{SocialNetworkUser}$, we essentially have that the probability of finding exceptional students not being young is higher than the one of finding exceptional students not using social networks, but both are typical properties of a student. In case the ontology engineer needs to formalize that typical students are not young person, he just need to have $\mathbf{T}(Student) \sqsubseteq_p \neg Young$ in his KB with a suitable p. As mentioned before, the correct reading of a typicality inclusion $\mathbf{T}(C) \sqsubseteq_p D$ is that the probability of finding exceptional Cs not being Ds is $1-p$, which is different from "typical Cs are also Ds with probability p": the language proposed in this work allows the user/the ontology engineer to ascribe typical properties of a concept C, and then to equip each property with a probability p of (not) finding exceptions. The meaning of probability here is significantly different from those of the DISPONTE semantics in [14] and, as we will discuss in Section 5.1, to define typicality in probabilistic DLs in [21]: in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ all typicality inclusions represent typicality properties, independently from probabilities equipping them. This is why one can make use of probabilities lower than 0.5, since it could be useful in situations like the one described in the following example.

Example 2. *Suppose that we want to automatically build an ontology in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{P}$ from the information of a web site dedicated to the villains of Disney cartoons*[3]*. Our objective is to build a prototype of the villain, extracting information from available resources, in the form of inclusions $\mathbf{T}(Villain) \sqsubseteq_{p_1} C_1$, $\mathbf{T}(Villain) \sqsubseteq_{p_2} C_2$, ..., $\mathbf{T}(Villain) \sqsubseteq_{p_n} C_n$, where probabilities $p_1, p_2, \ldots, p_n$ are automatically calculated in a suitable way. It could be the case that, processing resources reporting features of 30 different villains, 13 over 30 propose an "Intelligent" villain, whereas for 6 over 30 the web pages suggest that considered villains are not intelligent. We have no information about the remaining 11 characters about their being intelligent or not. The inaccuracy of the web resources could suggest to consider "Intelligent" as a typical property of villains, even if 13 over 30 is less than 50%. This could be justified by the fact that the difference between the percentage of intelligent villains and the one of not intelligent villains is significantly higher (over a given and fixed threshold). In this case, the system could capture this situation with an inclusion*

$$\mathbf{T}(Villain) \sqsubseteq_{0.43} Intelligent,$$

where 0.43 is computed as $13/30$. Following the same approach, suppose that 3 over 30 villains are good looking, whereas 10 over 30 are not: in my approach, this could be captured with an inclusion

$$\mathbf{T}(Villain) \sqsubseteq_{0.33} \neg GoodLooking.$$

Obviously, one can think of considering as typical only those properties occurring in more than the 50% (or even more) of the elements belonging to a given class: in the example, if we further discover that 27 over 30 villain characters has an hero as his opponent, we can have

$$\mathbf{T}(Villain) \sqsubseteq_{0.9} \exists hasOpponent.Hero,$$

[3] We want to stress that, as in any probabilistic formal framework, probabilities are assumed to come from an application domain. This is true also for other frameworks: probabilities for our typicality inclusions come out in the same way of probabilities in probabilistic extensions of logic programs or degrees of belief in fuzzy logics. In this paper, we focus on the proposal of the formalism itself, therefore the machinery for obtaining probabilities from a dataset of the application domain is out of the scope.

without adding any inclusion about the properties "Intelligent" and "Good Looking": our language allows the ontology engineer to choose what he wants to consider as typical, in particular in a context when available resources (especially in the www) have a low level of accuracy.

3.1. Extensions of ABox

Given a KB, we define the finite set $\mathfrak{Tip}$ of concepts occurring in the scope of the typicality operator, i.e. $\mathfrak{Tip} = \{C \mid \mathbf{T}(C) \sqsubseteq_p D \in \text{ KB}\}$. Given an individual a explicitly named in the ABox, we define the set of typicality assumptions $\mathbf{T}(C)(a)$ that can be minimally entailed from KB in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, with $C \in \mathfrak{Tip}$. We then consider an ordered set $\mathfrak{Tip}_{\mathcal{A}}$ of pairs (a, C) of all possible assumptions $\mathbf{T}(C)(a)$, for all concepts $C \in \mathfrak{Tip}$ and all individual constants a in the ABox.

Definition 2 (Assumptions in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$). *Given an $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{P}$ KB=($\mathcal{T}, \mathcal{A}$), let $\mathcal{T}'$ be the set of inclusions of $\mathcal{T}$ without probabilities, namely*

$$\mathcal{T}' = \{\mathbf{T}(C) \sqsubseteq D \mid \mathbf{T}(C) \sqsubseteq_p D \in \mathcal{T}\} \cup \{C \sqsubseteq D \in \mathcal{T}\}.$$

Given a finite set of concepts $\mathfrak{Tip}$, we define, for each individual name a occurring in $\mathcal{A}$:

$$\mathfrak{Tip}_a = \{C \in \mathfrak{Tip} \mid (\mathcal{T}', \mathcal{A}) \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{RaCl}} \mathbf{T}(C)(a)\}.$$

We also define

$$\mathfrak{Tip}_{\mathcal{A}} = \{(a, C) \mid C \in \mathfrak{Tip}_a \textit{ and } a \textit{ occurs in } \mathcal{A}\}$$

and we impose an order on its elements: $\mathfrak{Tip}_{\mathcal{A}} = [(a_1, C_1), (a_2, C_2), \ldots, (a_n, C_n)]$. Furthermore, we define the ordered multiset

$$\mathcal{P}_{\mathcal{A}} = [p_1, p_2, \ldots, p_n],$$

respecting the order imposed on $\mathfrak{Tip}_{\mathcal{A}}$, where

$$p_i = \prod_{j=1}^{m} p_{ij} \textit{ for all } \mathbf{T}(C_i) \sqsubseteq_{p_{i1}} D_1, \mathbf{T}(C_i) \sqsubseteq_{p_{i2}} D_2, \ldots, \mathbf{T}(C_i) \sqsubseteq_{p_{im}} D_m \textit{ in } \mathcal{T}.$$

The ordered multiset $\mathcal{P}_{\mathcal{A}}$ is a tuple of the form $[p_1, p_2, \ldots, p_n]$, where p_i is the probability of the assumption $\mathbf{T}(C)(a)$, such that $(a, C) \in \mathfrak{Tip}_{\mathcal{A}}$ at position i. p_i is the product of all the probabilities p_{ij} of typicality inclusions $\mathbf{T}(C) \sqsubseteq_{p_{ij}} D$ in the TBox.

Following the basic idea underlying surprising scenarios outlined in [20], we consider different extensions $\widetilde{\mathcal{A}_i}$ of the ABox and we equip them with a probability $\mathbb{P}_i$. Starting from $\mathcal{P}_{\mathcal{A}} = [p_1, p_2, \ldots, p_n]$, the first step is to build all alternative tuples where 0 is used in place of some p_i to represent that the corresponding typicality assertion $\mathbf{T}(C)(a)$ is no longer assumed (Definition 3). Furthermore, we define the *extension* of the ABox corresponding to a string so obtained (Definition 4). In this way, the highest probability is assigned to the extension of the ABox corresponding to $\mathcal{P}_{\mathcal{A}}$, where all typicality assumptions are considered. The probability decreases in the other extensions, where some typicality assumptions are discarded, thus 0 is used in place of the corresponding p_i. The probability of an extension $\widetilde{\mathcal{A}_i}$ corresponding to a string $\mathcal{P}_{\mathcal{A}_i} = [p_{i1}, p_{i2}, \ldots, p_{in}]$ is defined as the product of probabilities p_{ij} when $p_{ij} \neq 0$, i.e. the probability of the corresponding typicality assumption when this is selected for the extension, and $1 - p_j$ when $p_{ij} = 0$, i.e. the corresponding typicality assumption is discarded, that is to say the extension contains an exception to the inclusion.

Definition 3 (Strings of possible assumptions $\mathbb{S}$). *Given a KB=($\mathcal{T}, \mathcal{A}$), let the set $\mathfrak{Tip}_{\mathcal{A}}$ and $\mathcal{P}_{\mathcal{A}} = [p_1, p_2, \ldots, p_n]$ be as in Definition 2. We define the set $\mathbb{S}$ of all the* strings of possible assumptions *with respect to KB as*

$$\mathbb{S} = \{[s_1, s_2, \ldots, s_n] \mid \forall i = 1, 2, \ldots, n \textit{ either } s_i = p_i \textit{ or } s_i = 0\}$$

Definition 4 (Extension of ABox). *Let KB=($\mathcal{T}, \mathcal{A}$), $\mathcal{P}_{\mathcal{A}} = [p1, p_2, \ldots, p_n]$ and $\mathfrak{Tip}_{\mathcal{A}} = [(a_1, C_1), (a_2, C_2), \ldots, (a_n, C_n)]$ as in Definition 2. Given a string of possible assumptions $[s_1, s_2, \ldots, s_n] \in \mathbb{S}$ of Definition 3, we define the extension $\widetilde{\mathcal{A}}$ of $\mathcal{A}$ with respect to $\mathfrak{Tip}_{\mathcal{A}}$ and $\mathbb{S}$ as:*

$$\widetilde{\mathcal{A}} = \{\mathbf{T}(C_i)(a_i) \mid (a_i, C_i) \in \mathfrak{Tip}_{\mathcal{A}} \textit{ and } s_i \neq 0\}$$

We also define the probability of $\widetilde{\mathcal{A}}$ as $\mathbb{P}_{\widetilde{\mathcal{A}}} = \prod_{i=1}^{n} \chi_i$ where $\chi_i = \begin{cases} p_i & \textit{if } s_i \neq 0 \\ 1 - p_i & \textit{if } s_i = 0 \end{cases}$

It can be observed that, in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, the set of typicality assumptions that can be inferred from a KB corresponds to the extension of $\mathcal{A}$ corresponding to the string $\mathcal{P}_{\mathcal{A}}$ (no element is set to 0): all the typicality assertions of individuals occurring in the ABox, that are consistent with the KB, are assumed. On the contrary, in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$, no typicality assumptions can be derived from a KB, and this corresponds to extending $\mathcal{A}$ by the assertions corresponding to the string $[0, 0, \ldots, 0]$, i.e. by the empty set. It is easy to observe that we obtain a probability distribution over extensions of $\mathcal{A}$.

Example 3. *Given a KB=($\mathcal{T}, \mathcal{A}$), let the only typicality inclusions in $\mathcal{T}$ be:*

$\mathbf{T}(C) \sqsubseteq_{0.6} D$
$\mathbf{T}(E) \sqsubseteq_{0.85} F$.

Let a and b be the only individual constants occurring in $\mathcal{A}$. Suppose also that $\mathbf{T}(C)(a)$, $\mathbf{T}(C)(b)$, and $\mathbf{T}(E)(b)$ are entailed from KB in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$. We have that

$$\mathfrak{Tip}_{\mathcal{A}} = \{(a, C), (b, C), (b, E)\}$$

and

$$\mathcal{P}_{\mathcal{A}} = [0.6, 0.6, 0.85].$$

All possible strings, corresponding extensions of $\mathcal{A}$ and probabilities are shown in Table 1.

Table 1: Plausible extensions of the ABox of Example 3.

String	Extension	Probability
[0.6, 0.6, 0.85]	$\widetilde{\mathcal{A}_1} = \{\mathbf{T}(C)(a), \mathbf{T}(C)(b), \mathbf{T}(E)(b)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_1}} = 0.6 \times 0.6 \times 0.85 = 0.306$
[0, 0, 0.85]	$\widetilde{\mathcal{A}_2} = \{\mathbf{T}(E)(b)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_2}} = (1-0.6) \times (1-0.6) \times 0.85 = 0.136$
[0, 0.6, 0]	$\widetilde{\mathcal{A}_3} = \{\mathbf{T}(C)(b)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_3}} = (1-0.6) \times 0.6 \times (1-0.85) = 0.036$
[0.6, 0, 0]	$\widetilde{\mathcal{A}_4} = \{\mathbf{T}(C)(a)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_4}} = 0.6 \times (1-0.6) \times (1-0.85) = 0.036$
[0, 0.6, 0.85]	$\widetilde{\mathcal{A}_5} = \{\mathbf{T}(C)(b), \mathbf{T}(E)(b)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_5}} = (1-0.6) \times 0.6 \times 0.85 = 0.204$
[0.6, 0, 0.85]	$\widetilde{\mathcal{A}_6} = \{\mathbf{T}(C)(a), \mathbf{T}(E)(b)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_6}} = 0.6 \times (1-0.6) \times 0.85 = 0.204$
[0.6, 0.6, 0]	$\widetilde{\mathcal{A}_7} = \{\mathbf{T}(C)(a), \mathbf{T}(C)(b)\}$	$\mathbb{P}_{\widetilde{\mathcal{A}_7}} = 0.6 \times 0.6 \times (1-0.85) = 0.054$
[0, 0, 0]	$\widetilde{\mathcal{A}_8} = \emptyset$	$\mathbb{P}_{\widetilde{\mathcal{A}_8}} = (1-0.6) \times (1-0.6) \times (1-0.85) = 0.024$
	$\mathbb{P}_{\widetilde{\mathcal{A}_1}} + \mathbb{P}_{\widetilde{\mathcal{A}_2}} + \ldots + \mathbb{P}_{\widetilde{\mathcal{A}_8}} =$	1

3.2. Reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$

We are now ready to provide formal definitions for nonmonotonic entailment in the Description Logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$. Intuitively, given KB and a query F, we distinguish two cases:

- if F is an inclusion $C \sqsubseteq D$, then it is entailed from KB if it is minimally entailed from KB' in the nonmonotonic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, where KB' is obtained from KB by removing probabilities of exceptions, i.e. by replacing each typicality inclusion $\mathbf{T}(C) \sqsubseteq_p D$ with $\mathbf{T}(C) \sqsubseteq D$;

- if F is an ABox fact $C(a)$, then it is entailed from KB if it is entailed in the monotonic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ from the knowledge bases including the extensions of the ABox of Definition 4.

More in detail, we provide both (i) a notion of entailment restricted to scenarios whose probabilities belong to a given range and (ii), similarly to [14], a notion of probability of the entailment of a query $C(a)$, as the sum of the probabilities of all extensions from which $C(a)$ is so entailed.

Here below are the formal definitions of entailment of a query F in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$. Given a knowledge base KB and two real numbers p and q, we write KB $\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$ to represent that F follows – or is entailed – from KB restricting reasoning to scenarios whose probabilities range from p to q. We distinguish the case in which the query is a TBox inclusion from the one in which it is an ABox assertion.

Definition 5 (Entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$). *Given a KB=$(\mathcal{T}, \mathcal{A})$, two real numbers $p, q \in (0, 1]$, and a query F which is a TBox inclusion either $C \sqsubseteq D$ or $\mathbf{T}(C) \sqsubseteq D$, we say that F is entailed from KB in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ in range $\langle p, q \rangle$, written KB $\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$, if $(\mathcal{T}', \mathcal{A}) \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{RaCl}} F$, where $\mathcal{T}' = \{\mathbf{T}(C) \sqsubseteq D \mid \mathbf{T}(C) \sqsubseteq_r D \in \mathcal{T}\} \cup \{C \sqsubseteq D \in \mathcal{T}\}$.*

Definition 6 (Entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$). *Given a KB=($\mathcal{T}, \mathcal{A}$), given $\mathfrak{Tip}$ a set of concepts, and given $p, q \in (0, 1]$, let $\mathcal{E} = \{\widetilde{A_1}, \widetilde{A_2}, \ldots, \widetilde{A_k}\}$ be the set of extensions of $\mathcal{A}$ of Definition 4 with respect to $\mathfrak{Tip}$, whose probabilities are such that $p \leq \mathbb{P}_1 \leq q, p \leq \mathbb{P}_2 \leq q, \ldots, p \leq \mathbb{P}_k \leq q$. Let $\mathcal{T}' = \{\mathbf{T}(C) \sqsubseteq D \mid \mathbf{T}(C) \sqsubseteq_r D \in \mathcal{T}\} \cup \{C \sqsubseteq D \in \mathcal{T}\}$. Given a query F which is an ABox assertion $C(a)$, where $a \in \mathcal{O}$, we say that F is entailed from KB in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ in range $\langle p, q\rangle$, written KB $\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$, if $(\mathcal{T}', \mathcal{A} \cup \widetilde{\mathcal{A}_i}) \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}} F$ for all $\widetilde{\mathcal{A}_i} \in \mathcal{E}$.*
We also define the probability of the entailment of a query as $\mathbb{P}(F) = \sum\limits_{i=1}^{k} \mathbb{P}_i$.

It is worth noticing that, in Definition 5, probabilities p and q do not play any role: indeed, probabilities of scenarios are related to ABox extensions, that are not involved when we are reasoning about TBoxes. As already mentioned, in this case entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ corresponds to entailment in the nonmonotonic Description Logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$.

4. A Decision Procedure for Reasoning with Probabilities in Description Logics

In this section we describe a decision procedure for reasoning in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, in order to check whether a query F is entailed from a given KB as in Definitions 5 and 6. We then exploit such decision procedure to show that the problem of entailment in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ is in EXPTIME. This allows us to conclude that reasoning about typicality and defeasible inheritance with probabilities of exceptions is essentially inexpensive, in the sense that reasoning retains the same complexity class of the underlying standard Description Logic $\mathcal{ALC}$, which is known to be EXPTIME-complete [1].

Given an $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ KB=($\mathcal{T}, \mathcal{A}$) and a query F, we define a procedure computing the following four steps:

1. compute the set $\mathfrak{Tip}_a$ of all typicality assumptions that are minimally entailed from the knowledge base in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$;
2. compute all possible $\widetilde{A_i}$ extensions of the ABox and compute their probabilities;
3. select the extensions whose probabilities belong to a given range $\langle p, q\rangle$;
4. check whether the query F is entailed from all the selected extensions in the monotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$.

Step 4 is based on reasoning in the monotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$: to this aim, the procedure relies on a polynomial encoding of $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ into $\mathcal{ALC}$ introduced in [22]. Step 1 is based on reasoning in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$: in this case, the procedure computes the rational closure of an $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ knowledge base by means of the algorithm introduced in [11], which is sound and complete with respect to the minimal model semantics recalled in Section 2. Also the algorithm computing the rational closure relies on reasoning in the monotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$, then on the above mentioned polynomial encoding in $\mathcal{ALC}$. We first recall the procedures for reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ and $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, then we describe the overall procedure for reasoning in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$.

4.1. Reasoning in $\mathcal{ALC} + \mathbf{T_R}$

In order to reason in $\mathcal{ALC} + \mathbf{T_R}$, in [22] the authors provide the following polynomial encoding in standard $\mathcal{ALC}$ of KB[4]. The idea on which the encoding is based exploits the definition of the typicality operator $\mathbf{T}$ in terms of a Gödel-Löb modality $\Box$ as follows: $\mathbf{T}(C)$ is defined as $C \sqcap \Box \neg C$ where the accessibility relation of the modality $\Box$ is the preference relation $<$ in $\mathcal{ALC} + \mathbf{T_R}$ models.

Let KB=$(\mathcal{T}, \mathcal{A})$ be a knowledge base where $\mathcal{A}$ does not contain positive typicality assertions on individuals of the form $\mathbf{T}(C)(a)$. The encoding KB'=$(\mathcal{T}', \mathcal{A}')$ of KB in $\mathcal{ALC}$ is defined as follows. First of all, we let $\mathcal{A}' = \emptyset$. Then, for each $A \sqsubseteq B \in \mathcal{T}$, not containing $\mathbf{T}$, we introduce $A \sqsubseteq B$ in $\mathcal{T}'$. For each $\mathbf{T}(A)$ occurring in $\mathcal{T}$, we introduce a new atomic concept $Box_{\neg A}$ and, for each inclusion $\mathbf{T}(A) \sqsubseteq_p B \in \mathcal{T}$, we add to $\mathcal{T}'$ the inclusion

$$A \sqcap Box_{\neg A} \sqsubseteq B.$$

In order to capture the properties of the $\Box$ modality, a new role R is introduced to represent the relation $<$ in preferential models, and the following inclusions are introduced in $\mathcal{T}'$:

$$Box_{\neg A} \sqsubseteq \forall R.(\neg A \sqcap Box_{\neg A})$$
$$\neg Box_{\neg A} \sqsubseteq \exists R.(A \sqcap Box_{\neg A})$$

The first inclusion accounts for the transitivity of $<$. The second inclusion accounts for the well-foundedness, namely the fact that if an element is not a typical A element then there must be a typical A element preferred to it. For the encoding of the inclusions, if $C_l \sqsubseteq C_r$ is not a typicality inclusion, then $C_l' = C_l$ and $C_r' = C_r$; if $C_l \sqsubseteq C_r$ is a typicality inclusion $\mathbf{T}(A) \sqsubseteq C_r$, then $C_l' = A \sqcap Box_{\neg A}$ and $C_r' = C_r$.

The size of KB' is polynomial in the size of the KB. The same for C_l' and C_r', assuming the size of C_l and C_r be polynomial in the size of KB.

Given the above encoding, in [22] it is shown that (we write KB $\models_{\mathcal{ALC}} F$ to mean that F holds in all $\mathcal{ALC}$ models of KB):

$$\text{KB} \models_{\mathcal{ALC}+\mathbf{T_R}} C_l \sqsubseteq C_r \text{ if and only if KB' } \models_{\mathcal{ALC}} C_l' \sqsubseteq C_r'$$

and, as a consequence, that the problem of deciding entailment in $\mathcal{ALC} + \mathbf{T_R}$ is in ExpTime, since reasoning in $\mathcal{ALC}$ is ExpTime-complete. ExpTime-hardness follows from the fact that $\mathcal{ALC} + \mathbf{T_R}$ includes $\mathcal{ALC}$. In conclusion, the problem of deciding entailment in $\mathcal{ALC} + \mathbf{T_R}$ is ExpTime-complete.

4.2. Reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$

We have mentioned that the semantics of the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ corresponds to the *rational closure* of an $\mathcal{ALC} + \mathbf{T_R}$ knowledge base introduced in [11]. Here we recall this machinery, essentially an extension to $\mathcal{ALC} + \mathbf{T_R}$ of the definition of rational closure introduced by Lehmann and Magidor in [10] for the propositional case. We first consider the rational closure with respect to the TBox, in which essentially we

[4]The results provided in [22] are extended to the more expressive logic $\mathcal{SHIQ}$. Here we focus our attention on the basic $\mathcal{ALC}$.

only consider which inclusions belong to the rational closure of KB. Next we will consider rational closure with respect to the ABox, in which we consider the individuals explicitly named in the ABox itself.

Definition 7 (Exceptionality). *Let KB=($\mathcal{T}, \mathcal{A}$) be a knowledge base. A concept C is said to be* exceptional *for KB if and only if KB $\models_{\mathcal{ALC}+\mathbf{T}_R} \mathbf{T}(\top) \sqsubseteq \neg C$. An inclusion $\mathbf{T}(C) \sqsubseteq D$ is exceptional for KB if C is exceptional for KB. The set of typicality inclusions of KB which are exceptional in KB are denoted as $\mathcal{E}$(KB).*

Similarly to the rational closure for propositional logic in [10], we introduce a sequence of knowledge bases, starting from the initial one, KB, in order to iteratively use exceptionality in the construction of the rational closure. At each step, in order to reason about the following exceptional subset of KB, we remove the inclusions $\mathbf{T}(C) \sqsubseteq D$ of KB that are not exceptional for KB. Before we do this, if there is an assertion $\mathbf{T}(C)(a) \in \mathcal{A}$, we add to a all the typical properties of C that we are removing. In order to reason in the same way for equivalent concepts, we need the slightly more complicated formulation of $\mathcal{A}_i$ below.

Definition 8. *Given KB=($\mathcal{T}, \mathcal{A}$), it is possible to define a sequence of knowledge bases $E_0, \ldots, E_i, \ldots, E_n$ by letting $E_0 = (\mathcal{T}_0, \mathcal{A}_0)$ where $\mathcal{T}_0 = \mathcal{T}$ and $\mathcal{A}_0 = \mathcal{A}$ and, for $i > 0$, $E_i = (\mathcal{T}_i, \mathcal{A}_i)$ where*

- $\mathcal{T}_i = \mathcal{E}(E_{i-1}) \cup \{C \sqsubseteq D \in \mathcal{T} \mid \mathbf{T}$ *does not occur in* $C\}$
- $\mathcal{A}_i = \mathcal{A}_{i-1} \cup \{(\neg C \sqcup D)(a) \mid \mathbf{T}(C) \sqsubseteq D$ *in* $(E_{i-1} - E_i)$ *and there is a* $\mathbf{T}(B)(a) \in \mathcal{A}$ *such that* $E_{i-1} \not\models_{\mathcal{ALC}+\mathbf{T}_R} \mathbf{T}(\top) \sqsubseteq \neg B$ *and* $E_j \models_{\mathcal{ALC}+\mathbf{T}_R} \mathbf{T}(\top) \sqsubseteq \neg B$ *for all* $j < i-1\}$

(as a consequence of the next Definition 9, these are the Bs such that $rank(B) = i-1$).

Clearly $\mathcal{T}_0 \supseteq \mathcal{T}_1 \supseteq \mathcal{T}_2, \ldots$, while $\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2, \ldots$ Observe that, being KB finite, there is a least $n \geq 0$ such that, for all $m > n, \mathcal{T}_m = \mathcal{T}_n$ or $\mathcal{T}_m = \emptyset$. We take $(\mathcal{T}_n, \mathcal{A}_n)$ as the last element of the sequence of knowledge bases starting from KB.

Informally, for the definition of $\mathcal{A}_i$, if $\mathbf{T}(B)(a) \in \mathcal{A}$ (i.e., a is a typical B-element), and B has rank $i-1$, then, for all the inclusions $\mathbf{T}(C) \sqsubseteq D$ in $(E_{i-1} - E_i)$, since C has also rank $i-1$ we have that: if a is a C-element, then it is a typical C-element and the assertion $(\neg C \sqcup D)(a)$ must hold.

Definition 9 (Rank of a concept). *A concept C has* rank *i (denoted by $rank(C) = i$) for KB=($\mathcal{T}, \mathcal{A}$), if and only if i is the least natural number for which C is not exceptional for E_i. If C is exceptional for all E_i then $rank(C) = \infty$, and we say that C has no rank.*

Consider the least $n \geq 0$ such that, for all $m > n, \mathcal{T}_m = \mathcal{T}_n$ or $\mathcal{T}_m = \emptyset$. Then from the above definition it follows that if a concept C has a rank, its highest possible value is n. The notion of rank of a formula allows one to define the rational closure of a knowledge base KB with respect to the TBox.

Definition 10 (Rational closure of TBox). *Given $KB = (\mathcal{T}, \mathcal{A})$, we define the* rational closure $\overline{\mathcal{T}}$ *of* $\mathcal{T}$, *as* $\overline{\mathcal{T}} = \{\mathbf{T}(C) \sqsubseteq D \mid$ *either* $rank(C) < rank(C \sqcap \neg D)$ *or* $rank(C) = \infty\} \cup \{C \sqsubseteq D \mid KB \models_{\mathcal{ALC}+\mathbf{T}_R} C \sqsubseteq D\}$.

Let us now consider the rational closure of the ABox as defined in [11]:

Definition 11 (Rational closure of ABox). *Given $KB = (\mathcal{T}, \mathcal{A})$, let $a_1, \dots, a_m$ be the individuals explicitly named in $\mathcal{A}$. Let $k_1, k_2, \dots, k_h$ be all the possible rank assignments to the individuals occurring in $\mathcal{A}$.*

- *Given a rank assignment k_j we define:*
  - *for each a_i: $\mu_i^j = \{(\neg C \sqcup D)(a_i)$ s.t. $\mathbf{T}(C) \sqsubseteq D$ in $\overline{\mathcal{T}}$, and $k_j(a_i) = rank(C)\} \cup \{(\neg C \sqcup D)(a_i)$ s.t. $C \sqsubseteq D \in \mathcal{T}\}$;*
  - *let $\mu^j = \mu_1^j \cup \ldots \cup \mu_m^j$ for all $\mu_1^j \ldots \mu_m^j$ just calculated for all $a_1, \dots, a_m$ in $\mathcal{A}$*
- *We say that k_j is consistent with $(\overline{\mathcal{T}}, \mathcal{A})$ if:*
  - *if $\mathbf{T}(C)(a_i) \in \mathcal{A}$, then $k_j(a_i) = rank(C)$;*
  - *$\mathcal{T} \cup \mathcal{A} \cup \mu^j$ is consistent in $\mathcal{ALC} + \mathbf{T_R}$;*
- *We say that k_j is minimal and consistent with $(\overline{\mathcal{T}}, \mathcal{A})$ if k_j is consistent with $(\overline{\mathcal{T}}, \mathcal{A})$ and there is no k_i consistent with $(\overline{\mathcal{T}}, \mathcal{A})$ s.t. for all a_i, $k_i(a_i) \leq k_j(a_i)$ and for some b, $k_i(b) < k_j(b)$.*
- *The rational closure of $\mathcal{A}$ ($\overline{\mathcal{A}}$) is the set of all assertions derivable in $\mathcal{ALC} + \mathbf{T_R}$ from $\mathcal{T} \cup \mathcal{A} \cup \mu^j$ for all minimal consistent rank assignments k_j: $\overline{\mathcal{A}} = \bigcap_{k_j \textit{minimal consistent}} \{C(a) \mid \mathcal{T} \cup \mathcal{A} \cup \mu^j \models_{\mathcal{ALC}+\mathbf{T}_R} C(a)\}$.*

In [11] it is shown that, given a knowledge base KB=($\mathcal{T}, \mathcal{A}$), the semantics based on rational models is equivalent with the above notion of rational closure of KB, namely:

- given an inclusion $C \sqsubseteq D$, KB $\models_{\mathcal{ALC}} + \mathbf{T}_{\mathbf{R}}^{RaCl} C \sqsubseteq D$ if and only if $C \sqsubseteq D \in \overline{\mathcal{T}}$
- given an ABox fact $C(a)$, KB $\models_{\mathcal{ALC}} + \mathbf{T}_{\mathbf{R}}^{RaCl} C(a)$ if and only if $C(a) \in \overline{\mathcal{A}}$.

Moreover, it is shown that the problem of deciding whether $\mathbf{T}(C) \sqsubseteq D \in \overline{\mathcal{T}}$ is in EXPTIME and that an individual constant a and a concept C, the problem of deciding whether $C(a) \in \overline{\mathcal{A}}$ is EXPTIME-complete.

4.3. Reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$: the overall procedure

Let us finally introduce the overall procedure for reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ and then let us analyze its complexity.

Let KB=($\mathcal{T}, \mathcal{A}$) be an $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ knowledge base. Let $\mathcal{T}'$ be the set of inclusions of $\mathcal{T}$ without probabilities of exceptions: $\mathcal{T}' = \{\mathbf{T}(C) \sqsubseteq D \mid \mathbf{T}(C) \sqsubseteq_r D \in \mathcal{T}\} \cup \{C \sqsubseteq D \in \mathcal{T}\}$, that the procedure will consider in order to reason in $\mathcal{ALC} + \mathbf{T_R}$ and $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$ for checking query entailment and finding all plausible typicality assumptions, respectively. Other inputs of the procedure are the finite set of concepts $\mathfrak{Tip}$, a query F, and two real numbers $p, q \in (0, 1]$ describing a range of probabilities. If F is an inclusion $C \sqsubseteq D$ (where C could be $\mathbf{T}(C')$), we just need to check whether $(\mathcal{T}', \mathcal{A}) \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}} C \sqsubseteq D$ in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$. If F is an ABox formula of the form $C(a)$, Algorithm 1 builds all possible scenarios, computes their probabilities and then checks whether KB $\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$ if F holds in all those scenarios having a probability between p and q.

Algorithm 1 Entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$: KB $\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$

---

procedure ENTAILMENT($(\mathcal{T}, \mathcal{A}), \mathcal{T}', F, \mathfrak{Tip}, p, q$)
  if F is of the form $C \sqsubseteq D$ **then** ▷ *If F is an inclusion, we rely on $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$ for entailment*
    return $(\mathcal{T}', \mathcal{A}) \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}} F$
  ▷ *Otherwise, F is an ABox assertion of the form $C(a)$*
  $\mathfrak{Tip}_{\mathcal{A}} \leftarrow \emptyset$ ▷ build the set $\mathbb{S}$ of possible assumptions
  for each $C \in \mathfrak{Tip}$ **do**
    for each individual $a \in \mathcal{A}$ **do** ▷ Reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$
      if $(\mathcal{T}', \mathcal{A}) \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}} \mathbf{T}(C)(a)$ **then** $\mathfrak{Tip}_{\mathcal{A}} \leftarrow \mathfrak{Tip}_{\mathcal{A}} \cup \{\mathbf{T}(C)(a)\}$
  $\mathcal{P}_{\mathcal{A}} \leftarrow \emptyset$ ▷ compute the probabilities of Definition 2 given $\mathcal{T}$ and $\mathfrak{Tip}_{\mathcal{A}}$
  for each $C \in \mathfrak{Tip}$ **do**
    $\Pi_C \leftarrow 1$
    for each $\mathbf{T}(C) \sqsubseteq_p D \in \mathcal{T}$ **do** $\Pi_C \leftarrow \Pi_C \times p$
    $\mathcal{P}_{\mathcal{A}} \leftarrow \mathcal{P}_{\mathcal{A}} \cup \Pi_C$
  $\mathbb{S} \leftarrow$ *build strings of possible assumptions as in Definition 3 given $\mathfrak{Tip}_{\mathcal{A}}$ and $\mathcal{P}_{\mathcal{A}}$*
  $\mathcal{E} \leftarrow \emptyset$ ▷ build extensions of $\mathcal{A}$
  for each $s_i \in \mathbb{S}$ **do**
    build the extension $\widetilde{\mathcal{A}_i}$ corresponding to s_i and compute $\mathbb{P}_{\widetilde{\mathcal{A}_i}}$ as in Definition 4
    if $p \leq \mathbb{P}_{\widetilde{\mathcal{A}_i}} \leq q$ **then** $\mathcal{E} \leftarrow \mathcal{E} \cup \widetilde{\mathcal{A}_i}$ ▷ select extensions with probability in $\langle p, q \rangle$
  for each $\widetilde{\mathcal{A}_i} \in \mathcal{E}$ **do** ▷ query entailment in $\mathcal{ALC} + \mathbf{T_R}$
    if $(\mathcal{T}', \mathcal{A} \cup \widetilde{\mathcal{A}_i}) \not\models_{\mathcal{ALC}+\mathbf{T_R}} F$ **then return** KB $\not\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$
  return KB $\models^{\langle p,q \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$ ▷ F is entailed in all extensions

---

In the following example we focus on ABox reasoning in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$.

Example 4. *Let us consider a KB whose TBox is:*

$$\mathbf{T}(\mathit{ItalianTeenAger}) \sqsubseteq_{0.4} \exists \mathit{listenTo}.\mathit{TrapMusic}$$
$$\mathbf{T}(\mathit{ItalianTeenAger}) \sqsubseteq_{0.8} \mathit{SoccerLover}$$
$$\mathbf{T}(\mathit{Student}) \sqsubseteq_{0.75} \neg \mathit{TaxPayer}$$

and whose ABox is:

$$\{ItalianTeenAger(fabrizio), Student(fabrizio)\}$$

In the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ we have four different scenarios, combining the assumptions:

$$\mathbf{T}(ItalianTeenAger)(fabrizio)$$
$$\mathbf{T}(Student)(fabrizio)$$

As an example, we have that

$$KB \models^{\langle 0.5,1\rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} \neg TaxPayer(fabrizio),$$

whereas we have that

$$KB \not\models^{\langle 0.5,1\rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} \exists listenTo.TrapMusic(fabrizio).$$

As another example, we have that

$$KB \not\models^{\langle 0.01,0.2\rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} \neg TaxPayer(fabrizio),$$

whereas in the underlying nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$ we cannot restrict reasoning to suitable scenarios and, therefore, we have that

$$KB \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}} \exists listenTo.TrapMusic(fabrizio)$$

as well as

$$KB \models_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}} \neg TaxPayer(fabrizio).$$

We exploit the procedure of Algorithm 1 to show that the problem of entailment in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ is ExpTime complete. This allows us to conclude that reasoning about typicality and defeasible inheritance with probabilities of exceptions is essentially inexpensive, since reasoning retains the same complexity class of the underlying standard $\mathcal{ALC}$, which is known to be ExpTime-complete [1].

Theorem 2 (Complexity of entailment). *Given a KB in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, real numbers $p, q \in (0, 1]$ and a query F whose size is polynomial in the size of KB, the problem of checking whether KB $\models^{\langle p,q\rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} F$ is* ExpTime*-complete.*

Proof 1. *Let n be the size of KB, i.e. the length of the string representing it. Consider the operations computed by Algorithm 1. First, in line 2 the procedure distinguishes the following two cases:*
1. the query F is a TBox inclusion of the form $C \sqsubseteq D$, including the case in which C is a typicality inclusion $\mathbf{T}(C')$: as mentioned before, in this case the procedure relies on reasoning in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$, and checks whether such an inclusion belongs to the rational closure of the knowledge base. In [11] it is shown that query entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathit{RaCl}}$ is in ExpTime*, and we are done;*
2. the query F is an ABox assertion of the form $C(a)$, and the algorithm proceeds as follows:

- *lines 4-7: the algorithm checks, for each concept $C \in \mathfrak{Tip}$ and for each individual name a of the ABox whether $\mathbf{T}(C)(a)$ is minimally entailed from the KB in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$. The number of individual names in the ABox is $O(n)$. We have assumed that $\mathfrak{Tip}$ contains only concepts belonging to KB, therefore also the size of $\mathfrak{Tip}$ is $O(n)$. It follows that the number of different $\mathbf{T}(C)(a)$ considered is $O(n^2)$. For each $\mathbf{T}(C)(a)$ the algorithm relies on reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, which is in* EXPTIME*, therefore we make a polynomial number of computations in* EXPTIME*;*

- *lines 8-12: the algorithm builds the ordered multiset $\mathcal{P}_{\mathcal{A}}$ of Definition 2: obviously, this operation consists in computing the product of the probabilities of the inclusions $\mathbf{T}(C) \sqsubseteq_p D \in \mathcal{T}$, which are $O(n)$, for each $C \in \mathfrak{Tip}$, again $O(n)$. Therefore, this problem can be solved with $O(n^2)$ operations, i.e. in polynomial time;*

- *line 13: the algorithm builds the set $\mathbb{S}$ of possible assumptions (Definition 3). We have to consider all possible strings obtained by assuming (or not) each typicality assumption $\mathbf{T}(C)(a)$, that are $O(n^2)$. Consider a generic string $\langle s_1, s_2, \dots, s_{n^2} \rangle$. For each s_i, we have two options: we can choose either to not include the corresponding typicality assumption, then $s_i = 0$, or to include it, then s_i corresponds to the probability p_i for that concept. So we can build $2 \times 2 \times \ldots \times 2$ different strings, therefore $O(2^{n^2})$, that is to say the multiset $\mathbb{S}$ has exponential size in n;*

- *lines 14-16: the algorithm builds the extensions of the ABox corresponding to strings of $\mathbb{S}$, again an exponential number of extensions ($O(2^{n^2})$);*

- *line 17: the algorithm selects extensions whose probabilities $\mathbb{P}_{\widetilde{\mathcal{A}_i}}$ are in the range $[p, q]$: again, since $\mathbb{S}$ has exponential size in n, this operation can be solved in* EXPTIME*;*

- *steps 18-20: the algorithm relies on reasoning in monotonic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ in order to check whether the query F is entailed in all extensions in $\mathcal{E}$. Since the size of $\mathcal{E}$ is $O(2^n)$, we have $O(2^n)$ call to query entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$, which is an* EXPTIME*-complete problem. It is worth noticing that the size of the KB is augmented by the size of the extension $\widetilde{\mathcal{A}}_i$, which is however polynomial in n, precisely $O(n^2)$. We can conclude that these operations are in* EXPTIME.

EXPTIME *hardness immediately follows from the fact that the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ extends standard $\mathcal{ALC}$, which is* EXPTIME*-complete [1]. Indeed, we can consider a knowledge base without the $\mathbf{T}$ operator (therefore, without probabilities), and consider $\mathfrak{Tip} = \emptyset$.*
□

Let us now conclude Example 1 introduced in Section 3 in the light of the definitions provided above.

Example 5 (Reasoning in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ part 2). *Suppose that the ABox is*

$$\mathcal{A} = \{Bipolar(john), ProstateCancerPatient(greg)\},$$

we can consider two typicality assumptions:

(a) $\mathbf{T}(Bipolar)(john)$
(b) $\mathbf{T}(ProstateCancerPatient)(greg)$

then we can distinguish among four different extensions:

(i) *both* (a) *and* (b) *are assumed: in this scenario, whose probability is* $0.7 \times (0.6 \times 0.8) = 0.336$*, we conclude that both John and Greg have mood swings, and that Greg has nocturia;*
(ii) *we assume* (b) *but not* (a)*: this scenario has probability* $(1-0.7) \times (0.6 \times 0.8) = 0.144$*, and we can only conclude* $\exists hasSymptom.MoodReactivity(greg)$ *and* $\exists hasSymptom.Nocturia(greg)$*;*
(iii) *we assume* (a) *and not* (b)*: this scenario, having a probability* $0.7 \times (1-(0.6 \times 0.8)) = 0.364$*, allows us to conclude* $\exists hasSymptom.MoodReactivity(john)$*;*
(iv) *neither* (a) *nor* (b) *is added to* $\mathcal{A}$*: here the probability is* $(1-0.7) \times (1-(0.6 \times 0.8)) = 0.156$*, but we are not able to conclude anything about John and Greg.*

The probability that John has mood swings is defined as the sum of the probabilities of scenarios where such inference can be performed, namely scenarios (1) and (3), and it is therefore $0.336 + 0.364 = 0.7$.

Let us conclude this section with a further example that suggests another possible application of the Description Logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ in order to find a plausible but not trivial medical *diagnosis* to explain patients' symptoms and signs. In medical diagnosis, the most likely explanation for a set of symptoms is not always the solution to the problem, whereas reasoning about scenarios whose probabilities are such that they can be considered as plausible, but not the most probable/obvious could help the medical staff in taking alternative explanations into account. In the following example we exploit the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ in order to formulate a plausible diagnosis in order to explain the symptoms of a patient, as an alternative to the most obvious one.

Example 6. *Let us consider again the KB=($\mathcal{T}, \mathcal{A}$) of Example 1, that we recall and extend here for the sake of readability: $\mathcal{T}$ is as follows:*

$$Bipolar \sqsubseteq Depressed \quad (1)$$
$$\mathbf{T}(Depressed) \sqsubseteq_{0.85} \neg\exists hasSymptom.MoodReactivity \quad (2)$$
$$\mathbf{T}(Bipolar) \sqsubseteq_{0.7} \exists hasSymptom.MoodReactivity \quad (3)$$
$$\mathbf{T}(ProstateCancerPatient) \sqsubseteq_{0.6} \exists hasSymptom.MoodReactivity \quad (4)$$
$$\mathbf{T}(ProstateCancerPatient) \sqsubseteq_{0.8} \exists hasSymptom.Nocturia \quad (5)$$
$$\mathbf{T}(Depressed) \sqsubseteq_{0.65} Smart \quad (6)$$

whereas $\mathcal{A} = \{Depressed(greg), \neg Smart(greg)\}$.

Let us consider a set $\mathcal{V}$ of formulas of the form $C(a)$ representing patients' symptoms and signs. For instance, let $\mathcal{V}$ describe Greg's symptom, in particular that he has mood reactivity:

$$\mathcal{V} = \{\exists hasSymptom.MoodReactivity(greg)\}.$$

We have that $\mathcal{V}$ is not entailed by KB, but KB $\cup \mathcal{V}$ is consistent. Indeed, in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ we have that Greg is a bipolar person having mood swings, and this is consistent with all the inclusions in $\mathcal{T}$.

We are then interested in finding a diagnosis for Greg's symptoms, that is to say a set of assertions $\mathcal{D}$ such that $\mathcal{V}$ follows from KB $\cup \mathcal{D}$. For instance:

$$\mathcal{D}_1 = \{Bipolar(greg)\},$$

but also

$$\mathcal{D}_2 = \{ProstateCancerPatient(greg)\}$$

as well as

$$\mathcal{D}_3 = \{Bipolar(greg), ProstateCancerPatient(greg)\}$$

are examples of diagnosis, explaining those symptoms.

We exploit the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ in order to describe surprising/not trivial/not obvious *diagnosis, in order to suggest an alternative iter that could suggest further investigations in case the most plausible explanation is not the correct one.*

Let us first consider the set of typicality assumptions that can be entailed in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$. We have that:

- $\mathbf{T}(Bipolar)(greg)$ *is entailed from KB $\cup \mathcal{D}_1$ and KB $\cup \mathcal{D}_3$*
- $\mathbf{T}(ProstateCancerPatient)(greg)$ *is entailed from KB $\cup \mathcal{D}_2$ and KB $\cup \mathcal{D}_3$*
- $\mathbf{T}(Depressed)(greg)$ *is* not *entailed from* any *knowledge base, since assuming that Greg is a typical depressed person would necessarily imply that he is smart (by inclusion (6)), and this is inconsistent with the information of $\mathcal{A}$ that Greg is not ($\neg Smart(greg)$).*

In the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ we can reason about the following scenarios. Let us consider $\mathcal{D}_3$, and let $\mathfrak{Tip}_{\mathcal{A}}$ be as follows:

$$(greg, Bipolar), (greg, ProstateCancerPatient)$$

We have also $\mathcal{P}_{\mathcal{A}} = [0.7, 0.48]$, where 0.7 is the probability equipping the only typical property of the concept $Bipolar$, whereas $0.48 = 0.6 \times 0.8$ (0.6 and 0.8 equip the two typicality inclusions of the concept $ProstateCancerPatient$).

We can reason about the following scenarios:

$\widetilde{\mathcal{A}_1^3} = \{\mathbf{T}(Bipolar)(greg)\}$, *with* $\mathbb{P}_{\widetilde{\mathcal{A}_1^3}} = 0.7 \times 0.52 = 0.364$

$\widetilde{\mathcal{A}_2^3} = \{\mathbf{T}(ProstateCancerPatient)(greg)\}$, *with* $\mathbb{P}_{\widetilde{\mathcal{A}_2^3}} = 0.3 \times 0.48 = 0.144$

$\widetilde{\mathcal{A}_3^3} = \{\mathbf{T}(Bipolar)(greg), \mathbf{T}(ProstateCancerPatient)(greg)\}$, *with* $\mathbb{P}_{\widetilde{\mathcal{A}_3^3}} = 0.7 \times 0.48 = 0.336$.

In the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, we have that

$$KB \cup \mathcal{D}_i \models^{\langle 0,1 \rangle}_{\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}} \exists hasSymptom.MoodReactivity(greg)$$

for $i = 1, 2, 3$*, that is to say all the above set of assertions represent a diagnosis for the symptom* $\mathcal{V} = \{\exists hasSymptom.MoodReactivity(greg)\}$.

The logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ *could suggest an alternative – plausible but* not obvious *– diagnosis, in case of a failure of the most probable one. In this respect, the extension* $\widetilde{\mathcal{A}_2^3}$ *with the lowest probability could suggest that Greg has prostate cancer, and such a non trivial diagnosis could be confirmed by an evaluation of other typical symptoms of such a disease (e.g. nocturia).*

5. Discussion and Conclusions

In this work we have introduced the Description Logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, which extends the nonmonotonic Description Logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ by means of probabilities equipping typicality inclusions. In this setting, $\mathbf{T}(C) \sqsubseteq_p D$ means that "normally, Cs are Ds and we have a probability of $1 - p$ of having exceptional Cs not being Ds". From a knowledge representation point of view, as a difference from $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ allows one to distinguish among typicality inclusions by means of their probabilities: given two typical properties D_1 and D_2 of a given concept C, one can formalize the fact that the probability of having exceptional elements of C with respect to the property D_1 is higher than the one of having exceptional ones with respect to D_2 by means of a pair of inclusions

$$\mathbf{T}(C) \sqsubseteq_{p_1} D_1$$
$$\mathbf{T}(C) \sqsubseteq_{p_2} D_2$$

where $p_1 < p_2$.

Probabilities of exceptions are then used in order to reason about plausible scenarios, obtained by selecting only some – i.e., not necessarily all – typicality assumptions and whose probabilities belong to a given and fixed range. We have also introduced a decision procedure for reasoning in the Description Logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, and we have exploited it in order to estimate the complexity of the proposed logic; in detail we have shown that reasoning in DLs with rational closure and probabilities of exceptions is essentially inexpensive, in the sense that the complexity of entailment in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ remains in EXPTIME as in the underlying standard Description Logic $\mathcal{ALC}$.

It is worth noticing that the proposed logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ is not intended to replace existing extensions of DLs for representing and reasoning about prototypical properties and defeasible inheritance. The idea is that, in some applications, the need of reasoning about probabilities of exceptions and to restrict reasoning to plausible – but not necessarily the most probable – scenarios could help domain experts to achieve their goals, wherever standard reasoning is not enough to do it.

The logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, as well as the underlying $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, are based on the rational closure, then they inherit its virtues, but also its weakness. It is well known that the main advantage of the rational closure is related to its good computational properties. However, rational closure is affected by the "all or nothing" behavior, in the sense that it does not allow one to separately reason about the inheritance of different properties. For instance, let us recall the example in the Introduction: we have that typical students are sport lovers, and normally they are also social network users. Furthermore, we can consistently express that, normally, third age students are not social network users. As a consequence, third age students are recognized as untypical students, then no inheritance of typical properties is possible, for instance it is not possible to infer that they are sport lovers. The problem also affects the definition of scenarios in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$: if $\mathbf{T}(Student)(gary)$ is a typicality assumption to be considered in the construction of different scenarios (since it is entailed in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ from the knowledge base), then Gary inherits all the properties of typical students. On the contrary, if $\mathbf{T}(ThirdAgeStudent)(gary)$ is the typicality assertion to be considered for the scenarios generation, no inheritance of typical students is possible for Gary. In order to solve this problem, a strengthening of a rational closure-like algorithm with defeasible inheritance networks has been studied by [23]. In [24] the author has proposed an alternative semantics by considering models equipped with multiple preference relations, whence with multiple "typicality" operators. In this variant, it should be possible to distinguish different aspects of typicality/exceptionality and consequently to avoid the "all or nothing" behavior of rational closure with respect to property inheritance.

5.1. Related Works

The recent literature is rich of sophisticated approaches introduced in order to tackle the problem of reasoning under probabilistic uncertainty in Description Logics and ontologies, following the need of reasoning about vague and incomplete information available from web resources. Among them, a work that can be considered as strongly related to the one presented in this paper is [21], where the author introduces two probabilistic extensions of expressive Description Logics $\mathcal{SHIF}(\mathbf{D})$ and $\mathcal{SHOIN}(\mathbf{D})$. These extensions are semantically based on the notion of probabilistic lexicographic entailment [25] and allow to represent and reason about prototypical properties of classes that are semantically interpreted as lexicographic entailment introduced by Lehmann from conditional knowledge bases. Intuitively, the basic idea is to interpret inclusions of the TBox and facts in the ABox as probabilistic knowledge about random and concrete instances of concepts. As an example, an expression of the form

$$(SocialNetworkUser \mid Student)[0.7,\ 1]$$

represents that "typically, a randomly chosen student makes use of social networks with a probability of at least 70%", whereas default knowledge can be expressed as

$$(Young \mid Student)[1,\ 1] \qquad (*)$$

whose meaning is that prototypical students are young people (but we have no information about the probability of having or not exceptions). Obviously, also strict inclusions $C \sqsubseteq D$ are allowed.

In these extensions, we can also have ABox facts like

$$(Student \mid \top)[0.8,\ 1] \qquad (**)$$

for an individual name *chris*, representing that Chris is a student with a probability of at least 80%.

There are two main differences between the approach of [21] and our proposal:

- on the one hand, we have a significant difference in the meaning of the probability: here, the probability of an inclusion $C \sqsubseteq D$ is intended as the probability that an individual belonging to C also belongs to D, then admitting the presence of typical Cs not being Ds. As mentioned at the very beginning of the Introduction, in our framework the probability p equipping a typicality inclusion $\mathbf{T}(C) \sqsubseteq_p D$ is used to represent the probability of (not) finding exceptional Cs not being Ds, but *all* typical Cs are also Ds by the semantics of the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ underlying our logic. Furthermore, in our framework, each typicality inclusion is equipped by a probability, whereas in [21] we can have prototypical properties like $(*)$ where no probability is provided;

- reasoning about individuals in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ is based on the definition of an extension of an ABox, obtained by assuming (or not) typicality properties about the individuals themselves: given the individual name $chris$, we build different scenarios from the typicality assertions that can be *nonmonotonically* inferred from the knowledge base in the underlying *nonomonotonic* logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, for instance that Chris is a typical student $\mathbf{T}(Student)(chris)$ and a typical tennis player $\mathbf{T}(TennisPlayer)(chris)$. In the probabilistic extensions introduced in [21] one can express probabilities about facts of the ABox like in $(**)$, but reasoning about ABox facts is not related to typical properties of named individuals entailed by the knowledge base.

As the logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$, the lexicographic entailment defined in [21] inherits interesting and useful nonmonotonic properties from lexicographic entailment in [25], such as specificity, rational monotonicity and some forms of irrelevance. As mentioned above, the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ inherits, however, the main drawback of rational closure, namely the "all or nothing" behavior, whereas the notion of lexicographic entailment allows one to deal with overriding less specific properties without such inheritance blocking. On the contrary, in order to perform useful, stronger nonmonotonic inferences, the logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{RaCl}$ is obtained by adding an additional nonmonotonic machinery on top of the logic. It could be of interest to study a formal relation between these two approaches in order to reason about defeasible inheritance in Description Logics, as well as to evaluate the opportunity of using the latter as the basis for the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$.

Several other nonmonotonic extensions of DLs have been proposed in the literature in order to reason about inheritance with exceptions, essentially based on the integration of DLs with well established nonmonotonic reasoning mechanisms [2, 3, 4, 16, 5, 6, 8], ranging from Reiter's defaults to minimal knowledge and negation as failure. We remind to [9, 8] for a detailed discussion about extensions of DLs for defeasible inheritance, and to [16] for a formal and precise comparison between the approach based on the typicality operator $\mathbf{T}$ and circumscribed knowledge bases. In none of them, probability of exceptions in concept inclusions is taken into account, as far as we know.

Probabilistic extensions of DLs, allowing one to label inclusions (and facts) with degrees representing probabilities, have been introduced in [13, 14]. In this approach, called DISPONTE, the authors propose the integration of probabilistic information with DLs based on the distribution semantics for probabilistic logic programs [26].

The basic idea is to label inclusions of the TBox as well as facts of the ABox with a real number between 0 and 1, representing their probabilities, assuming that each axiom is independent from each others. The resulting knowledge base defines a probability distribution over *worlds*: roughly speaking, a world is obtained by choosing, for each axiom of the KB, whether it is considered as true of false. The distribution is further extended to queries and the probability of the entailment of a query is obtained by marginalizing the joint distribution of the query and the worlds. As an example, consider the following variant of the knowledge base inspired by the people and pets ontology in [14]:

$$0.3 \quad :: \quad \exists hasAnimal.Pet \sqsubseteq NatureLover \tag{1}$$
$$0.6 \quad :: \quad Cat \sqsubseteq Pet \tag{2}$$
$$0.9 \quad :: \quad Cat(tom) \tag{3}$$
$$hasAnimal(kevin, tom) \tag{4}$$

The inclusion (1) expresses that individuals that own a pet are nature lovers with a 30% probability, whereas (2) is used to state that cats are pets with probability 60%. The ABox fact (3) represents that Tom is a cat with probability 90%. Inclusions (1), (2) and (3) are *probabilistic* axioms, whereas (4) is a *certain* axiom, that must always hold. The KB has eight possible worlds, representing all possible combinations of considering/not considering each probabilistic axiom. For instance, the world $\{((1), 1), ((2), 0), ((3), 1)\}$ represents the situation in which we have that (1) and (3) hold, i.e. $\exists hasAnimal.Pet \sqsubseteq NatureLover$ and $Cat(tom)$, whereas (2) does not. The query $NatureLover(kevin)$ is true only in the last world, i.e. having that (1), (2) and (3) are all true, whereas it is false in all the other ones. The probability of such a query is $0.3 \times 0.6 \times 0.9 = 0.162$.

There are two main differences between the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ proposed in this work and probabilistic DLs. On the one hand, as already mentioned in the Introduction, in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ probabilities are used in order to express different degrees of admissibility of exceptions with respect to such typicality inclusions. Probabilities are then the basis of different scenarios built by assuming – or not – that individuals are typical instances of a given concept. On the contrary, in DISPONTE probabilities are used to capture a notion of uncertainty about information of the KB, therefore an inclusion $C \sqsubseteq D$ having a very low probability p has a significantly different meaning with respect to an inclusion $\mathbf{T}(C) \sqsubseteq_p D$, representing anyway a typical property: normally, Cs are Ds, even if with a high probability of having exceptions to such typical inclusion. On the other hand, in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ probabilities are restricted to typicality inclusions only. On the contrary, in DISPONTE probabilities can be associated to concept inclusions as well as to ABox facts.

In [20] a nonmonotonic procedure for reasoning about *surprising* scenarios in DLs has been proposed. In this approach, the Description Logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ is extended by inclusions of the form $\mathbf{T}(C) \sqsubseteq_d D$, where d is a *degree of expectedness*. Similarly to $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, a notion of extension of an ABox is introduced in order to assume typicality assertions about individuals satisfying cardinality restrictions on concepts, then degrees of expectedness are used in order to define a preference relation among extended ABoxes: entailment of queries is then restricted to ABoxes that are minimal with respect to such preference relations and that represent surprising scenarios. Also in this case, we have two main differences with the approach of the logic $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$: first, in $\mathcal{ALC} + \mathbf{T}^{\mathsf{exp}}_{\mathbf{R}}$ degrees of expectedness are non-negative integers used essentially to define a – partial – preference relation among extended ABoxes, whereas they are not used in order to estimate probabilities of typicality inclusions. Second, cardinality restrictions play a fundamental role in order to "filter" extended ABoxes. On the contrary, in the logic $\mathcal{ALC} + \mathbf{T}^{\mathsf{P}}_{\mathbf{R}}$, entailment is defined in terms of the probability of a given scenario and can be used to estimate the probability of a given query.

Several approaches in the literature exploit the well established paradigm of Answer Set Programming (ASP) to deal with incomplete information. Extensions of ASP making use of probabilities are proposed in [27, 28]. In [27] the authors introduce a declarative language, called P-log, which extends ASP by means of probabilistic constructs. More in detail, a P-log program contains *random attributes* - essentially, random variables – in addition to standard ASP statements: such random attributes have the form $a(X)$ where both X and the value of $a(X)$ range over finite domains. P-log is able to deal with nonmonotonic probabilistic inferences, namely an update of a P-log program/knowledge base can cause the generation of new possible worlds in the adopted probabilistic models. As an example, consider the following P-log program, inspired to the one proposed in [27]:

$$year : \{2018, 1978\}. \tag{1}$$
$$year = 2018 \longleftarrow \ not\ abnormal. \tag{2}$$
$$random(year) \longleftarrow \ abnormal. \tag{3}$$

In this program, $year$ has two possible values: 2018 and 1978 (statement 1). Rule 2 states that, in a typical/not abnormal situation, we are considering the current year, then the value of $year$ is 2018. Otherwise, we are in an untypical/abnormal situation, and rule 3 states that the value of $year$ will be randomly assigned. Since the program does not contain the atom $abnormal$, rule 2 allows the reasoner to conclude that $year$=2018, with a single possible world having probability 1. If we enrich the program by

$$abnormal. \tag{4}$$

rule 2 is no longer applicable, whereas rule 3 allows the reasoner to conclude that there are two possible worlds, one in which the value of $year$ is 2018 and another one in which the value of $year$ is 1978, both with a probability of 50%. As a difference with our approach, where probabilities aims at estimating probabilities of exceptions, in this work probabilities are intended as a measure of the degree of a belief of an agent.

In [28] the authors introduce an extension of ASP inspired by Markov Logic Networks (MLN) introduced in [29]. They propose a language, called LP^{MLN}, which combines the stable model semantics of ASP with the basic ideas underlying MLN, whose main aim is to combine first-order logic and probabilistic graphical models. In this work, the authors move this combination to the context of logic programming, where rules are equipped with *weights* that are closely related to probabilities of [27]. As the same authors point out, the language LP^{MLN} is strongly related to the language P-log, as well as to the language P$\mathcal{C}+$ introduced in [30] for probabilistic reasoning about actions.

Several works also discuss the combination of open and closed world reasoning in DLs. In particular, formalisms have been defined for combining DLs with logic programming rules (see, for instance, [31] and [32]). A grounded circumscription approach for DLs with local closed world capabilities has been defined in [33]. More in detail, in [32] the authors introduce the formalism of $MKNF^+$ knowledge bases, which allows for a flexible integration of DLs and Answer Set Programming. The semantics of the formalisms, based on the logic of MKNF [4], overcomes the discrepancy between the open world assumption of DLs and the closed world assumption of rules. [32] presents several algorithms for reasoning with $MKNF^+$ knowledge bases and establishes tight complexity results. In [31] the authors combine Answer Set Programming with the Description Logics $\mathcal{SHIF}(\mathbf{D})$ and $\mathcal{SHOIN}(\mathbf{D})$, introducing the notion of *description logic programs*, consisting in a DL knowledge base together with a generalized normal program P. While rule bodies may contain DL queries, nonmonotonicity is provided via negation-as-failure. [34] presents a non-monotonic extension of the description logic $\mathcal{SHROIQ}$ based on the logic MKNF, which encompasses some of the most prominent languages related to OWL, rules, non-monotonic reasoning, and their integrations. Given the relation among ASP and default logic, this approach has some similarities with the extensions of DLs based on defaults [35, 3]: the nonmonotonic inferences induced by program rules are limited to named individuals only. A common limitation of the nonmonotonic extensions of DLs based on minimal knowledge and negation as failure (including the integrations of DLs and rules) is that they provide no support for capturing specificity nor priorities. In [36], the authors exploit ASP for reasoning in an extension of the low-complexity Description Logic $\mathcal{SROEL}$ with the typicality operator $\mathbf{T}$ based on the rational closure. In order to strengthen the rational entailment, the authors consider a minimal model semantics. They rely on a Small Model result, where models correspond to answer sets of a suitable ASP encoding, and exploit Answer Set Preferences for reasoning under minimal entailment. They also provide complexity results for the problem of instance checking, which is Π_2^P-complete. As already mentioned at the very beginning of this section, none of these works take probabilities into account.

5.2. Future Works

In future work we aim at extending the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ to more expressive Description Logics, such as those underlying the standard language for ontology engineering OWL. As a first step, in [22] the logic with the typicality operator and the rational closure construction have been applied to the logic $\mathcal{SHIQ}$. Moreover, we aim at extending the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ with cardinality restrictions, in order to investigate the precise relation with the approach proposed in [20] and mentioned above.

We are currently developing a preliminary implementation of the reasoning machinery for the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, and a prototype will be soon available. The current version of the system is implemented in Pyhton and exploits the translation of an $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ knowledge base into standard $\mathcal{ALC}$ introduced in [11], summarized in Section 4.1 and adopted by the system RAT-OWL [37]. The system also makes use of the library `owlready2` [5] that allows one to rely on the services of efficient DL reasoners, e.g. the HermiT reasoner, in order to generate different scenarios and to reason about them as described in Section 4. As mentioned, this prototype represents a very preliminary attempt to implement reasoning services for the logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$, whereas a more mature version, obtained by investigating the application of techniques introduced in [38, 39] in order to improve its efficiency, will be addressed in future works.

As we have pointed out in the previous section, the approach based on the probabilities of typicality inclusions of the logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ and the DISPONTE semantics in [13, 14] could be combined in order to describe a probabilistic extension of DLs with typicalities and probabilities of having exceptions: a knowledge base can contain axioms labelled by probabilities that can be interpreted as "epistemic" ones, i.e. as degrees of our belief in those axioms, as in [13], as well as typicality inclusions with probabilities about exceptions. In this respect, an inclusion

$$p \ :: \ \mathbf{T}(C) \sqsubseteq_q D$$

represents that we have degree of belief p in the fact that typical Cs are also Ds with a probability q of not having exceptions. In this line of research, in [40, 41, 42, 43] we have introduced an extension of the logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ in order to tackle the problem of dealing with the composition of concepts in presence of prototypical properties: in this respect, the prototype of a compound concept cannot result from the composition of the prototypes of its components, take the *pet fish* as an example. It is well established in the literature that fuzzy-based approaches are not adequate to provide a solution to this problem, whereas an extension of the logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ with a DISPONTE semantics seems to be a good candidate for a solution to such problem. Moreover, combining the logic $\mathcal{ALC}+\mathbf{T}_{\mathbf{R}}^{\mathsf{P}}$ with the DISPONTE semantics should provide an alternative solution to the problem of the "all or nothing" behavior of rational closure with respect to property inheritance discussed above. Such a further extension will be material for future works.

Acknowledgements

I am extremely grateful to the anonymous referees for their careful reading and very constructive criticisms, which greatly helped me to improve the final version of this work.

References

[1] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, P. Patel-Schneider, The Description Logic Handbook - Theory, Implementation, and Applications, 2nd edition, Cambridge, 2007.

[2] P. A. Bonatti, C. Lutz, F. Wolter, The Complexity of Circumscription in DLs, Journal of Artificial Intelligence Research (JAIR) 35 (2009) 717–773.

[5] https://pythonhosted.org/Owlready2/

[3] F. Baader, B. Hollunder, Priorities on defaults with prerequisites, and their application in treating specificity in terminological default logic, Journal of Automated Reasoning (JAR) 15 (1) (1995) 41–68.

[4] F. M. Donini, D. Nardi, R. Rosati, Description logics of minimal knowledge and negation as failure, ACM Transactions on Computational Logics (ToCL) 3 (2) (2002) 177–225.

[5] G. Casini, U. Straccia, Rational closure for defeasible description logics, in: T. Janhunen, I. Niemelä (Eds.), Logics in Artificial Intelligence - Proceedings of the12th European Conference (JELIA 2010), Vol. 6341 of Lecture Notes in Computer Science (LNCS), Springer, 2010, pp. 77–90.

[6] G. Casini, U. Straccia, Defeasible Inheritance-Based Description Logics, Journal of Artificial Intelligence Research (JAIR) 48 (2013) 415–473.

[7] U. Straccia, Default inheritance reasoning in hybrid kl-one-style logics, in: Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI'93), Morgan Kaufmann, 1993, pp. 676–681.

[8] P. A. Bonatti, M. Faella, I. Petrova, L. Sauro, A new semantics for overriding in description logics, Artificial Intelligence 222 (2015) 1–48. doi:10.1016/j.artint.2014.12.010.

[9] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, ALC+T: a preferential extension of description logics, Fundamenta Informaticae 96 (2009) 341–372. doi:10.3233/FI-2009-185.

[10] D. Lehmann, M. Magidor, What does a conditional knowledge base entail?, Artificial Intelligence 55 (1) (1992) 1–60. doi:http://dx.doi.org/10.1016/0004-3702(92)90041-U.

[11] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, Semantic characterization of Rational Closure: from Propositional Logic to Description Logics, Artificial Intelligence 226 (2015) 1–33. doi:10.1016/j.artint.2015.05.001.

[12] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, Minimal model semantics and rational closure in description logics, in: T. Eiter, B. Glimm, Y. Kazakov, M. Krötzsch (Eds.), Informal Proceedings of the 26th International Workshop on Description Logics, Ulm, Germany, July 23 - 26, 2013, Vol. 1014 of CEUR Workshop Proceedings, CEUR-WS.org, 2013, pp. 168–180.
URL `http://ceur-ws.org/Vol-1014/paper_5.pdf`

[13] F. Riguzzi, E. Bellodi, E. Lamma, R. Zese, Reasoning with probabilistic ontologies, in: Q. Yang, M. Wooldridge (Eds.), Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015, AAAI Press, 2015, pp. 4310–4316.

[14] F. Riguzzi, E. Bellodi, E. Lamma, R. Zese, Probabilistic description logics under the distribution semantics, Semantic Web 6 (5) (2015) 477–501. doi:10.3233/SW-140154.

[15] G. L. Pozzato, Reasoning in description logics with typicalities and probabilities of exceptions, in: A. Antonucci, L. Cholvy, O. Papini (Eds.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 14th European Conference, ECSQARU 2017, Lugano, Switzerland, July 10-14, 2017, Proceedings, Vol. 10369 of Lecture Notes in Computer Science, Springer, 2017, pp. 409–420. doi:10.1007/978-3-319-61581-3_37.

[16] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, A NonMonotonic Description Logic for Reasoning About Typicality, Artificial Intelligence 195 (2013) 165 – 202. doi:10.1016/j.artint.2012.10.004.

[17] H. Rott, Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning, Oxford University Press, 2001.

[18] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, Analytic tableaux calculi for KLM logics of nonmonotonic reasoning, ACM Transactions on Computational Logic 10 (3) (2009) 18:1–18:47. doi:10.1145/1507244.1507248.
URL `https://doi.org/10.1145/1507244.1507248`

[19] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, Analytic tableaux for KLM preferential and cumulative logics, in: G. Sutcliffe, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning, 12th International Conference, LPAR 2005, Montego Bay, Jamaica, December 2-6, 2005, Proceedings, Vol. 3835 of Lecture Notes in Computer Science, Springer, 2005, pp. 666–681. doi:10.1007/11591191_46.
URL `https://doi.org/10.1007/11591191_46`

[20] G. L. Pozzato, Reasoning about surprising scenarios in description logics of typicality, in: G. Adorni, S. Cagnoni, M. Gori, M. Maratea (Eds.), Advances in Artificial Intelligence: Proceedings of the 15th International Conference of the Italian Association for Artificial Intelligence, Vol. 10037 of Lecture Notes in Artificial Intelligence LNAI, Springer, Genova, 2016, pp. 1–15. doi:10.1007/978-3-319-49130-1 31.

[21] T. Lukasiewicz, Expressive probabilistic description logics, Artificial Intelligence 172 (6-7) (2008) 852–883. doi:10.1016/j.artint.2007.10.017.

[22] L. Giordano, V. Gliozzi, N. Olivetti, G. L. Pozzato, Rational closure in $\mathcal{SHIQ}$., in: DL 2014, 27th International Workshop on Description Logics, Vol. 1193 of CEUR Workshop Proceedings, CEUR-WS.org, 2014, pp. 543–555.

[23] G. Casini, U. Straccia, Defeasible Inheritance-Based Description Logics, in: T. Walsh (Ed.), Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011), Morgan Kaufmann, Barcelona, Spain, 2011, pp. 813–818.

[24] V. Gliozzi, Reasoning about Multiple Aspects in Rational Closure for DLs, in: G. Adorni, S. Cagnoni, M. Gori, M. Maratea (Eds.), Advances in Artificial Intelligence: Proceedings of the 15th International Conference of the Italian Association for Artificial Intelligence, Vol. 10037 of Lecture Notes in Artificial Intelligence LNAI, Springer, Genova, 2016, pp. 392–405. doi:10.1007/978-3-319-49130-1 31.

[25] D. J. Lehmann, Another perspective on default reasoning, Annals of Mathematics and Artificial Intelligence 15 (1) (1995) 61–82. doi:10.1007/BF01535841.

[26] T. Sato, A statistical learning method for logic programs with distribution semantics, in: L. Sterling (Ed.), Logic Programming, Proc. of the 12th International Conference on Logic Programming, MIT Press, 1995, pp. 715–729.

[27] C. Baral, M. Gelfond, J. N. Rushton, Probabilistic reasoning with answer sets, TPLP 9 (1) (2009) 57–144. doi:10.1017/S1471068408003645.

[28] Y. Wang, J. Lee, Handling uncertainty in answer set programming, in: B. Bonet, S. Koenig (Eds.), Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA., AAAI Press, 2015, pp. 4218–4219.

[29] M. Richardson, P. M. Domingos, Markov logic networks, Machine Learning 62 (1-2) (2006) 107–136. doi:10.1007/s10994-006-5833-1.

[30] T. Eiter, T. Lukasiewicz, Probabilistic reasoning about actions in nonmonotonic causal theories, in: C. Meek, U. Kjærulff (Eds.), UAI '03, Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence, Acapulco, Mexico, August 7-10 2003, Morgan Kaufmann, 2003, pp. 192–199.

[31] T. Eiter, T. Lukasiewicz, R. Schindlauer, H. Tompits, Combining Answer Set Programming with Description Logics for the Semantic Web, in: D. Dubois, C. Welty, M. Williams (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the 9th International Conference (KR 2004), AAAI Press, Whistler, Canada, 2004, pp. 141–151.

[32] B. Motik, R. Rosati, Reconciling Description Logics and rules, Journal of the ACM 57 (5).

[33] A. A. Krisnadhi, K. Sengupta, P. Hitzler, Local closed world semantics: Keep it simple, stupid!, in: Proceedings of the 24th International Workshop on Description Logics (DL 2011), Vol. 745 of CEUR Workshop Proceedings, Barcelona, Spain, 2011, pp. 1–15.

[34] M. Knorr, P. Hitzler, F. Maier, Reconciling OWL and non-monotonic rules for the semantic web, in: ECAI 2012 - 20th European Conference on Artificial Intelligence. Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track, Montpellier, France, August 27-31 , 2012, Vol. 242 of Frontiers in Artificial Intelligence and Applications, IOS Press, 2012, pp. 474–479.

[35] F. Baader, B. Hollunder, Embedding defaults into terminological knowledge representation formalisms, Journal of Automated Reasoning (JAR) 14 (1) (1995) 149–180.

[36] L. Giordano, D. T. Dupré, ASP for minimal entailment in a rational extension of $\mathcal{SROEL}$, TPLP 16 (5-6) (2016) 738–754. doi:10.1017/S1471068416000399.

[37] L. Giordano, V. Gliozzi, G. L. Pozzato, R. Renzulli, An efficient reasoner for description logics of typicality and rational closure, in: A. Artale, B. Glimm, R. Kontchakov (Eds.), Proceedings of the 30th International Workshop on Description Logics, Montpellier, France, July 18-21, 2017, Vol. 1879 of CEUR Workshop Proceedings, CEUR-WS.org, 2017.
URL `http://ceur-ws.org/Vol-1879/paper25.pdf`

[38] M. Alberti, E. Bellodi, G. Cota, F. Riguzzi, R. Zese, cplint on SWISH: probabilistic logical inference with a web browser, Intelligenza Artificiale 11 (1) (2017) 47–64. doi:10.3233/IA-170106.
URL `https://doi.org/10.3233/IA-170106`

[39] E. Bellodi, E. Lamma, F. Riguzzi, R. Zese, G. Cota, A web system for reasoning with probabilistic OWL, Journal of Software: Practice and Experience 47 (1) (2017) 125–142. doi:10.1002/spe.2410.
URL `https://doi.org/10.1002/spe.2410`

[40] A. Lieto, G. L. Pozzato, A description logic of typicality for conceptual combination, in: M. Ceci, N. Japkowicz, J. Liu, G. A. Papadopoulos, Z. W. Ras (Eds.), Foundations of Intelligent Systems - 24th International Symposium, ISMIS 2018, Limassol, Cyprus, October 29-31, 2018, Proceedings, Vol. 11177 of Lecture Notes in Computer Science, Springer, 2018, pp. 189–199. doi:10.1007/978-3-030-01851-1_19.
URL `https://doi.org/10.1007/978-3-030-01851-1_19`

[41] A. Lieto, G. L. Pozzato, Creative concept generation by combining description logic of typicality, probabilities and cognitive heuristics, in: C. Ghidini, B. Magnini, A. Passerini, P. Traverso (Eds.), AI*IA 2018 - Advances in Artificial Intelligence - XVIIth International Conference of the Italian Association for Artificial Intelligence, Trento, Italy, November 20-23, 2018, Proceedings, Vol. 11298 of Lecture Notes in Computer Science, Springer, 2018, pp. 183–196. doi:10.1007/978-3-030-03840-3_14.
URL `https://doi.org/10.1007/978-3-030-03840-3_14`

[42] A. Lieto, G. L. Pozzato, A. Valese, COCOS: a typicality based concept combination system, in: P. Felli, M. Montali (Eds.), Proceedings of the 33rd Italian Conference on Computational Logic, Bolzano, Italy, September 20-22, 2018., Vol. 2214 of CEUR Workshop Proceedings, CEUR-WS.org, 2018, pp. 55–59.
URL `http://ceur-ws.org/Vol-2214/paper6.pdf`

[43] A. Lieto, G. L. Pozzato, A description logic framework for commonsense conceptual combination integrating typicality, probabilities and cognitive heuristics, arXiv preprint arXiv:1811.02366.