Reasoning About Plausible Scenarios in Description Logics of Typicality

Gian Luca Pozzato

Abstract. In some application domains it is sometimes useful to reason about plausible but “not obvious” scenarios, excluding the most trivial ones. In this work we introduce nonmonotonic procedures for preferential Description Logics in order to reason about plausible but – in some sense – “surprising” scenarios. We consider an extension, called $\mathcal{ALC} + \mathbf{T}_{\text{exp}}$, of the nonmonotonic logic of typicality $\mathcal{ALC} + \mathbf{T}_R$ by inclusions of the form $\mathbf{T}(C) \sqsubseteq_d D$, where $d$ is a degree of expectedness. We consider a notion of extension of an ABox, in order to assume typicality assertions about individuals satisfying cardinality restrictions on concepts, then we define a preference relation among such extended ABoxes based on the degrees of expectedness, then we restrict entailment to those extensions that are minimal with respect to this preference relation. We propose a decision procedure for reasoning in $\mathcal{ALC} + \mathbf{T}_{\text{exp}}$ and we exploit it to show that entailment is in $\text{EXPTIME}$ as for the underlying $\mathcal{ALC}$. Last, we introduce a further extension of the proposed approach in order to reason about all plausible extensions of the ABox, by restricting the attention to specific degrees of expectedness ranging from the most surprising scenarios to the most expected ones, and we show that entailment remains in $\text{EXPTIME}$ also in these cases.

Keywords: Description Logics, Nonmonotonic Reasoning, Typicality

1. Introduction

Nonmonotonic extensions of Description Logics (from now on, DLs for short) have been actively investigated since the early 90s [4,5,6,9,13,14,15,28] in order to tackle the problem of representing prototypical properties of classes and to reason about defeasible inheritance. A simple but powerful nonmonotonic extension of DLs is proposed in [16,17,18,19,21]: in this approach “typical” or “normal” properties can be directly specified by means of a “typicality” operator $\mathbf{T}$ enriching the underlying DL. The semantics of the $\mathbf{T}$ operator is characterized by the core properties of nonmonotonic reasoning axiomatized by either preferential logic [22] or rational logic [23]. We focus on the Description Logic $\mathcal{ALC} + \mathbf{T}_R$ introduced in [21]. In this logic one can express defeasible inclusions such as “typical depressed persons have sleep disorders”:

$$\mathbf{T}(\text{Depressed}) \sqsubseteq \exists \text{Symptom}. \text{SleepDisorder} \quad (*)$$

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, a patient affected by depression is not able to react to positive life events”, whereas “mood reactivity (ability to feel better temporarily in response to positive life events) is a typical symptom of atypical depression” as follows:

$$\mathbf{T}(\text{Depressed}) \sqsubseteq \neg \exists \text{Symptom}. \text{MoodReactivity}$$

$$\mathbf{T}(\text{AtypicalDepressed}) \sqsubseteq \exists \text{Symptom}. \text{MoodReactivity}$$

In the Description Logic $\mathcal{ALC} + \mathbf{T}_R$ standard models are extended by a function $f$ which selects the typ-
The function $f$ satisfies a set of postulates that are a restatement of Kraus, Lehmann and Magidor’s axioms of rational logic $\mathbf{R}$. This allows the typicality operator to inherit well-established properties of nonmonotonic reasoning: as an example, the property known as specificity, namely the choice of according preference to more specific information in case of conflicts among inherited properties, results “built-in” in the approach.

The logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}$ itself is too weak in several application domains. Indeed, although the operator $\mathbf{T}$ is nonmonotonic ($\mathbf{T}(C) \subseteq E$ does not imply $\mathbf{T}(C \cap D) \subseteq E$), the logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}$ is monotonic, in the sense that if the fact $F$ follows from a given knowledge base $\mathbf{KB}$, then $F$ also follows from any $\mathbf{KB}' \supseteq \mathbf{KB}$. As a consequence, unless a $\mathbf{KB}$ contains explicit assumptions about typicality of individuals, there is no way of inferring defeasible properties about them: in the above example (*), if $\mathbf{KB}$ contains the fact that Kate is a depressed woman, i.e. $\text{Depressed(kate)} \in \mathbf{KB}$, it is not possible to infer that she has sleep disorders ($\exists \text{Symptom.SleepDisorder(kate)}$). This would be possible only if the $\mathbf{KB}$ contained the stronger information that Kate is a typical depressed woman, i.e. $\mathbf{T} \text{(Depressed)}(\text{kate}) \in \mathbf{KB}$ to (or can be inferred from) $\mathbf{KB}$. In order to overcome this limit and perform useful inferences, in [21] the authors have introduced a nonmonotonic extension of the logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}$ based on a minimal model semantics, corresponding to a notion of rational closure as defined in [23] 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 $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$, supports typicality assumptions, so that if one knows that Kate is depressed, one can nonmonotonically assume that she is also a typical depressed if this is consistent, and therefore that she has sleep disorders. From a semantic point of view, the logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$ is based on a preference relation among $\mathbf{ALC} + \mathbf{T}_\mathbf{R}$ models and a notion of minimal entailment restricted to models that are minimal with respect to such preference relation.

The logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$ imposes to consider all typicality assumptions that are consistent with a given $\mathbf{KB}$. This seems to be too strong in several application domains, in particular when the need arises of bounding the cardinality of the extension of a given concept, that is to say the number of domain elements being members of such a concept, as introduced in [1]. As a further example, consider the following $\mathbf{KB}$, representing knowledge about toys and children preferences about them:

\[
\begin{align*}
\mathbf{T} & (\text{ActionFigure}) \sqsubseteq \text{BoysToy} \\
\mathbf{T} & (\text{TechToy}) \sqsubseteq \text{BoysToy} \\
\mathbf{T} & (\text{Vehicle}) \sqsubseteq \text{BoysToy} \\
\mathbf{T} & (\text{Dollhouse}) \sqsubseteq \neg \text{BoysToy}
\end{align*}
\]

representing that, normally, action figures, tech toys and vehicles are toys appreciated by boys, whereas dollhouses are not (last inclusion). Suppose the assertional part of the $\mathbf{KB}$ contains the facts:

\[
\begin{align*}
\text{ActionFigure} & (\text{techThor}) \\
\text{TechToy} & (\text{techThor}) \\
\text{Vehicle} & (\text{garagePlayset}) \\
\text{ActionFigure} & (\text{yoKaiJibanyan})
\end{align*}
\]

Facts (a) and (b) state that the action figure of the Avengers titan hero Thor is also a tech toy, since it really speaks. Fact (c) means that a garage playset is classified in the “vehicle” category, whereas fact (d) states that we can consider the action figure of Yo-kai Watch’s Jibanyan. In the nonmonotonic logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$ we are able to infer the following facts:

\[
\begin{align*}
\mathbf{T} & (\text{ActionFigure})(\text{techThor}) \\
\mathbf{T} & (\text{TechToy})(\text{techThor}) \\
\mathbf{T} & (\text{Vehicle})(\text{garagePlayset}) \\
\mathbf{T} & (\text{ActionFigure})(\text{yoKaiJibanyan})
\end{align*}
\]

and then that techThor, garagePlayset, and yoKaiJibanyan are all boys’ toys. This happens in $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$ because it is consistent to make the three assumptions above, that hold in all minimal models, however one should be interested in selecting only one toy (for instance, for a present), therefore considering three distinct scenarios that cannot be captured by $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$ as it is. One could think of extending the logic $\mathbf{ALC} + \mathbf{T}_\mathbf{R}^{\text{RCl}}$ by means of cardinality restrictions, in the example by imposing that there is only one member of the extension of the concept BoysToy, however the resulting knowledge base would be inconsistent. Furthermore, it is sometimes useful to restrict reasoning to surprising scenarios, excluding “trivial” / “obvious” ones. For instance, in the above example about toys, one could be interested in selecting an appreciated boys’ toy, but not the most “obvious” one, in order to minimize the risk of duplication of birthday gifts. The ontology engineer could need to represent different degrees of typical properties: action figures and vehicles are both typical boys’ toys, however there are more exceptions of action figures not intended for
In this work we introduce the logic $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$, which is based on the combination of two components. On the one hand, we allow one to express different degrees of expectedness of typicality inclusions, having the form $T(C) \subseteq_d D$ where $d$ is a positive integer such that an inclusion with degree $d$ is more “trivial” (or “obvious”) with respect to another one with degree $d' \leq d$: this allows one to describe several plausible scenarios by considering different combinations of typicality assumptions about individuals named in the ABox. Such degrees introduce a rank of expectedness among plausible scenarios, ranging from surprising to obvious ones. On the other hand, TBoxes are extended to allow restrictions about the cardinality of concepts, in order to “filter” such plausible scenarios. Finally, reasoning tasks are restricted to reasonable but not “trivial” in this choice, it could be useful to conclude that the action figure of Yo-kai Watch’s Libyann has to be preferred to a garage playset as well as to the “less unexpected” action figure of Thor. As another example, recently a great attention has been devoted to serendipitous search engines, that must be able to provide results that are “surprising, semantically cohesive, i.e. relevant to some information need of the user, or just interesting” [11]. In this sense, the scenario (among those satisfying cardinality restrictions) obtained by assuming the largest set of consistent typicality assumptions in $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$ corresponds to the most trivial one, whereas one could be interested in less expected ones, in which some typicality assumptions are discarded.

The plan of the paper is as follows. In Section 2 we recall Description Logics of typicality $\mathcal{ALC} + \mathcal{T}_{R}$ and $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$, then in Section 3 we introduce the logic $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$, allowing to express degrees of expectedness of typicality inclusions as well as cardinality restrictions, and to reason about surprising scenarios. In Section 4 we provide a decision procedure for checking query entailment in $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$, then we exploit it to show that the complexity of reasoning in $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$ is in Exptime. In Section 5 we extend the proposed approach in order to reason about scenarios “in the middle”, each one characterized by a specific degree of expectedness, ranging from the most surprising to the most expected situations, and we show that reasoning in $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$ is in Exptime also in this case. In Section 6 we mention related works, focusing on the differences between the logic $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$ and the closest approaches of DLs allowing the formalization of priorities among inclusions and probabilistic DLs under the DISPONTE semantics: in particular, for the former, we show that priorities among defaults are provided for free by the underlying logic of typicality, and not by means of degrees of expectedness; for the latter, we show that the logic $\mathcal{ALC} + \mathcal{T}_{R}^{\exp}$ allows one to perform useful reasoning that cannot be captured by reasoning in probabilistic DLs. We conclude in Section 7 with a brief discussion on how we aim at extending our approach in future research.

This work is an extended and improved version of [25], whereas preliminary ideas have been outlined by considering an extension of the lightweight DL-Lite-core for surprising scenarios in [24].

2. Preferential Description Logics

2.1. The monotonic logic $\mathcal{ALC} + \mathcal{T}_{R}$

The logic $\mathcal{ALC} + \mathcal{T}_{R}$ is obtained by adding to standard $\mathcal{ALC}$ the typicality operator $\mathcal{T}$ [16]. The intuitive idea is that $\mathcal{T}(C)$ selects the typical instances of a concept $C$. We can therefore distinguish between the properties that hold for all instances of concept $C \subseteq D$, and those that only hold for the normal or typical instances of $C \subseteq D$.

The semantics of the $\mathcal{T}$ 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 $\mathcal{R}$ [23]: in this respect an assertion of the form $\mathcal{T}(C) \subseteq D$ is equivalent to the conditional assertion $C \Rightarrow D$ in $\mathcal{R}$. The basic ideas are as follows: given a domain $\Delta^\mathcal{T}$ and an evaluation function $^\mathcal{T}$, one can define a function $f_{\mathcal{T}} : \text{Pow}(\Delta^\mathcal{T}) \mapsto \text{Pow}(\Delta^\mathcal{T})$ that selects the typical instances of any $S \subseteq \Delta^\mathcal{T}$; in case $S = C^\mathcal{T}$ for a concept $C$, the selection function selects the typical instances of $C$, namely:

$$\langle \mathcal{T}(C) \rangle^\mathcal{T} = f_{\mathcal{T}}(C^\mathcal{T}).$$

$f_{\mathcal{T}}$ has the following properties for all subsets $S$ of $\Delta^\mathcal{T}$, that are essentially a restatement of the properties characterizing rational logic $\mathcal{R}$.
(f_T - 1) f_T(S) \subseteq S
(f_T - 2) \text{if } S \neq \emptyset, \text{ then also } f_T(S) \neq \emptyset
(f_T - 3) \text{if } f_T(S) \subseteq R, \text{ then } f_T(S) = f_T(S \cap R)
(f_T - 4) f_T(\bigcup S_i) \subseteq \bigcup f_T(S_i)
(f_T - 5) \bigcap f_T(S_i) \subseteq f_T(\bigcap S_i)
(f_T - 6) \text{if } f_T(S) \cap R \neq \emptyset, \text{ then } f_T(S \cap R) \subseteq f_T(S)

The semantics of the \( T \) operator can be equivalently formulated in terms of rational models \[21\]: a model \( M \) is any structure \( (\Delta^2, \langle, \rangle) \) where \( \Delta^2 \) is the domain, \( < \) is an irreflexive, transitive, well-founded and modular (for all \( x, y, z \) in \( \Delta^2 \), if \( x < y \) then either \( x < z \) or \( z < y \) relation over \( \Delta^2 \)). 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 \( x \in \Delta^2 \) is a typical instance of some concept \( C \) if \( x \in \Delta^2 \) and there is no \( C \)-element in \( \Delta^2 \) more typical than \( x \). In detail, \( \mathcal{T} \) is the extension function that maps each concept \( C \) to \( C^2 \subseteq \Delta^2 \), and each role \( R \) to \( R^2 \subseteq \Delta^2 \times \Delta^2 \). For concepts of \( ALC, C^2 \) is defined as usual. For the \( T \) operator, we have

\[(T(C))^2 = Min_{<}(C^2),\]

where \( Min_{<}(C^2) = \{ x \in C^2 | \exists y \in C^2 \text{ s.t. } y < x \} \).

A model \( M \) can be equivalently defined by postulating the existence of a function \( k_M : \Delta^2 \rightarrow \mathbb{N} \), where \( k_M \) assigns a finite rank to each domain element: the rank function \( k_M \) and \( < \) can be defined from each other by letting \( x < y \) if and only if \( k_M(x) < k_M(y) \).

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

2.2. The nonmonotonic logic \( ALC + T_{R}^{\text{RatCl}} \)

Even if the typicality operator \( T \) itself is nonmonotonic (i.e. \( T(C) \subseteq E \) does not imply \( T(C \cap D) \subseteq E \)), what is inferred from a KB can still be inferred from any KB’ with KB \( \subseteq \text{KB}’ \), i.e. the logic \( ALC + T_{R} \) is monotonic. In order to perform useful nonmonotonic inferences, in \[21\] 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 \( ALC + T_{R}^{\text{RatCl}} \) and it corresponds to a notion of rational closure on top of \( ALC + T_{R} \). Such a notion is a natural extension of the rational closure construction provided in \[23\] for the propositional logic.

The nonmonotonic semantics of \( ALC + T_{R}^{\text{RatCl}} \) 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 \( M \) 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 \( KB \models_{ALC+T_{R}^{\text{RatCl}}} F \), if it holds in all minimal canonical models of KB. In \[21\] it is shown that query entailment in \( ALC + T_{R}^{\text{RatCl}} \) is in \text{ExpTime}. The construction of the rational closure and the correspondence between semantics and construction is recalled in Section 4.2.

3. Between \( ALC + T_{R} \) and \( ALC + T_{R}^{\text{RatCl}} \): the logic \( ALC + T_{R}^{\text{Exp}} \)

In this section we define an alternative semantics that allows us to express a degree of expectedness for the typicality inclusions and to limit the number of typicality assumptions in the ABox in order to obtain less predictable scenarios. The basic idea is similar to the one proposed in \[16\], where a completion of an \( ALC+T \) 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. Here we propose a similar, algorithmic construction in order to compute only some assumptions of typicality of domain elements/individual constants, in order to describe alternative, plausible, but not obvious scenarios. Constraints about the cardinality of the extensions of concepts are also introduced in order to filter scenarios, allowing to define eligible extensions of the ABox satisfying such constraints, and entailment is restricted to minimal scenarios, called perfect extensions, with

\(^1\)In Theorem 10 in \[21\] the authors have shown that for any consistent KB there exists a finite minimal canonical model of KB.
respect to an order relation among extensions: intuitively, an extension is preferred to another one if it represents a more surprising scenario.

The logic \( ALC + T^{exp}_R \) allows one to express cardinality restrictions in the TBox. More expressive DLs allow one to specify (un)qualified number restrictions, in order to specify the number of possible elements filling a given role \( R \). As an example, number restrictions allow one to express that a student attends to 3 courses. Number restrictions are therefore “localized to the fillers of one particular role” [1], for instance we can have \( Student \sqsubseteq 3 \text{Attends} \). Course as a restriction on the number of role fillers of the role \textit{Attends}. However one could need to express global restrictions on the number of domain elements belonging to a given concept, for instance to express that in the whole domain there are exactly 3 courses. In DLs not allowing cardinality restrictions one can only express that every student must attend to three courses, but that not all must attend to the same ones. In the logic \( ALC + T^{exp}_R \), cardinality restrictions on concepts are added to the TBox: they are assertions of the form either \( (\geq n \ C) \) or \( (\leq n \ C) \) or \( (= n \ C) \), where \( n \) is a positive integer and \( C \) is a concept, as proposed in [1].

This is formally defined in the next definition, where, given a set \( S \), \( \not\exists S \) is the cardinality of \( S \).

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

\[
A := A \mid T \mid \bot \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \forall R.C \mid \exists R.C
\]

An \( ALC + T^{exp}_R \) knowledge base is a pair \( (T, A) \). \( T \) contains axioms of the form:

- \( C \sqsubseteq C \);
- \( T(C) \sqsubseteq d \ C \), where \( d \in \mathbb{N}^+ \) is called the degree of expectedness;
- \( (\odot n \ C) \), where \( \odot \in \{\leq, \leq, \geq\} \) and \( n \in \mathbb{N}^+ \).

\( A \) contains assertions of the form \( C(a) \) and \( R(a, b) \), where \( a, b \in O \).

Given an inclusion \( T(C) \sqsubseteq d \ D \), the higher the degree of expectedness the more the inclusion is, in some sense, “obvious”/not surprising. Given another inclusion \( T(C') \sqsubseteq d' \ D' \), with \( d' < d \), we assume that this inclusion is less “obvious”, more surprising with respect to the other one. As an example, let KB contain \( T(Student) \sqsubseteq_4 SocialNetworkUser \) and \( T(Student) \sqsubseteq_2 PartyParticipant \), representing that typical students make use of social networks, and that normally they go to parties; however, the second inclusion is less obvious with respect to the first one. In other words, one can think of representing the fact that both are properties of a prototypical student, however there are more exceptions of students not taking part to parties with respect to the number of exceptions of students not being part of the social media ecosphere.

It is known that axioms expressing cardinality restrictions are even more expressive than inclusions \( C \sqsubseteq D \) or \( C \sqsubseteq D \), the last one shortening for the pair of inclusions \( C \sqsubseteq D \) and \( D \sqsubseteq C \) (thus expressing that the concepts \( C \) and \( D \) have the same extensions, i.e. \( C^D = D^C \)). Indeed, \( C \sqsubseteq D \) means that the set of domain elements that are \( C \)'s but not \( D \)’s is empty, and viceversa (the set of domain elements that are \( D \)’s but not \( C \)’s is empty). This can be expressed by the following cardinality restriction: \( (\leq 0 ((C \cap \neg D) \sqcup (\neg C \cap D))) \). The same for an inclusion of the form \( C \sqsubseteq D \), whose meaning is that every \( C \) element is also a \( D \) element, that can be expressed by \( (\leq 0 (C \cap \neg D)) \) (the intersection of \( C \) and the complement of \( D \) is empty). Therefore, we could restrict our language to TBoxes only containing cardinality restrictions, however we have decided to consider the extended language of Definition 1 for the sake of readability.

Before introducing technical details and formal definitions (sections 3.1 and 3.2), we provide an example in order to give an intuitive idea of what we mean for reasoning about not obvious scenarios in the logic \( ALC + T^{exp}_R \).

**Example 1 (Mysterious medical diagnosis)** Let \( KB = (T, \emptyset) \) where \( T \) is as follows:

\[
\begin{align*}
&\text{Depressed} \sqsubseteq \text{Condition} \\
&\text{ProstateCancerPatient} \sqsubseteq \text{Condition} \\
&\text{Bipolar} \sqsubseteq \text{Condition} \\
&\text{AtypicalDepressed} \sqsubseteq \text{Depressed} \\
&T(\text{Depressed}) \sqsubseteq_2 \neg \exists \text{Symptom.MoodReactivity} \\
&T(\text{Bipolar}) \sqsubseteq_3 \exists \text{Symptom.MoodReactivity} \\
&T(\text{AtypicalDepressed}) \sqsubseteq_4 \exists \text{Symptom.MoodReactivity} \\
&T(\text{ProstateCancerPatient}) \sqsubseteq_2 \exists \text{Symptom.MoodReactivity} \\
&(\geq 1 \text{ Condition}) \\
&(\leq 2 \text{ Condition})
\end{align*}
\]

the last ones stating that we want to focus on at least one/at most two conditions determining patient’s symptoms. We have that

\[
(*) \ T(\text{Depressed} \cap \text{AcromegalicGiant}) \sqsubseteq \neg \exists \text{Symptom.MoodReactivity}
\]
follows\(^2\) from KB, and this is a wanted inference, since being affected by acromegaly, a rare syndrome that results when the anterior pituitary gland produces excess growth hormone, is irrelevant with respect to mood reactivity as far as we know. This is a non-monotonic inference that does no longer follow if it is discovered that typical depressed people also affected by acromegaly are subject to mood reactivity: given \(T' = T \cup \{T(\text{Depressed} \land \text{AcromegalicGiant}) \sqsubseteq \exists \text{Symptom}. \text{MoodReactivity}\}\), we have that the inclusion (+) does no longer follow from the KB with \(T'\) in the logic \(\mathcal{ALC} + T^\text{exp}_R\).

As for rational closure, the set of facts/inclusions that are entailed from a \(\mathcal{ALC} + T^\text{exp}_R\) KB is closed under the property known as rational monotonicity: for instance, from KB and the fact that

\[
T(\text{Depressed}) \sqsubseteq \neg \text{Elder}
\]

is not entailed from KB in \(\mathcal{ALC} + T^\text{exp}_R\), it follows that the inclusion

\[
T(\text{Depressed} \land \text{Elder}) \sqsubseteq \neg \exists \text{Symptom}. \text{MoodReactivity}
\]

is entailed in \(\mathcal{ALC} + T^\text{exp}_R\).

Concerning ABox reasoning, we can think of exploiting the logic \(\mathcal{ALC} + T^\text{exp}_R\) in order to find a mysterious medical diagnosis to explain patients’ symptoms and signs, a set of formulas of the form \(C_r(a)\) that we call \(\mathcal{P}\). For instance, let \(\mathcal{P}\) describe Greg’s symptom, in particular that he has mood reactivity:

\[
\mathcal{P} = \{\exists \text{Symptom}. \text{MoodReactivity}(\text{greg})\}.
\]

We have that \(\mathcal{P}\) is not entailed by KB, but KB \(\cup \mathcal{P}\) is consistent. We are then interested in finding a diagnosis for Greg’s symptoms, that is to say a set of assertions \(D\) such that \(\mathcal{P}\) follows from KB \(\cup D\). The most trivial scenario suggests that Greg is affected by the bipolar disorder, i.e., \(D = \{\text{Bipolar}(\text{greg})\}\), however this scenario is discarded in the logic \(\mathcal{ALC} + T^\text{exp}_R\): in this context, the condition that is taken into consideration is prostatic cancer, i.e.,

\[
D = \{\text{ProstateCancerPatient}(\text{greg})\},
\]

and such a non trivial diagnosis could be confirmed by an evaluation of other typical symptoms of such a disease (e.g. nocturia).

One could object that no one would be interested in a medical diagnosis support system that discards the most likely explanation for a medical problem, however, as mentioned in the Introduction, the idea underlying the proposed approach is not to ignore the most expected explanation, rather to “go beyond” it in order to find (unexpected) alternative ones in case of a failure with the standard diagnosis. In other words: if the most likely explanation does not provide a solution, the logic \(\mathcal{ALC} + T^\text{exp}_R\) tries to provide less obvious alternatives that could be taken into account for further investigations.

### 3.1. Extensions of ABox

Given a KB, we define the finite set \(C\) of concepts occurring in the scope of the typicality operator, i.e., \(C = \{C \mid T(C) \sqsubseteq_d D \in \text{KB}\}\). These are the concepts whose atypical instances we want to minimize.

Given an individual \(a\) explicitly named in the ABox, we define the set of “plausible” typicality assumptions \(T(C)(a)\) that can be minimally entailed from KB without cardinality restrictions in the logic \(\mathcal{ALC} + T^\text{exp}_R\), with \(C \in C\). We then consider an ordered set \(\mathcal{C}_A\) of pairs \((a, C)\) of all possible assumptions \(T(C)(a)\), for all concepts \(C \in C\) and all individual constants \(a\) occurring in the ABox.

**Definition 2 (Assumptions in \(\mathcal{ALC} + T^\text{exp}_R\))** Given an \(\mathcal{ALC} + T^\text{exp}_R\) KB = \((T \cup T^\text{card}_C, A)\), where \(T^\text{card}_C\) is a set of cardinality restrictions and \(T\) does not contain cardinality restrictions, let \(T'\) be the set of inclusions of \(T\) without degrees of expectedness, namely \(T' = \{T(C) \sqsubseteq D \mid T(C) \sqsubseteq_d D \in T\} \cup \{C \sqsubseteq D \in T\}\). Given a finite set of concepts \(A\), we define, for each individual name \(a\) occurring in \(A\):

\[
\mathcal{C}_a = \{C \in C \mid (T', A) |=_{\mathcal{ALC} + T^\text{exp}_R} T(C)(a)\}
\]

We also define \(\mathcal{C}_A\):

\[
\mathcal{C}_A = \{(a, C) \mid C \in \mathcal{C}_a \text{ and } a \text{ occurs in } A\}
\]

and we impose an order on the elements of \(\mathcal{C}_A\): \(\mathcal{C}_A = \left\{(a_1, C_1), (a_2, C_2), \ldots, (a_n, C_n)\right\}\). Furthermore, we define the ordered multiset \(d_A = [d_1, d_2, \ldots, d_n]\), respecting the order imposed on \(\mathcal{C}_A\), where \(d_i = \text{avg}(\{d \in \mathbb{N}^+ \mid T(C_i) \sqsubseteq_d D \in T\})\).
Intuitively, the ordered multisets $d_A$ is a tuple of the form $[d_1, d_2, \ldots, d_n]$, where $d_i$ is the degree of expect- edness of the assumption $T(C)(a)$, such that $(a, C) \in \mathbb{C}_A$ at position $i$. $d_i$ corresponds to the average of all the degrees $d$ of typicality inclusions $T(C) \sqsubseteq_d D$ in the TBox.

In order to define alternative scenarios, where not all plausible assumptions are taken into account, we consider different extensions of the ABox and we introduce an order among them, allowing to range from unpredictable to trivial ones. Starting from $d_A = [d_1, d_2, \ldots, d_n]$, the first step is to build all alternative tuples where 0 is used in place of some $d_i$ to represent that the corresponding typicality assertion $T(C)(a)$ is no longer assumed (Definition 3). Furthermore, we define the extension of the ABox corresponding to a string so obtained (Definition 4). To give an intuitive idea, before introducing the formal definitions, let us consider the following example:

**Example 2** Given a $KB = (T, A)$, the only typicality inclusions in $T$ be $T(C) \sqsubseteq_1 D$ and $T(E) \sqsubseteq_2 F$. Let $a$ and $b$ be the only individual constants occurring in $A$. Suppose also that $T(C)(a)$, $T(C)(b)$, and $T(E)(b)$ are entailed in $ALC + T_{R_{UC}}$, more precisely:

- let $KB'$ be the knowledge base obtained by replacing $T(C) \sqsubseteq_1 D$ and $T(E) \sqsubseteq_2 F$ in $T$ by $T(C) \sqsubseteq D$ and $T(E) \sqsubseteq F$, respectively;
- suppose that:
  - $KB' \models_{ALC+T_{R_{UC}}} T(C)(a)$
  - $KB' \models_{ALC+T_{R_{UC}}} T(C)(b)$
  - $KB' \models_{ALC+T_{R_{UC}}} T(E)(b)$

We have that

$C_A = \{(a, C), (b, C), (b, E)\}$ and $d_A = [1, 1, 2]$

The other possible tuples (strings) are:

- [0, 0, 2], corresponding to extending $A$ with the assumptions $T(C)(a)$ and $T(E)(b)$;
- [0, 1, 0], corresponding to extending $A$ with the only assumption $T(C)(b)$;
- [1, 0, 0], corresponding to extending $A$ with the only assumption $T(C)(a)$;
- [0, 1, 2], corresponding to extending $A$ with the assumptions $T(C)(b)$ and $T(E)(b)$.

Other aggregation functions could be used to define $d_i$ (e.g. maximum/minimum degree). We aim at studying the impact of this choice on the reasoning machinery in future research.

Let us now introduce formal definitions for the above mentioned notions of string of assumptions and of extension of an ABox corresponding to a string.

**Definition 3 (Strings of possible assumptions)** Given a $KB = (T, A)$, let the set $C_A$ and $d_A = [d_1, d_2, \ldots, d_n]$ be as in Definition 2. We define the set $S$ of all the strings of possible assumptions with respect to $KB$ as

$S = \{[s_1, s_2, \ldots, s_n] | \forall i = 1, 2, \ldots, n \text{ either } s_i = d_i \text{ or } s_i = 0\}$

**Definition 4 (Extension of the ABox)** Let $KB = (T, A)$ and let $C_A = [(a_1, C_1), (a_2, C_2), \ldots, (a_n, C_n)]$ as in Definition 2. Given a string of possible assumptions $[s_1, s_2, \ldots, s_n] \in S$ of Definition 3, we define the extension $\tilde{A}$ of $A$ with respect to $C_A$ and $S$ as:

$\tilde{A} = \{T(C_i)(a_i) | (a_i, C_i) \in C_A \text{ and } s_i \neq 0\}$

It can be observed that, in $ALC + T_{R_{UC}}$, the set of typicality assumptions that can be inferred from a KB corresponds to the extension of $A$ corresponding to the string $d_A$ (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 $ALC + T_{R}$, no typicality assumptions can be derived from a KB, and this corresponds to extending $A$ by the assertions corresponding to the string $[0, 0, \ldots, 0]$, i.e. by the empty set.

### 3.2. Cardinality restrictions and perfect extensions

Let us now introduce models of the Description Logic $ALC + T_{R_{UC}}$ taking cardinality restrictions into account, as well as the notion of eligible extension of the ABox as a set of typicality assumptions satisfying cardinality restrictions.

**Definition 5** Given a model $M = (\Delta^T, \ll, T)$, it satisfies:

- (TBox)
  - an inclusion $C \sqsubseteq D$ if $C^T \sqsubseteq D^T$;
  - a typicality inclusion $T(C) \sqsubseteq_d D$ if $\text{Min}_<(C^T) \sqsubseteq D^T$;
and the tuples:

- a cardinality restriction of the form (\(\odot n C\)) if \(\exists^C \odot n,\) where \(\odot \in \{\leq, \geq, =\}\) and \(n \in \mathbb{N}^+\).

\[(\text{ABox})\]
- an assertion of the form \(C(a)\) if \(a^\exists \in C^\exists;\)
- an assertion of the form \(R(a, b)\) if \((a^R, b^R) \in R^R.\)

Given a \(\text{KB}=(T, A)\), we say that a model \(M\) satisfies \(\text{KB}\) if it satisfies all the inclusions in \(T\) and all the assertions in \(A\).

**Definition 6 (Eligible extension \(\hat{A}\))** Given an \(\text{ALC} + T^\text{R}_{\text{op}} \text{KB}=(T, A)\) and an extension \(\hat{A}\) of \(A\) as in Definition 4, we say that \(\hat{A}\) is eligible if there exists an \(\text{ALC} + T^\text{R}_{\text{op}}\) model \(M\) as in Definition 5 that satisfies \(\text{KB}=(\hat{T}, \hat{A} \cup \hat{A}).\)

**Definition 7 (Order between eligible extensions)** Given \(\text{KB}=(T, A)\) and the set \(S\) of Definition 3, let \(s = [s_1, s_2, \ldots, s_n]\) and \(r = [r_1, r_2, \ldots, r_n]\), with \(s, r \in S\). Let \(\hat{A}_s\) and \(\hat{A}_r\) be two eligible extensions of \(A\) corresponding to \(s\) and \(r\) (Definition 4). We say that \(s < r\) if there exists a bijection \(\delta\) between \(s\) and \(r\) such that, for each \((s_i, r_j) \in \delta\), it holds that \(s_i \leq r_j\), and there is at least one \((s_i, r_j) \in \delta\) such that \(s_i < r_j\). We say that \(\hat{A}_s\) is more surprising (or less trivial) than \(\hat{A}_r\) if \(s < r\).

Intuitively, a string \(s\) whose elements are “lower” than the ones of another string \(r\) corresponds to a less trivial ABox. For instance, let us consider again the knowledge base of Example 2: given the strings \(s = [1, 1, 0]\) and \(r = [1, 0, 2]\), we have that \(s < r\), because there exists a bijection \(((1, 1), (0, 0), (1, 2))\). The assumptions \(T(C)(a)\) and \(T(C)(b)\) corresponding to \(s\) are then considered less trivial than \(T(C)(a)\) and \(T(E)(b)\) corresponding to \(r\). It is worth noticing that the order of Definition 7 is partial: as an example, the strings \([1, 1, 0]\) and \([0, 0, 2]\) are not comparable, in the sense that \([1, 1, 0] \not< [0, 0, 2]\) and \([0, 0, 2] \not< [1, 1, 0]\). In order to choose between two incomparable situations, we introduce the following notion of weak order: intuitively, given two incomparable extensions \(\hat{A}_s\) and \(\hat{A}_r\), we assume that \(\hat{A}_s\) is weakly less trivial than \(\hat{A}_r\) if \(\hat{A}_s\) is strictly included in another eligible extension \(\hat{A}_u\), more trivial than \(\hat{A}_r\), i.e. \(\hat{A}_s \subset \hat{A}_u\), and \(s < u\). As an example, consider the KB whose TBox is \(\{T(Bird) \sqsubseteq F_{fly}\}\) and whose ABox is:

\[\{T(Bird)(\text{tweety}), T(Bird)(\text{jim}), T(Bird)(\text{mike})\}\]

and the tuples:

\[s_1 = [2, 2, 0] \quad s_2 = [2, 0, 0] \quad s_3 = [0, 0, 2]\]

corresponding to the following extensions:

\[\hat{A}_1 = \{T(Bird)(\text{tweety}), T(Bird)(\text{jim})\}\]
\[\hat{A}_2 = \{T(Bird)(\text{tweety})\}\]
\[\hat{A}_3 = \{T(Bird)(\text{mike})\}\]

we have that the extensions \(\hat{A}_2\) and \(\hat{A}_3\) are not comparable, however \(\hat{A}_3\) is a proper subset of \(\hat{A}_1\), which is more “obvious” than \(\hat{A}_3\) (indeed, \(s_3 < s_1\)). Therefore, we assume that the scenario represented by \(\hat{A}_3\) is weakly less trivial than the one represented by \(\hat{A}_2\).

This is formally stated in next definition:

**Definition 8 (Weak preference)** Given a \(\text{KB}=(T, A)\), let \(\hat{A}_s\) and \(\hat{A}_r\) be two eligible extensions of \(A\) such that neither \(\hat{A}_s\) is more surprising than \(\hat{A}_r\) nor \(\hat{A}_r\) is more surprising than \(\hat{A}_s\). We say that \(\hat{A}_s\) (weakly) more surprising (or (weakly) less trivial) than \(\hat{A}_r\) if there exists an eligible extension \(\hat{A}_u\) of \(A\) such that (i) \(\hat{A}_s\) is more surprising than \(\hat{A}_u\) (Definition 7) and (ii) \(\hat{A}_r \subset \hat{A}_u\).

**Definition 9 (Minimal (perfect) extensions)** Given a \(\text{KB}=(T, A)\) and the set \(S\) of strings of possible assumptions (Definition 3), we say that an eligible extension \(\hat{A}_s\) is minimal if there is no other eligible extension \(\hat{A}_r\) which is (weakly) more surprising (or (weakly) less trivial) than \(\hat{A}_r\).

Given the above definitions, we can define a notion of entailment in \(\text{ALC} + T^\text{R}_{\text{op}}\). Intuitively, given a query \(F\), we distinguish two cases:

- if \(F\) is a TBox inclusion \(C \sqsubseteq D\) (even if it is a typicality inclusion, i.e. \(C\) has the for \(T(C)\)), we rely on reasoning in the nonmonotonic logic \(\text{ALC} + T^\text{R}_{\text{op}}\) and we say that \(F\) is entailed by \(\text{KB}\) if it is minimally entailed in \(\text{ALC} + T^\text{R}_{\text{op}}\) from \(\text{KB}\) or, equivalently, \(F\) is in the rational closure of \(\text{KB}\);
- if \(F\) is an ABox fact \(C(a)\), we check whether \(F\) follows in the monotonic logic \(\text{ALC} + T^\text{R}\) from a given \(\text{KB}\), whose ABox is augmented with extensions that are minimal (perfect) as in Definition 9. We can reason either in a skeptical way, by asking that \(F\) is entailed if it follows in all \(\text{KBs}\), obtained by considering each minimal extension of the ABox, or in a credulous way, by assuming that
Definition 10 (Skeptical entailment in $\mathcal{ALC} + \mathcal{T}_{\text{exp}}$)

Given a KB=$\langle T, A \rangle$ and given $C$ a set of concepts, let $E$ the set of all extensions of $A$ that are minimal as in Definition 9. Given a query $F$, we say that $F$ is skeptically entailed from KB in $\mathcal{ALC} + \mathcal{T}_{\text{exp}}$, written KB $\models_{sk}^{\mathcal{ALC}+\mathcal{T}_{\text{exp}}} F$:
- if $F$ is a TBox inclusion $C \subseteq D$, if it holds that KB $\models_{\mathcal{ALC}+\mathcal{T}_{\text{exp}}} F$;
- if $F$ is an ABox fact $C(a)$, where $a \in \mathcal{O}$, if (T, A $\cup \hat{A}$) $\models_{\mathcal{ALC}+\mathcal{T}_{\text{exp}}} F$ for all $\hat{A} \in E$.

Definition 11 (Credulous entailment in $\mathcal{ALC} + \mathcal{T}_{\text{exp}}$)

Given a KB=$\langle T, A \rangle$ and given $C$ a set of concepts, let $E$ the set of all extensions of $A$ that are minimal as in Definition 9. Given a query $F$, we say that $F$ is credulously entailed from KB in $\mathcal{ALC} + \mathcal{T}_{\text{exp}}$, written KB $\models_{cr}^{\mathcal{ALC}+\mathcal{T}_{\text{exp}}} F$:
- if $F$ is a TBox inclusion $C \subseteq D$, if it holds that KB $\models_{\mathcal{ALC}+\mathcal{T}_{\text{exp}}} F$;
- if $F$ is an ABox fact $C(a)$, where $a \in \mathcal{O}$, if there exists $\hat{A} \in E$ such that (T, A $\cup \hat{A}$) $\models_{\mathcal{ALC}+\mathcal{T}_{\text{exp}}} F$.

At a first glance, one could have the impression that the notions of rank in the semantics of $\mathcal{ALC} + \mathcal{T}_{\text{exp}}$, where elements with lowest rank are the most typical ones, and the semantics of expectedness of Definitions 7 and 9, where lower ranks correspond to more surprising scenarios, are in conflict. However, this is not the case: ranks in the semantics are introduced in order to define extensions of typicality concepts, and this is also considered in the expectation semantics to select plausible typicality assumptions. The rank among extensions is rather used in order to choose less trivial scenarios, to restrict the number of typicality assumptions to satisfy cardinality restrictions: the unexpectedness is the additional ingredient to select surprising scenarios by fixing cardinality restrictions, where all candidates try to maximize the typicality of individuals.

Example 3 (Better Royal Rumble Match) The term sports entertainment is usually associated to professional wrestling, a combat sport combining athletics with theatrical performance. As a difference with typical athletics and games, which are conducted for competition, the main objective of sports entertainment is to entertain an audience. Wrestling matches are driven by storylines provided by a creative team, and their outcomes are predetermined: duration, sequence of athletic moves and, obviously, winners of the contests. One of the most attractive events in professional wrestling is the WWE Royal Rumble match: thirty athletes are involved in this competition, and the winner receives a title shot in the main annual event of the company. The objective of each participant is to eliminate all the other competitors by tossing them over the top rope of the ring; an athlete is eliminated if both his feet touch the floor outside the ring. The match starts with the two participants who have drawn entry numbers one and two, with the remaining competitors entering the ring at regular timed intervals, usually 90 seconds, according to their entrance number assigned by means of a lottery. On the contrary, the assignment of entrance numbers to the participants, the sequence of eliminations (who eliminates who), the last man being eliminated, as well as the winner himself are determined by the choices of a creative team and scheduled in all the details.

In the last years, the Royal Rumble match has been marked by an extremely negative audience reaction: the trivial sequence of eliminations, as well as the fact that the winners have been predicted before the match by professional wrestling web sites, lead the people in the arena to “boo” every single action of the show. In this example we exploit the logic $\mathcal{ALC} + \mathcal{T}_{\text{exp}}$ in order to tackle the problem of defining a better script for the Royal Rumble match, focusing on the selection of the winner. Let the TBox be as follows:

\[
\begin{align*}
\mathcal{T}(\text{FaceWrestler}) & \sqsubseteq_1 \text{RoyalRumbleWinner} \quad (T1) \\
\mathcal{T}(\text{Returning}) & \sqsubseteq_3 \text{RoyalRumbleWinner} \quad (T2) \\
\mathcal{T}(\text{Predicted}) & \sqsubseteq_2 \text{RoyalRumbleWinner} \quad (T3) \\
& \quad = 1 \text{RoyalRumbleWinner} \quad (T4)
\end{align*}
\]

the last one stating that there could be only one winner. Concerning typicality inclusions, we represent the facts that normally, a face wrestler wins the Royal Rumble match (T1), however this admits more exceptions with respect to the fact that, typically, an athlete whose victory has been predicted by wrestling web sites normally wins the Royal Rumble match (T3), which is in turn more surprising than the most “obvious” inclusion, namely that an athlete returning from
an injury wins the Royal Rumble match (T2). Let the ABox be:

\[
\begin{align*}
&\text{FaceWrestler}(\text{dean}) \\
&\text{Returning}(\text{seth}) \\
&\text{FaceWrestler}(\text{roman}) \\
&\text{Predicted}(\text{roman})
\end{align*}
\]

In the logic \(\mathcal{ALC} + T^\text{exp}_R\), we can infer that Dean is the winner of the Royal Rumble match:

\[
\begin{align*}
KB &\models_{sk}^{A\mathcal{L}C+T^\text{exp}_R} \text{RoyalRumbleWinner}(\text{dean}) \\
KB &\models_{CF}^{A\mathcal{L}C+T^\text{exp}_R} \text{RoyalRumbleWinner}(\text{dean})
\end{align*}
\]

since there is only one perfect extension. Let \(C = \{\text{FaceWrestler}, \text{Predicted}, \text{Returning}\}\). By Definition 2 above, we have that: \(C_{\text{dean}} = \{\text{FaceWrestler}\}\), \(C_{\text{roman}} = \{\text{FaceWrestler}, \text{Predicted}\}\), \(C_{\text{seth}} = \{\text{Returning}\}\). Furthermore, consider \(C_A = C_{\text{dean}} \cup C_{\text{roman}} \cup C_{\text{seth}}\) in this order. Concerning the degrees of expectedness, we have \(d_A = [1, 1, 2, 3]\). As mentioned, in \(\mathcal{ALC} + T^\text{exp}_R\) the minimal model semantics forces all the consistent typicality assumptions, namely we are considering an ABox extended with the following facts:

\[
\begin{align*}
T(\text{FaceWrestler})(\text{dean}) \\
T(\text{FaceWrestler})(\text{roman}) \\
T(\text{Predicted})(\text{roman}) \\
T(\text{Returning})(\text{seth})
\end{align*}
\]

Corresponding (in the sense of Definition 4) to the multiset \([1, 1, 2, 3]\). However, in \(\mathcal{ALC} + T_R\) we obtain that Dean, Roman and Seth are all winners, against the fact that we want to have only one winner: the extension corresponding to \([1, 1, 2, 3]\) is indeed not eligible in the sense of Definition 6. In order to find only one winner and to obtain a non-trivial outcome of the match, we consider the set \(\mathcal{S}\) of all possible strings of typicality assumptions (Definition 3). The only eligible extensions of the ABox are:

\[
\begin{align*}
\tilde{A}_1 &= \{T(\text{Returning})(\text{seth})\}, \text{by} \ [0, 0, 0, 3] \\
\tilde{A}_2 &= \{T(\text{Predicted})(\text{roman})\}, \text{by} \ [0, 0, 2, 0] \\
\tilde{A}_3 &= \{T(\text{FaceWrestler})(\text{dean})\}, \text{by} \ [1, 0, 0, 0] \\
\tilde{A}_4 &= \{T(\text{FaceWrestler})(\text{roman})\}, \text{by} \ [1, 0, 1, 0] \\
\tilde{A}_5 &= \{T(\text{FaceWrestler})(\text{roman})\}, \text{by} \ [0, 1, 2, 0] \\
\tilde{A}_6 &= \{T(\text{Predicted})(\text{roman})\}, \text{corresponding to} \ [0, 1, 2, 0]
\end{align*}
\]

We have that \(\tilde{A}_3\) and \(\tilde{A}_4\) are less trivial than \(\tilde{A}_5\), because \([1, 0, 0, 0] < [0, 1, 2, 0]\) and \([0, 1, 0, 0] < [0, 1, 2, 0]\). Furthermore, \(\tilde{A}_3\) and \(\tilde{A}_4\) are less trivial than \(\tilde{A}_1\) (again, \([1, 0, 0, 0] < [0, 0, 0, 3]\) and \([0, 1, 0, 0] < [0, 0, 0, 3]\). Moreover, \(\tilde{A}_3\) and \(\tilde{A}_4\) are less trivial than \(\tilde{A}_2\) (again, \([1, 0, 0, 0] < [0, 0, 0, 2]\) and \([0, 1, 0, 0] < [0, 0, 0, 2]\). The strings \([1, 0, 0, 0]\) and \([0, 1, 0, 0]\) are not comparable, however \(\tilde{A}_3\) is weakly less trivial than \(\tilde{A}_4\), since \(\tilde{A}_3 \subset \tilde{A}_4\) and \([1, 0, 0, 0] < [0, 1, 2, 0]\). This allows one to conclude that \(\tilde{A}_3\) is minimal (the perfect extension) and to suggest that Dean has to be chosen as the winner of the Royal Rumble match.

4. A Decision Procedure for \(\mathcal{ALC} + T^\text{exp}_R\)

In this section we describe a decision procedure for reasoning in the logic \(\mathcal{ALC} + T^\text{exp}_R\). We consider skeptical and credulous entailment. In both cases, we exploit the decision procedure to show that the problem of entailment in the logic \(\mathcal{ALC} + T^\text{exp}_R\) is in \(\text{EXPTIME}\). This allows us to conclude that reasoning about typicality and defeasible inheritance in surprising scenarios 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 \(\text{EXPTIME}\)-complete [2].

The procedure performs the following steps:

1. compute the set \(C_a\) of all typicality assumptions that are minimally entailed from the KB in the nonmonotonic logic \(\mathcal{ALC} + T^\text{R}_{\text{Rac}}\);
2. compute all possible extensions of the ABox and select perfect extensions;
3. check whether the query \(F\) is entailed from at least one extension/all the extensions of KB in the monotonic logic \(\mathcal{ALC} + T_R\) plus cardinality restrictions.

Step 3 is based on reasoning in the monotonic logic \(\mathcal{ALC} + T_R\): to this aim, the procedure relies on a polynomial encoding of \(\mathcal{ALC} + T_R\) into \(\mathcal{ALC}\) introduced in [20] and then on reasoning with cardinality restrictions. Step 1 is based on reasoning in the nonmonotonic logic \(\mathcal{ALC} + T^\text{R}_{\text{Rac}}\): in this case, the procedure computes the rational closure of an \(\mathcal{ALC} + T_R\) knowledge base by means of the algorithm introduced in [21], which is sound and complete with respect to the minimal model semantics recalled in Section 2.2. Also the algorithm computing the rational closure relies on reasoning in the monotonic logic \(\mathcal{ALC} + T_R\), then on the above mentioned polynomial encoding in \(\mathcal{ALC}\). We assume unary encoding of numbers in cardinality restrictions in order to exploit the results in [29],
namely that reasoning in \(\mathcal{AXC}^0\), extending \(\mathcal{AXC}\) with qualified number restrictions, is \(\text{ExpTime}\)-complete also with cardinality restrictions. Before introducing the overall procedure for reasoning in \(\mathcal{AXC}^0+\mathcal{T}_{\text{R}}^{\text{exp}}\) and analyze its complexity, we briefly recall procedures for reasoning in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) and \(\mathcal{AXC}+\mathcal{T}_{\text{R}}^{\text{expCl}}\).

4.1. Reasoning in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\)

In order to reason in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\), in [20] the authors provide the following polynomial encoding in standard \(\mathcal{AXC}\) of KB\(^4\). The idea on which the encoding is based exploits the definition of the typicality operator \(\mathcal{T}\) in terms of a Gödel-Löb modality \(\square\) as follows: \(\mathcal{T}(C)\) is defined as \(C \cap \square \neg C\) where the accessibility relation of the modality \(\square\) is the preference relation \(\prec\) in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) models.

Let KB\(=\langle T, A \rangle\) be a knowledge base where \(A\) does not contain positive typicality assertions on individuals of the form \(\mathcal{T}(C)(a)\). We define the encoding KB\(='#\langle T', A' \rangle\) of KB in \(\mathcal{AXC}\) as follows. First of all, we let \(A' = \emptyset\). Then, for each \(A \subseteq B \in T\), not containing \(T\), we introduce \(A \subseteq B\) in \(T'\). For each \(T(A)\) occurring in \(T\), we introduce a new atomic concept \(\boxtimes_{\rightarrow}A\) and, for each inclusion \(T(A) \subseteq d\ B \in T\), we add to \(T'\) the inclusion

\[
A \cap \boxtimes_{\rightarrow}A \subseteq B.
\]

In order to capture the properties of the \(\square\) modality, a new role \(R\) is introduced to represent the relation \(\prec\) in preferential models, and the following inclusions are introduced in \(T'\):

\[
\begin{align*}
\boxtimes_{\rightarrow}A & \subseteq \forall R. (\neg A \cap \boxtimes_{\rightarrow}A) \\
\neg \boxtimes_{\rightarrow}A & \subseteq \exists R. (A \cap \boxtimes_{\rightarrow}A)
\end{align*}
\]

The first inclusion accounts for the transitivity of \(\prec\). 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 \(C_t \subseteq C_r\) is not a typicality inclusion, then \(C_t' = C_t\) and \(C_r' = C_r\); if \(C_t \subseteq C_r\) is a typicality inclusion \(T(A) \subseteq C_r\), then \(C_t' = A \cap \boxtimes_{\rightarrow}A\) and \(C_r' = C_r\).

The size of KB\(^4\) is polynomial in the size of the KB. The same for \(C_t'\) and \(C_r'\), assuming the size of \(C_t\) and \(C_r\) be polynomial in the size of KB.

Given the above encoding, in [20] it is shown that (we write KB \(\models_{\mathcal{AXC}} F\) to mean that \(F\) holds in all \(\mathcal{AXC}\) models of KB):

\[
\text{KB} \models_{\mathcal{AXC}+\mathcal{T}_{\text{R}}} C_t \subseteq C_r \text{ if and only if } \text{KB'} \models_{\mathcal{AXC}} C_t' \subseteq C_r'
\]

and, as a consequence, that the problem of deciding entailment in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) is in \(\text{ExpTime}\), since reasoning in \(\mathcal{AXC}\) is \(\text{ExpTime}\)-complete. \(\text{ExpTime}\)-hardness follows from the fact that \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) includes \(\mathcal{AXC}\). In conclusion, the problem of deciding entailment in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) is \(\text{ExpTime}\)-complete. This also holds if cardinality restrictions are taken into account: in [29], it is shown that reasoning in the logic \(\mathcal{AXC}^0\), extending \(\mathcal{AXC}\) with qualified number restrictions, is \(\text{ExpTime}\)-complete also with cardinality restrictions. Therefore, we can again conclude that reasoning in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) with cardinality restrictions is \(\text{ExpTime}\)-complete. However, as far as we know, the only “concrete” decision procedure for \(\mathcal{AXC}^0\) with cardinality restrictions is the set of transformation rules described in [1], which is in \(\text{NExpTime}\).

4.2. Reasoning in \(\mathcal{AXC}+\mathcal{T}_{\text{R}}^{\text{expCl}}\)

We have mentioned that the semantics of the logic \(\mathcal{AXC}+\mathcal{T}_{\text{R}}^{\text{expCl}}\) corresponds to the rational closure of an \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) knowledge base introduced in [21]. Here we recall this machinery, essentially an extension to \(\mathcal{AXC}+\mathcal{T}_{\text{R}}\) of the definition of rational closure introduced by Lehmann and Magidor in [23] for the propositional case. We first consider the rational closure with respect to the TBox, in which essentially we 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 12 (Exceptionality)** Let KB\(=\langle T, A \rangle\) be a knowledge base. A concept C is said to be exceptional for KB if and only if KB \(\models_{\mathcal{AXC}+\mathcal{T}_{\text{R}}} T(\top) \subseteq \neg C\). An inclusion T(C) \(\subseteq D\) 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 \(E(KB)\).

Similarly to the rational closure for propositional logic in [23], 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 in-

---

\(^4\)The results provided in [20] are extended to the more expressive logic \(\mathcal{SHIQ}\). Here we focus our attention on the basic \(\mathcal{AXC}\).
clusions $T(C) \subseteq D$ of KB that are not exceptional for KB. Before we do this, if there is an assertion $T(C)(a) \in A$, 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 $A_i$ below.

**Definition 13** Given $KB=(T, A)$, it is possible to define a sequence of knowledge bases $E_0, E_1, \ldots, E_n$ by letting $E_0 = (T_0, A_0)$ where $T_0 = T$ and $A_0 = A$ and, for $i > 0$, $E_i = (T_i, A_i)$ where

- $T_i = \mathcal{E}(E_{i-1}) \cup \{C \subseteq D \in T \mid T \text{ does not occur in } C\}$
- $A_i = A_{i-1} \cup \{\neg (C \cup D) \mid T(C) \subseteq D \text{ in } (E_{i-1} - E_i) \}$

and there is a $T(B)(a) \in A$ such that $E_{i-1} \models \neg A_i \land \neg T(B)$ and $E_i \models A_i \land \neg T(B)$ for all $j < i - 1$.

(as a consequence of the next Definition 14, these are the $B$s such that $\text{rank}(B) = i - 1$.)

Clearly $T_0 \supseteq T_1 \supseteq T_2, \ldots$, while $A_0 \subseteq A_1 \subseteq A_2, \ldots$. Observe that, being KB finite, there is a least $n \geq 0$ such that, for all $m > n, T_m = T_n$ or $T_m = \emptyset$. We take $(T_n, A_n)$ as the last element of the sequence of knowledge bases starting from KB.

Informally, for the definition of $A_i$, if $T(B)(a) \in A$ (i.e., $a$ is a typical $B$-element), and $B$ has rank $i - 1$, then, for all the inclusions $T(C) \subseteq D$ 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 $\neg (C \cup D)$ must hold.

**Definition 14 (Rank of a concept)** A concept $C$ has rank $i$ (denoted by $\text{rank}(C) = i$) for $KB=(T, A)$, 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 $\text{rank}(C) = \infty$, and we say that $C$ has no rank.

Consider the least $n \geq 0$ such that, for all $m > n, T_m = T_n$ or $T_m = \emptyset$. 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 15 (Rational closure of TBox)** Given $KB=(T, A)$, we define the rational closure $\mathcal{T}$ of $T$, as $\mathcal{T} = \{T(C) \subseteq D \mid \text{either } \text{rank}(C) < \text{rank}(C \cap \neg D) \text{ or rank}(C) = \infty \} \cup \{C \subseteq D \mid KB \models A_i \land \neg T(B)$ $C \subseteq D\}$.

Let us now consider the rational closure of the ABox as defined in [21]:

**Definition 16 (Rational closure of ABox)** Given $KB=(T, A)$, let $a_1, \ldots, a_m$ be the individuals explicitly named in $A$. Let $k_1, k_2, \ldots, k_n$ be all the possible rank assignments to the individuals occurring in $A$.

- Given a rank assignment $k_j$ we define:
  * for each $a_i$: $\mu_j = \{(\neg C \cup D)(a_i) \mid T(C) \subseteq D \text{ in } \mathcal{T}, \text{ and } k_j(a_i) = \text{rank}(C)\} \cup \{(\neg C \cup D)(a_i) \mid T(C) \subseteq D \text{ in } \mathcal{T}\}$;
  * let $\mu^j = \mu^1 \cup \ldots \cup \mu^m$ for all $\mu^1, \ldots, \mu^m$ just calculated for all $a_1, \ldots, a_m$ in $A$.

- We say that $k_j$ is consistent with $(\mathcal{T}, A)$ if:
  * if $T(C)(a_i) \in A$, then $k_j(a_i) = \text{rank}(C)$;
  * $T \cup A \land \mu^j$ is consistent in $A_i \land \neg C$.

- The rational closure of $A(\mathcal{T})$ is the set of all assertions derivable in $A_i \land \neg C$ from $T \cup A \land \mu^j$ for all minimal consistent rank assignments $k_j$:

$$\overline{A} = \bigcap_{k_j, \text{minimal consistent}} \{C(a) \mid T \cup A \land \mu^j \models A_i \land \neg C\}.$$
procedure are a finite set of concepts \( C \) and a query \( F \). If \( F \) is an inclusion \( C \subseteq D \) (even with typicality, i.e. \( C \) has the form \( T^X(C) \)), by Definitions 10 and 11 we just need to reason in \( ALC + T^{Roc}R \) to check whether \( KB \models ALC + T^{Roc}R C \subseteq D \). If \( F \) is an ABox formula of the form \( C(a) \), then we exploit Algorithms 1 and 2. Algorithm 1 checks whether \( F \) is skeptically entailed from the KB in the logic \( ALC + T^{exp}R \), namely whether \( KB \models ALC + T^{exp}R C \models F \). Algorithm 2 is used to check whether \( F \) is credulously entailed from the KB, that is to say \( KB \models ALC + T^{exp}R C \models F \).

By exploiting the procedures described by Algorithms 1 and 2, we can estimate the complexity of reasoning in the logic \( ALC + T^{exp}R \).

**Theorem 1 (Complexity of skeptical entailment)** Given a KB in \( ALC + T^{exp}R \) and a query \( F \) whose size is polynomial in the size of KB, assuming the unary encoding of numbers in cardinality restrictions of KB, the problem of checking skeptically whether KB \( \models ALC + T^{exp}R C \models F \) is \( \text{ExpTime-complete} \).

**Proof.** If \( F \) is an inclusion \( C \subseteq D \), we check whether \( KB \models ALC + T^{Roc}R C \models F \), and we immediately conclude since entailment in \( ALC + T^{Roc}R \) is in \( \text{ExpTime} \). Otherwise, let \( n \) be the size of KB, i.e. the length of the string representing it. Consider the operations computed by Algorithm 1:

- lines 2-5: the algorithm checks, for each concept \( C \subseteq C \) and for each individual name \( a \) of the ABox whether \( T^X(C)(a) \) is minimally entailed from the KB (without cardinality restrictions and degrees of expectedness) in the non-monotonic logic \( ALC + T^{Roc}R \). The number of individual names in the ABox is \( O(n) \). We have assumed that \( C \) contains only concepts belonging to KB, therefore also the size of \( C \) is \( O(n) \). It follows that the number of different \( T^X(C)(a) \) considered is \( O(n^2) \). For each \( T^X(C)(a) \) the algorithm relies on reasoning in \( ALC + T^{Roc}R \), which is in \( \text{ExpTime} \), therefore we make a polynomial number of computation in \( \text{ExpTime} \);
- line 6: the algorithm builds the ordered multiset \( d_A \) of Definition 2: obviously, this operation consists in computing the average of the degrees of expectedness of the inclusions \( T^X(C) \subseteq D \in T \), which are \( O(n) \), for each \( C \subseteq C \), again \( O(n) \). Therefore, this problem can be solved with \( O(n^2) \) operations, i.e. in polynomial time;
- line 7: the algorithm builds the set \( S \) of possible assumptions (Definition 3). We have to consider all possible strings obtained by assuming (or not) each typicality assumption \( T(C)(a) \) that are \( O(n^2) \). Consider a generic string \((d_1, d_2, \ldots, d_n)\). For each \( d_i \), we have two options: we can choose either to not include the corresponding typicality assumption, then \( d_i = 0 \), or to include it, then \( d_i \) corresponds to the average of degrees of expectedness for that concept. So we can build \( 2 \times 2 \times \cdots \times 2 \) different strings, therefore \( O(2^{n^2}) \), that is to say the multiset \( S \) has exponential size in \( n \);
- lines 8-11: the algorithm builds the extensions of the ABox corresponding to strings of \( S \), again an exponential number of extensions \( O(2^{n^2}) \);
- lines 12-15: the algorithm checks, for each extension of the ABox, whether it satisfies cardinality constraints in \( T^{card} \): therefore, the algorithm makes \( O(2^{n^2}) \) calls to a procedure checking the satisfiability of a knowledge base in \( ALC + T^{Roc}R \) plus cardinality restrictions, that is in \( \text{ExpTime} \). Therefore, these problems are in \( \text{ExpTime} \);
- lines 17-18: the algorithm compares each pair of strings \((d_i, d_j)\) of possible assumptions, in order to check whether \( d_i < d_j \) and, therefore, conclude that the extension \( \hat{A}_i \) is more surprising than \( \hat{A}_j \). To this aim, we consider the following procedure. First, elements (integers) of \( d_i \) and \( d_j \) are ordered, and since they are \( O(n^2) \), this can be done in \( O(n^2 \log n^2) \) steps. Call \( d'_i \) and \( d'_j \) the ordered strings. Then, we proceed as follows: an index \( k \) is used to scan elements in \( d'_i \) and \( d'_j \); at each step, if \( d'_i[k] \) (integer in position \( k \) of \( d'_i \)) is \( \leq \) of \( d'_j[k] \), then \( k \) is incremented in order to index the subsequent integer in \( d'_i \) and \( d'_j \); otherwise the procedure answers that \( d'_i \neq d'_j \). If the scan of \( k \) is completed, and all integers of the strings have been considered, the procedure answers that \( d'_i < d'_j \). The procedure requires \( O(n^2) \) steps, and since there are \( O(2^{n^2}) \) strings in \( S \), these operations are performed in \( \text{ExpTime} \);
- line 19: the algorithm builds the set \( \mathcal{E} \) of perfect extensions, that is to say extensions of the ABox that are minimal with respect to the relation \( \prec \) introduced in line 18 by the algorithm. The algorithm checks, for each extension \( \hat{A}_i \), whether it is minimal: to this aim, for each \( \hat{A}_j \), the algorithm checks whether \( \hat{A}_i \not< \hat{A}_j \), and this can be done in constant time (we could think of implementing
Algorithm 1 Skeptical entailment in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$.

1: procedure $\text{SKEPTICAL.Entailment}((T \cup T_{\text{card}}, A), T', F, C)$
2: $\mathbb{C}_A \leftarrow \emptyset$
3: for each $C \in C$ do
4:   for each individual $a \in A$ do
5:     if $(T', A) \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} \{ T(C)(a) \}$ then $\mathbb{C}_A \leftarrow \mathbb{C}_A \cup \{ T(C)(a) \}$
6: $d_A \leftarrow$ build the ordered multiset of avg degrees of Definition 2 given $T$ and $\mathbb{C}_A$
7: $\mathbb{S} \leftarrow$ build strings of possible assumptions as in Definition 3 given $\mathbb{C}_A$ and $d_A$
8: $A_{pl} \leftarrow \emptyset$
9: for each $d_i \in \mathbb{S}$ do
10:   build the extension $\hat{A}_i$ corresponding to $d_i$
11: $A_{pl} \leftarrow A_{pl} \cup \hat{A}_i$
12: $A_{el} \leftarrow \emptyset$
13: for each $\hat{A}_i \in A_{el}$ do
14:   if $(T' \cup T_{\text{card}}, A \cup \hat{A}_i)$ is satisfiable in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$ then
15:     $A_{el} \leftarrow A_{el} \cup \hat{A}_i$
16: for each $\hat{A}_i \in A_{el}$ do
17:   for each $\hat{A}_j \in A_{el}$ do
18:     if $d_i \leq d_j$ then let $\hat{A}_i < \hat{A}_j$
19:     $E \leftarrow \{ \hat{A}_i | \exists \hat{A}_j \in A_{el} \text{ such that } \hat{A}_j < \hat{A}_i \}$
20: for each $\hat{A}_i \in \hat{E}$ do
21:   if $(T' \cup T_{\text{card}}, A \cup \hat{A}_i) \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} F$ then
22:     return $\mathbb{K}_B \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} F$
23: return $\mathbb{K}_B \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} F$

Algorithm 2 Credulous entailment in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$.

1: procedure $\text{CREDOUS.Entailment}((T \cup T_{\text{card}}, A), T', F, C)$
2: $\mathbb{C}_A \leftarrow \emptyset$
3: for each $C \in C$ do
4:   for each individual $a \in A$ do
5:     if $(T', A) \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} \{ T(C)(a) \}$ then $\mathbb{C}_A \leftarrow \mathbb{C}_A \cup \{ T(C)(a) \}$
6: $d_A \leftarrow$ build the ordered multiset of avg degrees of Definition 2 given $T$ and $\mathbb{C}_A$
7: $\mathbb{S} \leftarrow$ build strings of possible assumptions as in Definition 3 given $\mathbb{C}_A$ and $d_A$
8: $A_{pl} \leftarrow \emptyset$
9: for each $d_i \in \mathbb{S}$ do
10:   build the extension $\hat{A}_i$ corresponding to $d_i$
11: $A_{pl} \leftarrow A_{pl} \cup \hat{A}_i$
12: $A_{el} \leftarrow \emptyset$
13: for each $\hat{A}_i \in A_{el}$ do
14:   if $(T' \cup T_{\text{card}}, A \cup \hat{A}_i)$ is satisfiable in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$ then
15:     $A_{el} \leftarrow A_{el} \cup \hat{A}_i$
16: for each $\hat{A}_i \in A_{el}$ do
17:   for each $\hat{A}_j \in A_{el}$ do
18:     if $d_i \leq d_j$ then let $\hat{A}_i < \hat{A}_j$
19:     $E \leftarrow \{ \hat{A}_i | \exists \hat{A}_j \in A_{el} \text{ such that } \hat{A}_j < \hat{A}_i \}$
20: for each $\hat{A}_i \in \hat{E}$ do
21:   if $(T' \cup T_{\text{card}}, A \cup \hat{A}_i) \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} F$ then
22:     return $\mathbb{K}_B \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} F$
23: return $\mathbb{K}_B \models_{\mathcal{ALC} + T_{\text{exp}}^{\text{R}}} F$

$\triangleright$ build the set $\mathbb{S}$ of possible assumptions
$\triangleright$ Reasoning in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$
$\triangleright$ build plausible extensions of $A$
$\triangleright$ select eligible extensions checking cardinality restrictions
$\triangleright$ Reasoning in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$
$\triangleright$ check preference among extensions of $A$
$\triangleright$ select perfect extensions
$\triangleright$ query entailment in $\mathcal{ALC} + T_{\text{exp}}^{\text{R}}$
$\triangleright$ a perfect extension not entailing $F$
$\triangleright$ $F$ is entailed in all perfect extensions
5. Reasoning about Scenarios “in the middle”

The approach described in the previous sections could be extended in order to take also into account all the extensions of the ABox satisfying cardinality restrictions, that is to say the eligible extensions of Definition 6. The idea is to reason about all plausible scenarios, each one equipped with a degree of expectedness, representing a sort of probability, allowing the user to choose the one he considers more adequate for his application, ranging from the most trivial scenario to the most surprising one.

We iteratively define sets of extensions representing scenarios “in the middle”: intuitively, at each step, we consider the extensions of ABox that are minimal w.r.t. the weak preference among extensions of ABox in Definition 8. At the next step, only remaining eligible extensions are considered, and so on. This is formally stated as follows:

**Definition 17 (Extensions “in the middle”)** Given a KB=(\(T, A\)), let \(E\) be the set of all eligible extensions as in Definition 6. We define extensions “in the middle” as follows:

- we let \(E_0 = E\), where \(E \subseteq E\) is the set of all eligible extensions that are minimal as in Definition 9;
- while \(E_i\) is not empty, let \(E_{i+1}\) be the extensions in \(E - (E_0 \cup E_1 \cup \ldots \cup E_i)\) that are minimal with respect to the order relation of Definition 8.

Definition 17 describes a sequence of sets of eligible extensions \(E_0, E_1, \ldots, E_n\) with a degree of expectedness \(i\) associated to each one. We can formally define what we mean for reasoning in more or less surprising scenarios:

**Definition 18 (Entailment in \(ALC + T_{R}^{\exp}\) at degree \(i\))** Given a KB=(\(T, A\)) and a query \(F\), we say that (i) \(F\) is skeptically entailed from KB in \(ALC + T_{R}^{\exp}\) at degree \(i\), for \(i \in \mathbb{N}\), written \(KB \models_{ALC + T_{R}^{\exp}}^{ske}\ F\), if \((T, A \cup \hat{A}) \models_{ALC + T_{R}^{\exp}}^{\hat{A}} F\) for all \(\hat{A} \in E_i\); (ii) \(F\) is credulously entailed from KB in \(ALC + T_{R}^{\exp}\), written \(KB \models_{ALC + T_{R}^{\exp}}^{cr}\ F\), if there exists \(\hat{A} \in E_i\) such that \((T, A \cup \hat{A}) \models_{ALC + T_{R}^{\exp}} F\).

Since it computes all the extensions of an ABox, Algorithm 1 can be slightly modified in order to compute the sets of extensions \(E_0, E_1, \ldots, E_n\) in order to reason in plausible scenarios with a given degree of expectedness, according to Definition 18. Algorithm 3 shows the complete procedure for skeptical reasoning, that can be exploited in order to show that also reasoning about scenarios “in the middle” is EXPTime-complete. The Algorithm for reasoning credulously at
degree $i$ can be easily obtained by replacing lines 26–29 with the following ones:

26:  for each $\hat{A}_i \in E$ do
27:     if $(\mathcal{T}_i \cup \mathcal{T}_{\text{card}}, \mathcal{A} \cup \hat{A}_i) \models_{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$ then
28:         return $KB \models_{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$
29: return $KB \not\models_{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$

Theorem 3 (Complexity of reasoning at degree $i$)

Given a $KB$ in $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$, a query $F$ whose size is polynomial in the size of $KB$, and an integer $i$, assuming the unary encoding of numbers in cardinality restrictions of $KB$, the problem of checking skeptically (resp. credulously) whether $KB \models^{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$ (resp. $KB \not\models^{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$) is EXPTime-complete.

Proof. Algorithm 3 differs from Algorithm 1 by the while loop at starting at line 21, where the sets $E_0, E_1, \ldots, E_k$ of eligible extensions are computed. As in the proof of Theorem 1, we have that the number of different eligible extensions in $A_{ij}$ is $O(2^n)$, so is the maximum number of execution of the body of the while loop. At each step, the algorithm only selects the minimal extensions (among those remaining, thus belonging to the set $E$) with respect to the order $<\text{ among extensions introduced in step 18.}$ We proceed similarly to what done in Algorithm 1 to check, for each extension $\hat{A}_i$, whether it is minimal: for each $\hat{A}_j$, the algorithm checks whether $\hat{A}_i <\hat{A}_j$, and this can be done in constant time (we could think of implementing the relation $<$ by means of a bi-dimensional array, therefore this operation corresponds to a direct access to raw $i$ and column $j$ of such array). For each extension $\hat{A}_i$ (and they are $O(2^n)$) the algorithm considers all the other $O(2^n)$ extensions, therefore the algorithm computes $2^n \times 2^n = 2^{2n}$ steps. In the worst case, there is a total order among eligible extensions, and at each step only one extension is added to the corresponding $E_i$. In this case, the body of the while loop is executed $O(2^n)$ times, each one performing $2^{2n}$ steps, then we have $2^n \times 2^{2n} = 2^{3n}$, thus in EXPTime.

EXPTIME hardness again follows from the fact that the logic $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$ extends standard $\mathcal{ALC}$, which is EXPTime-complete: as for the proof of Theorem 1, we can consider a knowledge base without the $\mathcal{T}$ operator and consider $\mathcal{C} = \emptyset$.

It is easy to observe that reasoning in perfect extensions, corresponding to the most “surprising” scenarios, as in Definitions 10 and 11, is an instance of entailment at degree $i$, namely checking whether a query $F$ is entailed in perfect extensions $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$ corresponds to checking whether $F$ is entailed at degree $i = 0$ with respect to entailment in Definition 18.

A further generalization of the reasoning procedure allows us also to define (credulous and skeptical) entailment in combinations of plausible scenarios. As an example, one could be interested in reasoning in all scenarios $\mathcal{E}_i$ such that $i < k$ for a given and fixed $k$.

As a general case, we define a notion of entailment restricted to scenarios with degrees of expectations between $n$ and $m$, as in the following definition:

Definition 19 (Entailment at degrees $n..m$) Given a $KB = (T, \mathcal{A})$ and a query $F$, we say that (i) $F$ is skeptically entailed from $KB$ in $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$ at degrees $n..m$ if, for all $\mathcal{A} \in \mathcal{E}$, for all $i$ such that $n \leq i \leq m$; (ii) $F$ is credulously entailed from $KB$ in $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$, written $KB \models^{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$, if there exists $i$ such that $n \leq i \leq m$ such that for all $\mathcal{A} \in \mathcal{E}_i$ we have that $(T, \mathcal{A} \cup \mathcal{A}) \models_{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$.

The resulting procedure is formally stated by Algorithm 4, and such a procedure again allows us to show that the complexity of reasoning in the logic $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$ remains EXPTime-complete.

Theorem 4 (Complexity at degrees $n..m$) Given a $KB$ in $\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}$, a query $F$ whose size is polynomial in the size of $KB$, and two integers $n \leq m$, assuming the unary encoding of numbers in cardinality restrictions of $KB$, the problem of checking skeptically (resp. credulously) whether $KB \models^{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$ (resp. $KB \not\models^{\mathcal{ALC} + \mathcal{T}_{\text{R}}^{\exp}} F$) is EXPTime-complete.

Proof. The proof is straightforward, since Algorithm 4 differs from Algorithm 3 only because the eligible extensions taken into account are no longer restricted to those belonging to a single $\mathcal{E}_i$, but extended to those belonging to degrees between $n$ and $m$. To this aim, the while loop starting at line 27 is executed, however it only adds eligible extensions (already computed by the Algorithm) to the set $\mathcal{E}$. ■
6. Related Works

The need for representing prototypical properties and to reason about inheritance with exceptions in DLs has motivated the study of nonmonotonic extensions of DLs. Several nonmonotonic extensions of DLs have been proposed in the literature, essentially based on the integration of DLs with well established nonmonotonic reasoning mechanisms. Here we focus on the approaches that can be considered more similar to the peculiarities of the logic $\mathcal{ALC} + T_{\text{exp}}$ introduced in this work, whereas we refer to [16,19] for a detailed discussion about other extensions of DLs for defeasible inheritance. We consider:

- approaches allowing degrees/priorities among TBox inclusions;
- probabilistic extensions of DLs, allowing to label inclusions (and facts) with degrees representing probabilities.

In [3] an extension of DL with Reiter’s default logic is proposed. Being based on a minimal models mechanism, our approach is strongly related to the one based on circumscription\footnote{A formal and precise comparison between the logics with the typicality operator $T$ and circumscribed knowledge bases can be found in [19].} [6,7,8,10] and to the semantics proposed in [5], where defeasible inclusions are allowed besides classical concept inclusions. The prob-
lem of conflicting defaults and of priorities among inclusions has been studied since seminal works in the context of inheritance networks [12]. In the field of nonmonotonic extensions of DLs, in order to overcome the problem of modeling inheritance with exceptions, extensions of DLs with prioritized defaults have been studied in [4,28], whereas priorities among defeasible extensions of DLs with prioritized defaults have been 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 [26]:

\[
\begin{align*}
0.3 & : \exists \text{hasAnimal.Pet} \sqsubseteq \text{NatureLover} \\
0.6 & : \text{Cat} \sqsubseteq \text{Pet} \\
0.9 & : \text{Cat(tom)} \\
\text{hasAnimal(kevin, tom)}
\end{align*}
\]

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 the following eight possible worlds:

\[
\begin{align*}
&\{(1,0), ((2),0), ((3),0)\} \\
&\{(1), (2), (3), 1\} \\
&\{(1), (2), (3), 0\} \\
&\{(1), (2), (3), 1\} \\
&\{(1), (2), (3), 0\} \\
&\{(1), (2), (3), 1\} \\
&\{(1), (2), (3), 0\} \\
&\{(1), (2), (3), 1\}
\end{align*}
\]

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 \text{hasAnimal.Pet} \sqsubseteq \text{NatureLover} and \text{Cat(tom)}, whereas (2) does not. The query

\[
\text{NatureLover(kevin)}
\]

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 \(P(\text{NatureLover(kevin)}) = 0.3 \times 0.6 \times 0.9 = 0.162\).

To the best of our knowledge, the literature lacks a formalization of surprising scenarios in probabilistic formalizations of knowledge, however it is worth observing that a surprising scenario could be defined as a set of facts with a low probability, then one can
think of restricting the attention to less probable outcomes corresponding, for instance, to worlds where corresponding probabilities are lower than a given threshold. However, as mentioned above, degrees in the logic $\mathcal{ALC} + T^p_\mathcal{R}$ are used in order to represent a level of expectedness in typicality inclusions, whereas in DISPONTE KBs we can express inclusions like (1) above that cannot express prototypical properties.

In Section 3: we can think of representing the typicality inclusions

\[ T(\text{FaceWrestler}) \sqsubseteq_1 \text{RoyalRumbleWinner} \quad (T1) \]
\[ T(\text{Returning}) \sqsubseteq_3 \text{RoyalRumbleWinner} \quad (T2) \]
\[ T(\text{Predicted}) \sqsubseteq_2 \text{RoyalRumbleWinner} \quad (T3) \]

with the following ones in the language of probabilistic DLs:

\[ 0.6 :: \text{FaceWrestler} \sqsubseteq \text{RoyalRumbleWinner} \quad (a) \]
\[ 0.8 :: \text{Returning} \sqsubseteq \text{RoyalRumbleWinner} \quad (b) \]
\[ 0.7 :: \text{Predicted} \sqsubseteq \text{RoyalRumbleWinner} \quad (c) \]

The ABox facts

\[ \text{FaceWrestler}(\text{dean}) \quad \text{FaceWrestler}(\text{roman}) \quad \text{Predicted}(\text{roman}) \quad \text{Returning}(\text{seth}) \]

are all certain axioms, therefore differences among worlds are only introduced by the choices about (a), (b) and (c). Given the cardinality constraint (= 1 $\text{RoyalRumbleWinner}$), it can be observed that all the worlds where (a) is considered as true are inconsistent, since both Roman and Dean are face wrestlers. The world $\{((a), 0), ((b), 0), ((c), 0)\}$ excludes all the TBox inclusions, whereas the world $\{((a), 0), ((b), 1), ((c), 1)\}$ does not allow to fulfill the cardinality restriction too (from (b) and Returning(seth) we conclude $\text{RoyalRumbleWinner}(\text{seth})$ and from (c) and Predicted(roman) we have $\text{RoyalRumbleWinner}(\text{roman})$). The only worlds satisfying the restriction about a single Royal Rumble winner are

\[ \{((a), 0), ((b), 0), ((c), 1)\} \quad \text{and} \quad \{((a), 0), ((b), 1), ((c), 0)\} \]

In the first case, we have that the predicted winner, Roman, wins the match, whereas in the second one the winner is the returning athlete Seth. In none of them, the most surprising outcome in which Dean wins the Royal Rumble match is taken into account.

As an alternative, one can think of formalizing the initial TBox with certain axioms, as follows:

\[ \text{FaceWrestler} \sqsubseteq \text{RoyalRumbleWinner} \]
\[ \text{Returning} \sqsubseteq \text{RoyalRumbleWinner} \]
\[ \text{Predicted} \sqsubseteq \text{RoyalRumbleWinner} \]

and to capture degrees of expectedness with the following probabilistic ABox:

\[ (i)0.1 :: \text{FaceWrestler}(\text{dean}) \quad (ii)0.1 :: \text{FaceWrestler}(\text{roman}) \quad (iii)0.2 :: \text{Predicted}(\text{roman}) \quad (iv)0.3 :: \text{Returning}(\text{seth}) \]

In this case, we have 16 different worlds, representing all possible combinations of ABox facts. Only the following worlds satisfy the cardinality constraint of having only one winner:

\[ \{((i), 1), ((ii), 0), ((iii), 0), ((iv), 0)\} \]
\[ \{((i), 0), ((ii), 1), ((iii), 0), ((iv), 0)\} \]
\[ \{((i), 0), ((ii), 0), ((iii), 1), ((iv), 0)\} \]
\[ \{((i), 0), ((ii), 0), ((iii), 0), ((iv), 1)\} \]

In the first world, Dean is the winner with a probability of 100%, whereas in the last one Seth is the winner with a probability of 30%. In all other worlds, Roman is the winner. It is worth noticing that, with this knowledge base, in order to reason about typicality we need to “weaken” certain facts of the ABox to probabilistic ones, losing the opportunity to reason about them with respect to other properties. For instance, if the TBox further contains $\text{FaceWrestler} \sqsubseteq \text{KidsFavourite}$, then in the logic $\mathcal{ALC} + T^p_\mathcal{R}$ we are still able to infer that kids are fans of both Dean and Roman, since eligible extensions select only typicality assumptions: only one of them is a typical face wrestler, but both are face wrestlers, then obtaining support by the kids. On the contrary, in the DLs with DISPONTE semantics with the above formulation, Dean and Roman are face wrestlers with a certain probability, and this also affects the probabilities of being kids’ favourites or not.

7. Conclusions and Future Issues

In this work we have provided a nonmonotonic procedure for preferential Description Logics in order to reason about surprising scenarios in presence of car-
dinality restrictions on concepts. We have introduced the Description Logic of typicality $\text{ALC} + T_{\text{R}}^{\text{exp}}$, an extension of $\text{ALC}$ with a typicality operator $T$ allowing to (i) express typicality inclusions of the form $T(C) \sqsubseteq_d D$, where $d$ is a positive integer representing a degree of expectedness; (ii) reason in presence of restrictions on the cardinality of concepts. It is worth noticing that the proposed logic $\text{ALC} + T_{\text{R}}^{\text{exp}}$ 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 plausible but unexpected scenarios could help domain experts to achieve their goals, wherever standard reasoning is not enough to do it: as an example, in medical diagnosis, the most likely explanation for a set of symptoms is not always the solution to the problem, whereas reasoning about surprising scenarios could help the medical staff in taking alternative explanations into account. In other words, the logic $\text{ALC} + T_{\text{R}}^{\text{exp}}$ is not intended to replace existing nonmonotonic DLs, but to tile them in order to reason about alternative, plausible scenarios when it is needed to go beyond most likely solutions.

We have also described a procedure for reasoning in $\text{ALC} + T_{\text{R}}^{\text{exp}}$ exploiting reasoning mechanisms in the logics $\text{ALC} + T_{\text{R}}^{\text{D}}$ and $\text{ALC} + T_{\text{R}}^{\text{defCl}}$, the last one relying on a notion of rational closure for Description Logics. This procedure allowed us to show that entailment in $\text{ALC} + T_{\text{R}}^{\text{exp}}$ is ExpTime-complete as the underlying $\text{ALC}$, therefore it is essentially inexpensive, once unary encoding of numbers in cardinality restrictions is assumed. We have also provided an extension of our framework in order to reason about alternative scenarios “in the middle”, ranging from the most surprising to the most expected ones; we have proposed reasoning procedures and ExpTime completeness results also for this extension.

The extension of DLs of typicality with cardinality restrictions is of its own interest, and one can think of considering cardinality restrictions not limited to plausible scenarios of the logic $\text{ALC} + T_{\text{R}}^{\text{exp}}$, but directly applied to nonmonotonic semantics of $\text{ALC} + T_{\text{R}}^{\text{defCl}}$. Furthermore, we aim at studying also cardinality restrictions on roles.

In future work we aim at extending this approach to more expressive Description Logics, in particular those underlying the standard language for ontology engineering OWL. As a first step, in [20] the logic with the typicality operator and the rational closure construction have been applied to the logic $\text{SHIQ}$. 

References


