Two semantic interpretations of probabilities in description logics of typicality

ANTONIO LIETO*, *Università di Salerno, DISPC, Cognition Interaction and Intelligent Technologies Laboratory (CIIT Lab).*

GIAN LUCA POZZATO**, *Università di Torino, Dipartimento di Informatica.*

Abstract

In this work we focus on extensions of Description Logics (DLs) of typicality by means of probabilities. We introduce a novel extension of the logic of typicality $\mathcal{ALC} + \mathbf{T_R}$, able to represent and reason about typical properties and defeasible inheritance in DLs. The novel logic ($\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$: Typical $\mathcal{ALC}$ with Probabilities as Proportions) allows inclusions of the form $\mathbf{T}(C) \sqsubseteq^{p} D$, with probability p representing a proportion, meaning that "all the typical Cs are Ds, and the probability that a C element is not a D element is $1 - p$". We also compare and confront this novel logic with a similar one already presented in the literature ($\mathbf{T}^{\mathsf{CL}}$, introduced in Lieto and Pozzato (2020, *J. Exp. Theor. Artif. Intell.*, 32, 769–804)), inspired by the DISPONTE semantics and that allows inclusions of the form $p\ :\ \mathbf{T}(C) \sqsubseteq D$ with probability p, where p represents a degree of belief, whose meaning is that "we believe with a degree p that typical Cs' are also Ds.". We then show that the proposed $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ extension (like the previous $\mathbf{T}^{\mathsf{CL}}$) can be applied in order to tackle a specific and challenging problem in the field of common-sense reasoning, namely the combination of prototypical concepts, that have been shown to be problematic to model for other symbolic approaches like fuzzy logic. We show that, for the proposed extension, the complexity of reasoning remains EXPTIME-complete as for the underlying standard monotonic DL $\mathcal{ALC}$.

Keywords: Description logics, typicality, nonmonotonic reasoning, probabilistic reasoning, commonsense reasoning.

1 Introduction

Nonmonotonic extensions of Description Logics (DLs) have been actively investigated since the early 90s [1, 3, 4–7, 9, 39] 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 [12, 13, 15]: in this approach 'typical' or 'normal' properties can be directly specified by means of a 'typicality' operator $\mathbf{T}$ enriching the underlying DL. As a difference with standard DLs, one can consistently express typical properties of a class, admitting the presence of exceptional elements and reason about defeasible inheritance as well. For instance, a knowledge base can consistently express that 'normally, a wrestler is fit', whereas 'a typical sumo wrestler is not fit' as follows:

$$SumoWrestler \sqsubseteq Wrestler$$
$$\mathbf{T}(Wrestler) \sqsubseteq Fit$$
$$\mathbf{T}(SumoWrestler) \sqsubseteq \neg Fit$$

*E-mail: alieto@unisa.it
**E-mail: gianluca.pozzato@unito.it

Vol. 00, No. 00,
https://doi.org/10.1093/jigpal/jzae086

The key point of this approach relies on the fact that the semantics of the typicality operator $\mathbf{T}$ is essentially a restatement of the semantics of nonmonotonic entailment as defined by Kraus, Lehmann and Magidor in the early nineties [18], corresponding to a set of properties that any concrete nonmonotonic reasoning mechanism should satisfy. Furthermore, the semantics of $\mathcal{ALC} + \mathbf{T_R}$ [15] is strengthened by a minimal model machinery corresponding to a notion of *rational closure*, which extends to this context the notion introduced by Lehmann and Magidor in [21] for the propositional case.

The resulting nonmonotonic DL, called $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\textsc{RaCl}}$, allows the user to perform some useful nonmonotonic inferences about defeasible inheritance in presence of exceptions. For instance, in the above example, the logic allows one to infer that a typical wrestler is not a sumo wrestler ($\mathbf{T}(\mathit{Wrestler}) \sqsubseteq \neg \mathit{SumoWrestler}$), and that a typical swiss sumo wrestler is not fit ($\mathbf{T}(\mathit{SumoWrestler} \sqcap \mathit{Swiss}) \sqsubseteq \neg \mathit{Fit}$), since being Swiss is irrelevant with respect to being fit or not. Moreover, if one knows that Claudio is a wrestler ($\mathit{Wrestler}(\mathit{claudio})$), then the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\textsc{RaCl}}$ allows one to infer that Claudio is fit ($\mathit{Fit}(\mathit{claudio})$); however, if we discover that Claudio is a sumo wrestler, i.e. $\mathit{SumoWrestler}(\mathit{claudio})$ is added to the knowledge base, then the previously inferred information is retracted, whereas the conclusion $\neg \mathit{Fit}(\mathit{claudio})$ can be derived.

The logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\textsc{RaCl}}$ imposes to consider *all* typicality assumptions that are consistent with a given knowledge base, capturing the attitude of human beings of assuming that, in case there is no evidence to the contrary, if an individual a is a member of the category/class C, then he is a typical one. This seems to be too strong in several application domains, especially when the domain contains a lot of individuals: in these cases, to assume that *each and every* individual is a typical member of the classes he belongs to—in other words, if not explicitly stated, there are no exceptions—does not seem to be the most adequate solution. As an example, when reasoning about a domain containing hundreds of wrestlers, the assumption that they are all typical ones being fit, without exceptions, seems to be too strong and counter-intuitive.

In several application domains, it could be useful to reason about scenarios being plausible but, in some sense *surprising*, containing exceptions to typical properties. Furthermore, one could need to express *probabilities* about typicality inclusions, in order to also reason about the probability of inferred information as well as to restrict reasoning to scenarios whose probabilities belong to a given and fixed range.

In this work, we tackle this limitations of the DL of typicality by introducing a new DL called $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$, which extends $\mathcal{ALC}$ by means of typicality inclusions equipped by *probabilities of exceptions* of the form $\mathbf{T}(C) \sqsubseteq^p D$, where $p \in (0.5, 1)$, whose intuitive meaning is:

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

intending probabilities as *proportions*, more precisely capturing the fact that we have a p% of elements belonging to the class/concept C being exceptional, i.e. not belonging to D, with respect to the set of all members of C. 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 with respect to the cardinality of the set of C elements 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. From a semantic point of view, as in the underlying DL of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCl}}$, standard DLs models are enriched with a preference relation $<$ among domain elements, whose intuitive meaning is that $x < y$ means that the element x is 'more normal' than the element y. Typical elements of a concept C are then those elements belonging to the extension of C that are minimal with respect to the preference relation $<$. Given a knowledge base, we then restrict our attention only to models satisfying it, in particular probabilities/proportions of typicality inclusions. In order to perform useful nonmonotonic inferences, the resulting semantics is further strengthened by a minimal models semantics: we introduce a preference relation among models—intuitively, a model is preferred to another one if it contains fewer exceptional individuals—then we restrict entailment to minimal models with respect to such a preference relation.

We show that the novel proposed logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ is essentially inexpensive, in the sense that reasoning in them is ExpTime-complete as for the underlying standard DL $\mathcal{ALC}$. Furthermore, we show that such logic is a promising candidate to tackle the problem of *typicality-based concept combination* by using a different interpretation of the probabilistic statement if compared to other formalisms like $\mathbf{T}^{\mathrm{CL}}$. In particular, the problem of commonsense compositionality under focus consists in handling the harmonization of two conflicting requirements that are hardly accommodated in symbolic systems, namely the need of a syntactic and semantic compositionality (typical of logical systems) and that one concerning the exhibition of typicality effects. The problem can be summarized by using a well-known argument in the context of formal and cognitive semantics [31], namely the fact that prototypes (i.e. typical representations of concepts) are not compositional. Consider, for example, a concept like *pet fish*. It results from the composition of the concept *pet* and of the concept *fish*. However, the prototype of *pet fish* cannot result from the composition of the prototypes of a pet and a fish, e.g. a typical pet is furry and warm, a typical fish is grayish, but a typical pet fish is neither furry and warm nor grayish (typically, it is red).

The plan of the paper is as follows. In Section 2 we recall the DL of typicality called $\mathbf{T}^{\mathrm{CL}}$, representing the first logic proposing a solution to the pet fish problem, where probabilities are intended as degrees of belief. In Section 3 we present the novel DL of typicality called $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$, where probabilities are intended as proportions. In Section 5 we show that both the logics are suitable for accounting a specific aspect of commonsense reasoning, namely the problem of typicality-based concept combination. We conclude with a discussion, a comparison with related approaches as well as with some pointers to plausible future developments in Section 6.

2 Probabilities as degree of belief: the logic $\mathbf{T}^{\mathrm{CL}}$

Before introducing the novel proposed probabilistic DL (and in order to better understand the differences from a semantic view point of our novel proposed approach), we recall in this chapter the logic $\mathbf{T}^{\mathrm{CL}}$ (introduced in [23, 24]) that has been the first one to propose a solution for the commonsense compositionality problem entailed by the pet fish phenomenon [10]. It combines the semantics based on the rational closure of $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ [12, 13, 15] with the DISPONTE semantics [35, 36] of probabilistic DLs.

By taking inspiration from the heterogeneous proxytypes hypothesis [22], in our representational assumptions we consider two types of properties associated to a given concept: rigid and typical. Rigid properties are those defining a concept, e.g. $C \sqsubseteq D$ (all Cs are Ds). Typical properties are represented by inclusions equipped by a degree of belief expressed through probabilities with their epistemic meaning, like in the DISPONTE semantics. Before providing a formal presentation of the logic $\mathbf{T}^{\mathrm{CL}}$, we provide an example extending the knowledge base of the Introduction.

Let us consider a knowledge base containing the following inclusion relations and facts about individuals:

$$\mathit{SumoWrestler} \sqsubseteq \mathit{Athlete} \quad (1)$$
$$\mathit{Athlete} \sqsubseteq \mathit{HumanBeing} \quad (2)$$
$$0.8 \;:\; \mathbf{T}(\mathit{Athlete}) \sqsubseteq \mathit{Fit} \quad (3)$$
$$0.8 \;:\; \mathbf{T}(\mathit{SumoWrestler}) \sqsubseteq \neg \mathit{Fit} \quad (4)$$
$$0.95 \;:\; \mathbf{T}(\mathit{Athlete}) \sqsubseteq \mathit{YoungPerson} \quad (5)$$
$$\mathit{Athlete}(\mathit{roberto}) \quad (6)$$
$$\mathit{SumoWrestler}(\mathit{hiroyuki}) \quad (7)$$

Inclusions (1) and (2) are intended as usual in standard $\mathcal{ALC}$: all sumo wrestlers are athletes, and all athletes are human beings. Typicality properties (3), (4) and (5) represent the following facts, respectively:

- usually, athletes are fit, and this is believed with a probability of 80%;
- typical sumo wrestlers are not fit, and this is believed with a probability of 80%; and
- we believe with a probability/degree of 95% that, normally, athletes are young persons.

The ABox facts (6) and (7) are used to represent that Roberto is an athlete, whereas Hiroyuky is a sumo wrestler.

The language of the logic $\mathbf{T}^{\mathsf{CL}}$ extends the basic DL $\mathcal{ALC}$ by *typicality inclusions* of the form $\mathbf{T}(C) \sqsubseteq D$ equipped by a real number $p \in (0.5, 1]$, representing its probability with its epistemic meaning, that is to say "we believe with a degree/probability p that, normally, Cs are also D" [1].

DEFINITION 1 (Language of $\mathbf{T}^{\mathsf{CL}}$).
We consider an alphabet of concept names $\mathtt{C}$, of role names $\mathtt{R}$, and of individual constants $\mathtt{O}$. Given $A \in \mathtt{C}$ and $R \in \mathtt{R}$, we define:

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

We define a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ where:

- $\mathcal{R}$ is a finite set of rigid properties of the form $C \sqsubseteq D$;
- $\mathcal{T}$ is a finite set of typicality properties of the form

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

  where $p \in (0.5, 1] \subseteq \mathbb{R}$ is the probability/degree of belief of the typicality inclusion; and
- $\mathcal{A}$ is the ABox, i.e. a finite set of formulas of the form either $C(a)$ or $R(a, b)$, where $a, b \in \mathtt{O}$ and $R \in \mathtt{R}$.

Following from the DISPONTE semantics, each axiom is independent from each other. This avoids the problem of dealing with probabilities of inconsistent inclusions. Let us consider the following knowledge base:

$\mathit{WorkingStudent} \sqsubseteq \mathit{Student}$
(i) $0.8 \;:\; \mathbf{T}(\mathit{Student}) \sqsubseteq \neg \mathit{WorkingTaxesPayer}$
(ii) $0.9 \;:\; \mathbf{T}(\mathit{WorkingStudent}) \sqsubseteq \mathit{WorkingTaxesPayer}$

[1] The reason why we only allow typicality inclusions equipped with probabilities $p > 0.5$ is detailed in Section 4.

Also, in the scenarios where both the conflicting typical inclusions (*i*) and (*ii*) are considered, the two probabilities describe, respectively, the degree of belief 0.8 in that typical students do not pay working taxes, and that we believe with degree 0.9 that, normally, working students do pay working taxes, and those probabilistic inclusions are both acceptable due to the independence assumption. The two probabilities will contribute to a definition of probability of such scenario (as we will describe in Definition 6). It is worth noticing that the underlying logic of typicality allows us to get for free the correct way of reasoning in this case, namely if the ABox contains the information that *Mark* is a working student, we obtain that he pays working taxes, i.e. $\mathit{WorkingTaxesPayer}(\mathit{mark})$.

A model $\mathcal{M}$ in the logic $\mathbf{T}^{\mathsf{CL}}$ extends standard $\mathcal{ALC}$ models by a preference relation among domain elements as in the logic of typicality [15, 16]. 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[2] . 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 normal* than x. Formally:

DEFINITION 2 (Model of $\mathbf{T}^{\mathsf{CL}}$).
A model $\mathcal{M}$ is any structure

$$\langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$$

where:

- $\Delta^{\mathcal{I}}$ is a nonempty set of items called 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 $\Delta^{\mathcal{I}}$; and
- $.^{\mathcal{I}}$ is the extension function that maps each atomic concept C to $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and each role R to $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$, and is extended to complex concepts as follows:
  - $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
  - $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$
  - $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$
  - $(\exists R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \exists (x, y) \in R^{\mathcal{I}}$ such that $y \in C^{\mathcal{I}}\}$
  - $(\forall R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \forall (x, y) \in R^{\mathcal{I}}$ we have $y \in C^{\mathcal{I}}\}$
  - $(\mathbf{T}(C))^{\mathcal{I}} = Min_<(C^{\mathcal{I}})$, where $Min_<(C^{\mathcal{I}}) = \{x \in C^{\mathcal{I}} \mid \nexists y \in C^{\mathcal{I}}$ s.t. $y < x\}$.

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 [15]: the rank of x is the length of the longest chain $x_0 < \cdots < x$ from x to a minimal x_0, i.e. such that there is no x' such that $x' < x_0$. The rank function $k_{\mathcal{M}}$ and $<$ can be defined from each other by letting $x < y$ if and only if $k_{\mathcal{M}}(x) < k_{\mathcal{M}}(y)$.

[2] It could be possible to consider an alternative semantics whose models are equipped with multiple preference relations, whence with multiple typicality operators. In this case, it should be possible to distinguish different aspects of exceptionality, however the approach based on a single preference relation in [15] ensures good computational properties (reasoning in the resulting nonmonotonic logic $\mathcal{ALC} + \mathbf{T_R}$ has the same complexity of the standard $\mathcal{ALC}$), whereas adopting multiple preference relations could lead to higher complexities.

DEFINITION 3 (Model satisfying a knowledge base in $\mathbf{T}^{\text{CL}}$).
Let $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a KB as above. Given a model $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$, we assume that $.^{\mathcal{I}}$ is extended to assign a domain element $a^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$ to each individual constant a of $\mathtt{O}$. We say that:

- $\mathcal{M}$ satisfies $\mathcal{R}$ if, for all $C \sqsubseteq D \in \mathcal{R}$, we have $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$;
- $\mathcal{M}$ satisfies $\mathcal{T}$ if, for all $q\ :\ \mathbf{T}(C) \sqsubseteq D \in \mathcal{T}$, we have $\mathbf{T}(C)^{\mathcal{I}} \subseteq D^{\mathcal{I}}$, i.e. $Min_{<}(C^{\mathcal{I}}) \subseteq D^{\mathcal{I}}$; and
- $\mathcal{M}$ satisfies $\mathcal{A}$ if, for all assertion $F \in \mathcal{A}$, if $F = C(a)$ then $a^{\mathcal{I}} \in C^{\mathcal{I}}$, otherwise if $F = R(a,b)$ then $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$.

Let us now define the notion of *scenario*. Intuitively, a scenario is a knowledge base obtained by adding to all rigid properties in $\mathcal{R}$ and to all ABox facts in $\mathcal{A}$ only *some* typicality properties. More in detail, we define an *atomic choice* on each typicality inclusion, then we define a *selection* as a set of atomic choices in order to select which typicality inclusions have to be considered in a scenario.

DEFINITION 4 (Atomic choice).
Given $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{E_1 = q_1\ :\mathbf{T}(C_1) \sqsubseteq D_1, \dots, E_n = q_n\ :\mathbf{T}(C_n) \sqsubseteq D_n\}$ we define $(E_i,\, k_i)$ an *atomic choice*, where $k_i \in \{0,1\}$.

DEFINITION 5 (Selection).
Given $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{E_1 = q_1\ :\mathbf{T}(C_1) \sqsubseteq D_1, \dots, E_n = q_n\ :\mathbf{T}(C_n) \sqsubseteq D_n\}$ and a set of atomic choices ν, we say that ν is a *selection* if, for each E_i, one decision is taken, i.e. either $(E_i,\, 0) \in \nu$ and $(E_i,\, 1) \notin \nu$ or $(E_i,\, 1) \in \nu$ and $(E_i,\, 0) \notin \nu$ for $i = 1,2,\dots,n$. The probability of ν is $P(\nu) = \prod_{(E_i,1)\in\nu} q_i \prod_{(E_i,0)\in\nu} (1-q_i)$.

DEFINITION 6 (Scenario).
Given $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{E_1 = q_1\ :\mathbf{T}(C_1) \sqsubseteq D_1, \dots, E_n = q_n\ :\mathbf{T}(C_n) \sqsubseteq D_n\}$ and given a selection σ, we define a *scenario* $w_\sigma = \langle \mathcal{R}, \{E_i \mid (E_i, 1) \in \sigma\}, \mathcal{A} \rangle$. We also define the probability of a scenario w_σ as the probability of the corresponding selection, i.e. $P(w_\sigma) = P(\sigma)$. Last, we say that a scenario is *consistent* with respect to $\mathcal{K}$ when it admits a model in the logic $\mathbf{T}^{\text{CL}}$ satisfying $\mathcal{K}$.

We denote with $\mathcal{W}_{\mathcal{K}}$ the set of all scenarios. It immediately follows that the notion of probability of a scenario $P(w_\sigma)$ introduces a probability distribution over scenarios, that is to say $\sum_{w\in\mathcal{W}_{\mathcal{K}}} P(w) = 1$.

Given a query F, either an inclusion relation or an ABox fact, and a scenario w, we say that F is entailed by w when F is entailed in the underlying DL of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\text{RaCl}}$ from the knowledge base obtained from w by removing probabilities in typicality inclusions. We recall that, given a knowledge base $\mathcal{K}$, in [15] it is defined that a query F is entailed from $\mathcal{K}$, written $\mathcal{K} \models F$, if F holds in all minimal canonical models satisfying $\mathcal{K}$. Moreover, we define the probability of F given a knowledge base $\mathcal{K}$ as the sum of the probabilities of scenarios from which F is entailed.

DEFINITION 7
Given a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ and a query F, let w be a scenario as in Definition 6. Let w' be the knowledge base $\langle \mathcal{R} \cup \mathcal{T}', \mathcal{A} \rangle$ in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\text{RaCl}}$, where $\mathcal{T}' = \{\mathbf{T}(C) \sqsubseteq D \mid p\ :\ \mathbf{T}(C) \sqsubseteq D \in \mathcal{T}\}$. We say that F is entailed from w, written $w \models_{\mathbf{T}^{\text{CL}}} F$, in the logic $\mathbf{T}^{\text{CL}}$ if F is entailed from w' in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\text{RaCl}}$, i.e. $w' \models F$.

DEFINITION 8
Given a knowledge base $\mathcal{K}$ and a query F, we say that F is entailed from $\mathcal{K}$ with probability p, written $\mathcal{K} \models_p F$, if F is entailed in $\mathbf{T}^{\mathsf{CL}}$ in scenarios $w_1, w_2, \ldots, w_n$, i.e. $w_1 \models_{\mathbf{T}^{\mathsf{CL}}} F$, $w_2 \models_{\mathbf{T}^{\mathsf{CL}}} F, \ldots$, $w_n \models_{\mathbf{T}^{\mathsf{CL}}} F$, and $P(w_1) + P(w_2) + \cdots + P(w_n) = p$.

Let us extend the example of the Introduction about wrestlers and sumo wrestlers with the increased expressive power of the logic $\mathbf{T}^{\mathsf{CL}}$:

$$\begin{array}{ll} \mathit{SumoWrestler} \sqsubseteq \mathit{Wrestler} & \\ 0.6 \ : \ \mathbf{T}(\mathit{Wrestler}) \sqsubseteq \mathit{Fit} & (1) \\ 0.9 \ : \ \mathbf{T}(\mathit{Wrestler}) \sqsubseteq \mathit{Young} & (2) \\ 0.9 \ : \ \mathbf{T}(\mathit{SumoWrestler}) \sqsubseteq \neg \mathit{Fit} & (3) \end{array}$$

The three typicality inclusions allow one to describe eight different scenarios. As an example, the scenario including (1) and (3) has a probability $0.6 \times (1-0.9) \times 0.9 = 0.054$: in this scenario, if the ABox contains the fact that Seth is both a wrestler and a sumo wrestler, i.e. $\mathit{Wrestler}(\mathit{seth})$ and $\mathit{SumoWrestler}(\mathit{seth})$, it is entailed that Seth is not fit, i.e. $\neg \mathit{Fit}(\mathit{seth})$. In another scenario, where (1) and (2) are considered whereas (3) is discarded, whose probability is $0.6 \times 0.9 \times (1-0.9) = 0.054$, given the same ABox we infer that Seth is fit, and he is also young ($\mathit{Fit}(\mathit{seth})$ and $\mathit{Young}(\mathit{seth})$). We can also observe that $\mathcal{K} \models_{0.9} \neg \mathit{Fit}(\mathit{seth})$; indeed, such a query holds in the four scenarios including (3), therefore the probability of such a query is $0.054 + 0.324 + 0.486 + 0.036 = 0.9$.

In [24] we have provided a proof showing that $\mathbf{T}^{\mathsf{CL}}$ is computationally inexpensive compared to $\mathcal{ALC}$ in the sense that reasoning remains *Exptime* complete as in the underlying standard monotonic version of the logic.

3 Probabilities as proportion: the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$

In this section we introduce the second extension of the logic of typicality, where probabilities are intended as *proportions*. In this logic, that we call $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, typicality inclusions have the form

$$\mathbf{T}(C) \sqsubseteq^p D,$$

where p is a real number between 0.5 and 1 representing the probability of having atypical C-elements not being D, namely being exceptional with respect to the typical property D. Differently from the logic $\mathbf{T}^{\mathsf{CL}}$, here p does not represent a degree of belief about the inclusion, rather it provides the proportion between typical and atypical members of a given class/concept with respect to a specific property. Before introducing the formalism in detail, let us consider the following example.

In the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ we can have a knowledge base containing the following inclusions:

$$\begin{array}{ll} \mathit{Bird} \sqsubseteq \mathit{Vertebrate} & (1) \\ \mathit{Penguin} \sqsubseteq \mathit{Bird} & (2) \\ \mathbf{T}(\mathit{Bird}) \sqsubseteq^{0.7} \neg \mathit{Swim} & (3) \\ \mathbf{T}(\mathit{Bird}) \sqsubseteq^{0.9} \neg \mathit{BackwardFlier} & (4) \\ \mathbf{T}(\mathit{Penguin}) \sqsubseteq^{0.95} \mathit{Swim} & (5) \end{array}$$

Inclusions (1) and (2) are rigid properties with the usual semantics, representing that *all* birds are vertebrate and that penguins are birds, with no exceptions. The typicality inclusions (3) and (4) represent two typical properties of birds, namely that they usually do not swim and that are not able to fly backwards, respectively. In (3) the probability equipping the inclusion is 0.7, whereas in

(4) it is 0.9: this means that we have a higher probability of having exceptional birds swimming, 30% coming from 0.7, with respect to having exceptional birds able to fly backwards, and 10% coming from 0.9. Intuitively, (3) represents that the proportion between the number of atypical birds swimming (e.g. penguins, ducks) and the cardinality of the set of all birds, is at most $1 - 0.7 = 0.3$; similarly, (4) represents that the proportion between the number of atypical birds able to fly backwards (e.g. the hummingbirds) and the cardinality of the set of all birds does not exceed $1 - 0.9 = 0.1$. As we will formally describe in Definition 11, in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ and differently from $\mathbf{T}^{\mathsf{CL}}$, we restrict our concern to models where such proportions are fulfilled. In the example, for instance, we discard scenarios/models where the number of exceptional birds swimming is higher than 30%. As in the underlying logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{RaCL}}$, the inclusion (5) allows one to represent and reason about defeasible inheritance, by expressing that, normally, penguins are able to swim, and exceptional penguins not swimming (for instance, because they are ill) are, in proportion, no more than the 5%.

Let us now formally introduce the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$. The language of $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ extends the basic DL $\mathcal{ALC}$ by *typicality inclusions* of the form $\mathbf{T}(C) \sqsubseteq^p D$, where p is a real number $p \in (0.5, 1)$, representing its probability, whose meaning is that 'typical Cs are also Ds, and the probability of finding atypical Cs not being Ds is $1 - p$'. Observe that, as a difference with the logic $\mathbf{T}^{\mathsf{CL}}$, the probability p is chosen in an open interval $(0.5, 1)$. Indeed, if the right extreme 1 were allowed, as we will see, the typicality inclusion $\mathbf{T}(C) \sqsubseteq^1 D$ would collapse into a classic inclusion $C \sqsubseteq D$. It is also worth noticing that, in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, *all typical* Cs are Ds: i.e. we do not allow the presence of a typical C not being D.

DEFINITION 9 (Language of $\mathbf{T}^{\mathsf{CL}}$).
We consider an alphabet of concept names $\mathtt{C}$, of role names $\mathtt{R}$, and of individual constants $\mathtt{O}$. Given $A \in \mathtt{C}$ and $R \in \mathtt{R}$, we define:

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

We define a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ where:

- $\mathcal{R}$ is a finite set of rigid properties of the form $C \sqsubseteq D$;
- $\mathcal{T}$ is a finite set of typicality properties of the form

$$\mathbf{T}(C) \sqsubseteq^p D$$

  where $p \in (0.5, 1) \subseteq \mathbb{R}$ is the probability of not finding exceptions to the typicality inclusion; and
- $\mathcal{A}$ is the ABox, i.e. a finite set of formulas of the form either $C(a)$ or $R(a, b)$, where $a, b \in \mathtt{O}$ and $R \in \mathtt{R}$.

A model $\mathcal{M}$ in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ is similar to the one of the logic $\mathbf{T}^{\mathsf{CL}}$, since it extends standard $\mathcal{ALC}$ models by a preference relation among domain elements as in the logic of typicality [15]. Again, $x < y$ 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 normal* than x.

DEFINITION 10 (Model of $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$).
A model $\mathcal{M}$ is any structure (interpretation) ($\langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ where $\Delta^{\mathcal{I}}$ is a nonempty set of items called the domain; $<$ is an irreflexive, transitive, well-founded and modular relation over $\Delta^{\mathcal{I}}$; $.^{\mathcal{I}}$ is the

extension function that maps each atomic concept C to $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and each role R to $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ and is extended to complex concepts as in Definition 2, in particular for the typicality operator:

$$(\mathbf{T}(C))^{\mathcal{I}} = Min_<(C^{\mathcal{I}}), \text{ where } Min_<(C^{\mathcal{I}}) = \{x \in C^{\mathcal{I}} \mid \nexists y \in C^{\mathcal{I}} \text{ s.t. } y < x\}.$$

As in the logic $\mathbf{T}^{\mathsf{CL}}$, 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 [15]: the rank of x is the length of the longest chain $x_0 < \cdots < x$ from x to a minimal x_0, i.e. such that there is no x' such that $x' < x_0$. The rank function $k_{\mathcal{M}}$ and $<$ can be defined from each other by letting $x < y$ if and only if $k_{\mathcal{M}}(x) < k_{\mathcal{M}}(y)$. As a difference with $\mathbf{T}^{\mathsf{CL}}$, a model so defined is a model of a given knowledge base only if proportions defined by probabilities equipping typicality inclusions are satisfied. This is formally stated as follows:

DEFINITION 11 (Model satisfying a knowledge base in $\mathcal{ALCT}^{\mathsf{P}}$).
Let $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a KB. Given a model $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$, we assume that $.^{\mathcal{I}}$ is extended to assign a domain element $a^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$ to each individual constant a of $\mathtt{O}$. We say that:

- $\mathcal{M}$ satisfies $\mathcal{R}$ if, for all $C \sqsubseteq D \in \mathcal{R}$, we have $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$;
- $\mathcal{M}$ satisfies $\mathcal{T}$ if, for all $\mathbf{T}(C) \sqsubseteq^p D \in \mathcal{T}$, we have that:
  1. $\mathbf{T}(C)^{\mathcal{I}} \subseteq D^{\mathcal{I}}$, i.e. $Min_<(C^{\mathcal{I}}) \subseteq D^{\mathcal{I}}$
  2. $$\frac{\mid \{x \in C^{\mathcal{I}} \mid x \notin (\mathbf{T}(C))^{\mathcal{I}} \text{ and } x \in (\neg D)^{\mathcal{I}}\} \mid}{\mid C^{\mathcal{I}} \mid} \leq 1 - p;$$
- $\mathcal{M}$ satisfies $\mathcal{A}$ if, for all assertion $F \in \mathcal{A}$, if $F = C(a)$ then $a^{\mathcal{I}} \in C^{\mathcal{I}}$, otherwise if $F = R(a, b)$ then $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$.

As mentioned before, in $\mathcal{ALCT}^{\mathsf{P}}$ we do not allow inclusions equipped with the probability 1, otherwise such inclusions would correspond to a rigid one, as stated by the following proposition[3] .

PROPOSITION 1
Given a model $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$, we have that $\mathcal{M}$ satisfies $\mathbf{T}(C) \sqsubseteq^1 D$ if and only if $\mathcal{M}$ satisfies $C \sqsubseteq D$.

PROOF. By Definition 11, since $\mathcal{M}$ satisfies $\mathbf{T}(C) \sqsubseteq^1 D$ we have that

$$\frac{\mid \{x \in C^{\mathcal{I}} \mid x \notin (\mathbf{T}(C))^{\mathcal{I}} \text{ and } x \in (\neg D)^{\mathcal{I}}\} \mid}{\mid C^{\mathcal{I}} \mid} \leq 1 - 1 = 0$$

that is to say, $\mid \{x \in C^{\mathcal{I}} \mid x \notin (\mathbf{T}(C))^{\mathcal{I}} \text{ and } x \in (\neg D)^{\mathcal{I}}\} \mid = 0$, in other words the set $\alpha = \{x \in C^{\mathcal{I}} \mid x \notin (\mathbf{T}(C))^{\mathcal{I}} \text{ and } x \in (\neg D)^{\mathcal{I}}\}$ of elements not belonging to $(\mathbf{T}(C))^{\mathcal{I}}$ but belonging to $(\neg D)^{\mathcal{I}}$ is empty. For each $y \in C^{\mathcal{I}}$, we distinguish two cases: (i) y is a typical C, i.e. $y \in (\mathbf{T}(C))^{\mathcal{I}}$: in this case, by Definition 10, it follows that $y \in D^{\mathcal{I}}$; (ii) y is an atypical C, i.e. $y \notin (\mathbf{T}(C))^{\mathcal{I}}$: since the set α is empty, we have that necessarily $y \in D^{\mathcal{I}}$. Therefore, for each $y \in C^{\mathcal{I}}$, we have $y \in D^{\mathcal{I}}$, that is to say $\mathcal{M}$ satisfies $C \sqsubseteq D$.

[3] In other words: in this case we have that every inclusion of the form $\mathbf{T}(C) \sqsubseteq^1 D$ is replaced by $(C) \sqsubseteq D$ as shown possible by the following proposition.

For the other direction, we have that, if $\mathcal{M}$ satisfies $C \sqsubseteq D$, this means that all elements of $C^{\mathcal{I}}$ also belong to $D^{\mathcal{I}}$, therefore: (i) condition 1 in Definition 11 holds, since $(\mathbf{T}(C))^{\mathcal{I}} \subseteq C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$; (ii) condition 2 in Definition 11 holds, since the set α of elements of C that are not D is empty, therefore $\mid \{x \in C^{\mathcal{I}} \mid x \notin (\mathbf{T}(C))^{\mathcal{I}} \text{ and } x \in (\neg D)^{\mathcal{I}}\} \mid = 0$. □

In order to perform useful nonmonotonic inferences like those described in the Introduction, capturing specificity and some forms of irrelevance, as in [15], we describe a true nonmonotonic reasoning machinery on top of the semantics introduced here above, corresponding to an extension to DLs of the well-established mechanism of *rational closure* introduced in [21]. The nonmonotonic semantics relies on minimal rational models that minimize the rank of domain elements. Informally, given two models $\mathcal{M}_1$ and $\mathcal{M}_2$ of KB, the first one in which a given domain element x has rank $k_{\mathcal{M}_1}(x) = 2$ (because for instance $z < y < x$), with $k_{\mathcal{M}_1}(y) = 1$ and $k_{\mathcal{M}_1}(z) = 0$, and the other one in which $k_{\mathcal{M}_2}(x) = 1$ (because only $y < x$ with $k_{\mathcal{M}_2}(y) = 0$), we prefer $\mathcal{M}_2$, as in this model the element x is assumed to be 'more typical' than in the former.

DEFINITION 12
Given $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ and $\mathcal{M}' = \langle \Delta^{\mathcal{I}'}, <', .^{\mathcal{I}'} \rangle$ we say that $\mathcal{M}$ is preferred to $\mathcal{M}'$, written $\mathcal{M} < \mathcal{M}'$, if: (i) $\Delta^{\mathcal{I}} = \Delta^{\mathcal{I}'}$; (ii) $C^{\mathcal{I}} = C^{\mathcal{I}'}$ for all concepts C; (iii) for all $x \in \Delta^{\mathcal{I}}$, it holds that $k_{\mathcal{M}}(x) \leq k_{\mathcal{M}'}(x)$ whereas there exists $y \in \Delta^{\mathcal{I}}$ such that $k_{\mathcal{M}}(y) < k_{\mathcal{M}'}(y)$. Given a KB $\mathcal{K}$, we say that $\mathcal{M}$ is a minimal model of $\mathcal{K}$ with respect to $<$ if it is a model satisfying $\mathcal{K}$ and there is no $\mathcal{M}'$ model satisfying $\mathcal{K}$ such that $\mathcal{M}' < \mathcal{M}$.

Exactly as in [15], we restrict our attention to *canonical* models. The intuition is that a canonical model contains all the individuals that enjoy properties that are consistent with the knowledge base. This is needed when reasoning about the rank of the concepts: it is important to have them all represented. We consider all the sets of concepts $\{C_1, C_2, \ldots, C_n\}$ that are consistent with $\mathcal{K}$:

DEFINITION 13
Given a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, a model $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ satisfying $\mathcal{K}$ is canonical if, for each set of concepts $\{C_1, C_2, \ldots, C_n\}$ consistent with $\mathcal{K}$, we have $(C_1 \sqcap C_2 \sqcap \cdots \sqcap C_n)^{\mathcal{I}} \neq \emptyset$, where $C_1, C_2, \ldots, C_n$ belong to the the set of all the concepts and subconcepts occurring in $\mathcal{K}$ together with their complements.

DEFINITION 14
Given a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, a model $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ satisfying $\mathcal{K}$ is a minimal canonical model of $\mathcal{K}$ if it is minimal with respect to Definition 12 and it is canonical according to Definition 13.

DEFINITION 15
Given a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, let $\mathcal{M} = \langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ and $\mathcal{M}' = \langle \Delta^{\mathcal{I}'}, <', .^{\mathcal{I}'} \rangle$ be two canonical models of $\mathcal{K}$ that are minimal with respect to Definition 14. We say that $\mathcal{M}$ is preferred to $\mathcal{M}'$ with respect to the ABox, written $\mathcal{M} <_{\mathcal{A}} \mathcal{M}'$, if, for all individual constants a occurring in $\mathcal{A}$, it holds that $k_{\mathcal{M}}(a^{\mathcal{I}}) \leq k_{\mathcal{M}'}(a^{\mathcal{I}'})$ and there is at least one individual constant b occurring in $\mathcal{A}$ such that $k_{\mathcal{M}}(b^{\mathcal{I}}) < k_{\mathcal{M}'}(b^{\mathcal{I}'})$.

The following property holds:

THEOREM 1
Given a knowledge base $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, if it admits a model, then there exists a finite minimal canonical model of KB minimally satisfying the ABox.

PROOF. We proceed exactly as in Theorem 10 in [15]. □

A query F is minimally entailed from a KB $\mathcal{K}$, written $\mathcal{K} \models_{\mathcal{ALC}\mathbf{T}^{\mathsf{P}}} F$, if it holds in all minimal canonical models of $\mathcal{K}$.

It is worth observing that the semantics in Definition 11 strengthens the monotonic semantics underlying $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCl}}$, in the sense that here we restrict reasoning to models satisfying the restrictions on the proportions defined by probabilities, differently from models in $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCl}}$ where probabilities are not considered. We provide an example witnessing such a difference:

Let us consider another extension of the knowledge base of the Introduction about wrestlers and sumo wrestlers:

$SumoWrestler \sqsubseteq Wrestler$
$\mathbf{T}(Wrestler) \sqsubseteq^{0.6} Fit$
$\mathbf{T}(SumoWrestler) \sqsubseteq^{0.9} \neg Fit$
$\mathbf{T}(Wrestler) \sqsubseteq^{0.95} \forall hasSon.(Fit \sqcap Wrestler)$
$\mathbf{T}(Wrestler) \sqsubseteq^{0.95} \exists hasSon.(WrestlingLover)$
$\mathbf{T}(Wrestler) \sqsubseteq^{0.95} \exists hasSon.(\neg WrestlingLover)$

Concerning TBox reasoning, it can be shown that $\mathcal{K} \models_{\mathcal{ALC}\mathbf{T}^{\mathsf{P}}} \mathbf{T}(SumoWrestler \sqcap Wrestler) \sqsubseteq^{0.9} \neg Fit$. Concerning ABox reasoning, suppose to have the following facts about wrestlers and sumo wrestlers:

$Wrestler(hulk)$
$Fit(hulk)$
$\neg SumoWrestler(hulk)$

$SumoWrestler(paolo)$
$\neg Fit(paolo)$
$Wrestler(gino)$

$Wrestler(franco)$
$Fit(gino)$
$Fit(franco)$

Let us consider the following models $\mathcal{M}_1$ and $\mathcal{M}_2$:

- in $\mathcal{M}_1$, Hulk is the only typical wrestler, he is fit and he is not a sumo wrestler. Paolo is the only sumo wrestler, therefore he is a typical one and he is not fit. Gino and Franco are wrestlers, they are fit but they are neither typical wrestlers nor sumo wrestlers: in particular, we have that $hulk <_1 paolo$, $hulk <_1 franco$, $franco <_1 gino$;
- in $\mathcal{M}_2$, again Hulk is the only typical wrestler, he is fit and he is not a sumo wrestler, as well as Paolo is the only sumo wrestler, therefore he is a typical one and he is not fit. Gino and Franco here are wrestlers, they are fit but they are also sumo wrestlers, in particular atypical sumo wrestlers being fit. As in $\mathcal{M}_1$, we have that $hulk <_2 paolo$, $hulk <_2 franco$, $franco <_2 gino$, $paolo <_2 franco$.

We can observe that both $\mathcal{M}_1$ and $\mathcal{M}_2$ are models of KB in the logic $\mathcal{ALC} + \mathbf{T_R}$, whereas $\mathcal{M}_1$ is also a model of KB in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ but $\mathcal{M}_2$ is not. Indeed, in $\mathcal{M}_2$ the fact that both Gino and

Franco belong to the extension of *SumoWrestler* does not respect the probability 0.9 with respect to the inclusion $\mathbf{T}(\mathit{SumoWrestler}) \sqsubseteq^{0.9} \neg \mathit{Fit}$, that would allow one to have at most the 10% of sumo wrestlers being fit: if we had a model with 100 sumo wrestlers, models satisfying KB contain at most 10 of them being fit; in this example, where the model $\mathcal{M}_2$ contains only 3 sumo wrestlers, only models not containing sumo wrestlers being also fit are allowed. The following fact:

$$(*)\ (\neg \exists \mathit{hasSon}.\mathit{SumoWrestler})(\mathit{hulk})$$

stating that Hulk does not have a son being a sumo wrestler, does not hold in $\mathcal{M}_2$, therefore it cannot be inferred in the logic $\mathcal{ALC} + \mathbf{T_R}$, whereas it holds in $\mathcal{M}_1$, as well as in all models of KB in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, therefore we can conclude that $(*)$ can be inferred in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ from KB, i.e. $\mathcal{K} \models_{\mathcal{ALC}\mathbf{T}^{\mathsf{P}}} (\neg \exists \mathit{hasSon}.\mathit{SumoWrestler})(\mathit{hulk})$.

It can be shown that the nonmonotonic semantics of the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ corresponds to an extension of the rational closure for DLs introduced in [15]. Therefore, by proceeding as in Theorem 9 and Theorem 13 in [15], as in the case of the logic $\mathbf{T}^{\mathsf{CL}}$, we can prove that reasoning in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ is in the same complexity class of the underlying standard $\mathcal{ALC}$:

THEOREM 2
The problem of checking whether a query F is minimally entailed from a KB $\mathcal{K} \models_{\mathcal{ALC}\mathbf{T}^{\mathsf{P}}} F$ is EXPTIME-complete.

4 Allowed values of probabilities

In both the semantics $\mathbf{T}^{\mathsf{CL}}$ and $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, we only allow typicality inclusions equipped with probabilities $p > 0.5$. The reasons guiding this choice are the following:

- The very cognitive notion of typicality derives from that one of probability distribution [37], in particular typical properties attributed to entities are those characterizing the majority of instances involved. In addition, the notion of probability distribution is also intrinsically connected to the one concerning the level of uncertainty/degree of belief associated to typicality inclusions (i.e. typical knowledge is known to come with a low degree of uncertainty [19]); and
- In our effort of integrating two different semantics – DISPONTE and typicality logic in $\mathbf{T}^{\mathsf{CL}}$ and a proportional interpretation of probabilities with typicality in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ – the choice of having probabilities higher than 0.5 for typicality inclusions seems to be the only one compliant with both the formalisms. In fact, in the first case, despite the DISPONTE semantics allows one to assign also low probabilities/degrees of belief to standard inclusions, in the logic $\mathbf{T}^{\mathsf{CL}}$, for what explained above, it would be at least counter-intuitive to also allow low degrees of belief for typicality inclusions (simply because typicality inclusions with high uncertainty do not describe any typical knowledge). For example, the logic $\mathbf{T}^{\mathsf{CL}}$ does not allow an inclusion like $0.3\ :\ \mathbf{T}(\mathit{Student}) \sqsubseteq \mathit{YoungPerson}$, that could be interpreted as ‘we have a low degree of belief that typical students are young people’. Such interpretation would not have sense since, by definition, typical information is a statistically relevant one and is a carrier of trust and belief attribution. Please, note that this is not a limitation of the expressivity of the logic $\mathbf{T}^{\mathsf{CL}}$: we can in fact represent properties not holding for typical members of a category, for instance if one needs to represent that typical students are not married, we can have that $0.8\ :\ \mathbf{T}(\mathit{Student}) \sqsubseteq \neg \mathit{Married}$, rather than $0.2\ :\ \mathbf{T}(\mathit{Student}) \sqsubseteq \mathit{Married}$.
- Also, in the case of $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, on the other hand, it is not possible to represent inclusions of the form $\mathbf{T}(\mathit{Student}) \sqsubseteq^{0.3} \mathit{YoungPerson}$. This statement, in fact, would be interpreted in such logic

as 'we have many exceptions to the fact that typical students are young people', violating also in this case the probabilistic assumptions from which the notion of typicality derives (since 'being young' is a typical property associated to the class of students). As in the case of the logic $\mathbf{T}^{\mathsf{CL}}$, this does not affect the expressivity of the language, since we can represent such information by using the negation of the involved property, in the example, for instance, with an inclusion $\mathbf{T}(\mathit{Student}) \sqsubseteq^{0.7} \neg \mathit{YoungPerson}$.

Concerning the upper limit of the interval, we remind that, in the logic $\mathbf{T}^{\mathsf{CL}}$ the value 1 is allowed. Indeed, an inclusion $1 : \mathbf{T}(C) \sqsubseteq D$ can be used in order to capture the fact that there is no uncertainty about the fact that, normally, Cs are also Ds. The semantics of the logic $\mathbf{T}^{\mathsf{CL}}$ will consider scenarios either including or not including such properties, but obviously scenarios not containing them will have a probability of 0. As already mentioned, as a difference with the logic $\mathbf{T}^{\mathsf{CL}}$, the upper extreme 1 is not allowed in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ in order to avoid the collapse of Proposition 1. Indeed, in this latter semantics, a typicality inclusion $\mathbf{T}(C) \sqsubseteq^{1} D$ would correspond to a classical, 'rigid' inclusion $C \sqsubseteq D$, where no uncertain information—and therefore no typicality—is involved.

It is worth noticing that this is one of the main differences between the two proposed semantics: on the one hand, in the logic $\mathbf{T}^{\mathsf{CL}}$, the independence between axioms—coming from the DISPONTE semantics—allows one to consider scenarios with different epistemic interpretations on typicality inclusions (also having a degree 1). On the other hand, in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, the independence among inclusions is no longer imposed, therefore we don't have different scenarios, but only one knowledge base containing all inclusions. As a consequence, we need to avoid the above mentioned collapse by excluding inclusions with probability 1.

5 A case study on concept combination with probabilities

In this section we exploit the DL $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ in order to tackle a well-known problem in the field of common-sense reasoning, namely the problem of typicality-based concept combination. This generative phenomenon highlights some crucial aspects of the knowledge processing capabilities in human cognition and concerns high-level capacities associated to creative thinking and problem solving. Still, it represents an open challenge in the field of artificial intelligence where other formalisms, such as fuzzy logics, have failed to model the phenomenon [17, 31]. As mentioned in the Introduction, dealing with this problem requires, from an AI perspective, the harmonization of two conflicting requirements that are hardly accommodated in symbolic systems (including formal ontologies): the need of a syntactic and semantic compositionality (typical of logical systems) and that one concerning the exhibition of typicality effects [11]. A formal solution for the mentioned problem has been proposed, for the first time in $\mathbf{T}^{\mathsf{CL}}$ [24]. In the following, however, we show that also the semantics based on nonepistemic interpretation of the probability values assumed in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ can be another formal framework able to account for this phenomenon.

5.1 Concept combination with the novel $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$

As in the case of $\mathbf{T}^{\mathsf{CL}}$, we consider the distinction between rigid and prototypical properties, the latter ones considered the interpretation of probabilities as proportion provided in Section 3. Differently from $\mathbf{T}^{\mathsf{CL}}$, as we will see, $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ provides a built-in procedure to deal with the phenomenon of common-sense concept combination (i.e. it does not require to adopt the HEAD-MODIFIER heuristics [30]). Here we first exploit the nonmonotonic semantics of $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ in order to check which properties of the concepts to be composed are inherited by the combined concept. We fix the average of the probabilities p_i as a threshold, then we ascribe to the combined concept all properties D_i such that p_i is above or equal to the average. As in the case of the logic $\mathbf{T}^{\mathsf{CL}}$, the overall output of this mechanism corresponds to a revised knowledge base including typical properties of the combined concept.

Before providing the formal definitions, let us exploit the previous example of the *pet fish* in the light of the new logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ presented here. Let $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a KB, where the ABox $\mathcal{A}$ is empty, the set of rigid inclusions is $\mathcal{R} = \{Fish \sqsubseteq \forall livesIn.Water\}$ and the set of typicality properties $\mathcal{T}$ is as follows:

1. $\mathbf{T}(Pet) \sqsubseteq^{0.9} \forall livesIn.(\neg Water)$
2. $\mathbf{T}(Pet) \sqsubseteq^{0.8} Affectionate$
3. $\mathbf{T}(Pet) \sqsubseteq^{0.8} Warm$
4. $\mathbf{T}(Fish) \sqsubseteq^{0.7} \neg Affectionate$
5. $\mathbf{T}(Fish) \sqsubseteq^{0.7} Greyish$
6. $\mathbf{T}(Fish) \sqsubseteq^{0.9} Scaly$
7. $\mathbf{T}(Fish) \sqsubseteq^{0.8} \neg Warm$

As for the logic $\mathbf{T}^{\mathsf{CL}}$, the semantics underlying the typicality operator $\mathbf{T}$ allows us to capture useful nonmonotonic inferences, as well as to let the combined concept inherit all rigid properties, for instance we have that $\mathbf{T}(Pet \sqcap Fish) \sqsubseteq \forall livesIn.Water$.

First of all, we compute the average of the probabilities of all inclusions involving the concepts to combine. The underlying rationale for this choice is that we only want to let the combined concept inherit not all properties of the compound concepts, but those showing a 'stronger' typicality effect. In the example, the average is $avg = 0.9 \times 0.8 \times 0.8 \times 0.7 \times 0.7 \times 0.9 \times 0.8 = 0.8$. For each property of either *Pet* or *Fish*, we check whether it can be inherited from $\mathcal{K}$ in the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$. We have:

1. $\mathcal{K} \not\models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.9} \forall livesIn.(\neg Water)$ (as mentioned, a rigid property blocks this one)
2. $\mathcal{K} \not\models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.8} Affectionate$
3. $\mathcal{K} \not\models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.8} Warm$
4. $\mathcal{K} \not\models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.7} \neg Affectionate$, furthermore the probability is below the average
5. $\mathcal{K} \models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.7} Greyish$, but here again the probability is below the average
6. $\mathcal{K} \models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.9} Scaly$
7. $\mathcal{K} \not\models \mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.8} \neg Warm$

Properties in 4 and 5 are not inherited since the probabilities are lower than the fixed threshold. It is also worth noticing that conflicting properties *Affectionate* and *Warm* are not entailed for the combined concept *Pet* $\sqcap$ *Fish*: indeed, the underlying logic of typicality $\mathcal{ALC} + \mathbf{T_R}$ allows one to skeptically conclude nothing about a conflicting property in absence of further information (it is an instance of the classic example of the *Nixon diamond* [38]). The other side of the coin is that, as mentioned at the very beginning of the section, in combining concepts by exploiting the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$, we do not have to resort to the HEAD/MODIFIER heuristics. The resulting $(Pet \sqcap Fish)$-revised knowledge base extends $\mathcal{T}$ with the inclusion $\mathbf{T}(Pet \sqcap Fish) \sqsubseteq^{0.9} Scaly$. Also, in this case, we can say that the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ is able to tackle the problem of defining a prototype of a combined concept: only some properties are inherited from the initial concepts C_1 and C_2 involved in the combination, and we could still be able to extend the knowledge base, for instance, with the information that a typical pet fish is red. By the nonmonotonicity of the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathsf{RaCl}}$ underlying this approach, also adding a typicality property contradicting the same property in (at least) one of the initial concepts would not be problematic; as an example, we could add a further inclusion $\mathbf{T}(\mathit{Pet} \sqcap \mathit{Fish}) \sqsubseteq^{0.7} \neg \mathit{Affectionate}$.

It is worth noticing that the resulting revised knowledge is different with respect to the one obtained by means of the logic $\mathbf{T}^{\mathsf{CL}}$. Here we exploit the semantics of the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ in order to restrict the attention only to models satisfying the probabilities as proportions equipping typicality inclusions, then only inclusions having a probability higher than the average contribute to the inheritance of typical properties for the combined concept. It is also worth noticing that, as the example shows, $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ proposes a more conservative solution: intuitively, the logic $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ inherits reasoning capabilities of the underlying $\mathcal{ALC} + \mathbf{T_R}$, but restricted to allowed models satisfying proportions. In this respect, conflicting properties are systematically discarded: indeed, as already mentioned, the logic does not take any decision in case of a Nixon diamond, whereas in the logic $\mathbf{T}^{\mathsf{CL}}$ this decision is taken by considering the HEAD/MODIFIER heuristics.

Formally, the mechanism for concept combination in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ is as follows:

DEFINITION 16 (Average of probabilities).
Given a KB $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ and given two concepts C_1 and C_2 occurring in $\mathcal{K}$, we define the average of probabilities equipping inclusions of C_1 and C_2. Let $\mathbf{T}(C_1) \sqsubseteq^{p_1} D_1, \mathbf{T}(C_1) \sqsubseteq^{p_2} D_2, \ldots, \mathbf{T}(C_1) \sqsubseteq^{p_m} D_m, \mathbf{T}(C_2) \sqsubseteq^{p_{m+1}} D_{m+1}, \mathbf{T}(C_2) \sqsubseteq^{p_{m+2}} D_{m+2}, \ldots, \mathbf{T}(C_2) \sqsubseteq^{p_n} D_n$ be all typicality inclusions in $\mathcal{T}$ involving concepts C_1 and C_2, we define:

$$avg = \frac{\sum\limits_{i=1,2,\ldots,n} p_i}{n}$$

DEFINITION 17 (Candidate properties).
Given a KB $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ and given two concepts C_1 and C_2 occurring in $\mathcal{K}$, let *avg* be as in Definition 16. We define the set $\mathcal{S}_C$ of candidate properties for the combined concept $C_1 \sqcap C_2$ as

$$\mathcal{S}_C = \{D \mid \mathcal{K} \models \mathbf{T}(C_1 \sqcap C_2) \sqsubseteq^p D \text{ and } p \geq avg, \text{for all } \mathbf{T}(C) \sqsubseteq^p D\},$$

where C is either C_1 or C_2.

Last, we define the set of typicality inclusions of the form $\mathbf{T}(C_1 \sqcap C_2) \sqsubseteq^p D_i$, for D_i belonging to the set of candidate properties:

DEFINITION 18 (C-revised knowledge base).
Given a KB $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ and given two concepts C_1 and C_2 occurring in $\mathcal{K}$, let $\mathcal{S}_C$ be the set of candidate properties for the combined concept $C = C_1 \sqcap C_2$ as in Definition 17. We define the set $\mathcal{T}'$ of typicality inclusions as $\mathcal{T}' = \{\mathbf{T}(C_1 \sqcap C_2) \sqsubseteq^p D_i \mid D_i \in \mathcal{S}_C\}$. The C-revised knowledge base is defined as $\mathcal{K}_C = \langle \mathcal{R}, \mathcal{T} \cup \mathcal{T}', \mathcal{A} \rangle$.

Since we directly exploit the semantics underlying the logic, we have that concept combination in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ remains *ExpTime* complete:

THEOREM 3
Concept combination in $\mathcal{ALC}\mathbf{T}^{\mathsf{P}}$ is EXPTIME-complete.

PROOF. In order to combine concepts C_1 and C_2 we need to:

- check whether each $\mathbf{T}(C_1) \sqsubseteq D_i$ (respectively, $\mathbf{T}(C_2) \sqsubseteq D_j$), for all typical properties of C_1 (respectively C_2) is entailed from the KB in the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCl}}$, which is ExpTime-complete [15]; and
- compute the average of probabilities as in Definition 16, in order to select only those inferred inclusions whose probability is higher than such an average: given that the number of properties and of typicality inclusions is polynomial in the size of the KB, we have that this operation has polynomial complexity.

This provides an exponential upper bound complexity. Concerning a lower bound, just consider that the logic extends the basic DL $\mathcal{ALC}$, in which reasoning is ExpTime-complete [2]. Therefore, we conclude that computing concept combination in the logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ is ExpTime-complete. □

6 Related and future works

In this work we have introduced two extensions of the DL of typicality, which is able to represent and reason about prototypical properties and defeasible inheritance, by means of *probabilities*. On the one hand, we have proposed the logic $\mathbf{T}^{\mathrm{CL}}$, where typicality inclusions are equipped by probabilities representing a *degree of belief* in such inclusions. On the other hand, we have introduced the logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$, where probabilities are used to capture the *proportion* of exceptions to typicality properties. We have shown that the proposed DLs seem to be promising from both a computational and a practical point of view: indeed, we have proved that reasoning in the two proposed extensions remains ExpTime-complete as in the underlying standard DL $\mathcal{ALC}$; furthermore, we have shown that both DLs are suitable for accounting for the phenomenon of typicality-based concept combination, which turns out to be hardly accommodated in symbolic systems.

Several approaches have been introduced in the recent literature in order to deal with reasoning under probabilistic uncertainty in DL. In this section we will briefly recall the approaches that are close to the logics $\mathbf{T}^{\mathrm{CL}}$ and $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ introduced in this work, by emphasizing the main differences between such proposals and ours. In [33, 34], it is introduced a nonmonotonic extension of the logic of typicality $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}$ by inclusions of the form $\mathbf{T}(C) \sqsubseteq_p D$ with a probability p. This work inspired the logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ that we propose here, but there is a very significant difference between the two proposals: in the works [33, 34], probabilities are essentially used in order to define a syntactic completion of an ABox containing only *some* typicality assertions; a notion of entailment restricted to those extensions whose probabilities belong to a given and fixed range is then defined. However, such ABox extension does not have a semantical counterpart, in other words the semantics do not consider probabilities, whereas in the logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ models satisfying a knowledge base have to fulfill all the constraints defined by probabilities as proportions. In [29] the author introduces two probabilistic extensions of expressive DLs $\mathcal{SHIF}(\mathbf{D})$ and $\mathcal{SHOIN}(\mathbf{D})$. These extensions are semantically based on the notion of probabilistic lexicographic entailment [20] and allow one 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. In this logic, an expression of the form $(\mathit{SocialNetworkUser} \mid \mathit{Student})[0.8, 1]$ represents that 'typically, a randomly chosen student makes use of social networks with a probability of at least 80%', whereas default knowledge can be expressed as $(\mathit{Young} \mid \mathit{Student})[1, 1]$ whose meaning is that prototypical students are young people (but we have no information about the probability of having or not exceptions). As the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCl}}$, the lexicographic entailment defined in [29] inherits interesting and useful nonmonotonic properties from lexicographic entailment in [20], such as specificity, rational monotonicity and some forms of irrelevance. The logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCL}}$ 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 inheritance blocking. These extensions are more related to the logic $\mathcal{ALC} + \mathbf{T}_{\mathbf{R}}^{\mathrm{RaCL}}$, rather than to the logics $\mathbf{T}^{\mathrm{CL}}$ and $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ defined here, in the sense that it could be of interest to evaluate the alternative of having the system of [29] as the one for underlying $\mathbf{T}^{\mathrm{CL}}$ and $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ in order to reason about defeasible inheritance in DLs.

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 [1, 3, 5–7, 9, 13], ranging from Reiter's defaults to minimal knowledge and negation as failure (see [5, 12] for details). To the best of our knowledge, none of them consider probability of exceptions in concept inclusions. Probabilistic extensions of DLs, allowing one to label inclusions (and facts) with probabilities, have been introduced in the DISPONTE semantics [35, 36] inspiring the logic $\mathbf{T}^{\mathrm{CL}}$. The main difference between the logic $\mathbf{T}^{\mathrm{CL}}$ and probabilistic DLs is that, in $\mathbf{T}^{\mathrm{CL}}$, probabilities are restricted to typicality inclusions and, therefore, only values higher than 0.5 are allowed. On the contrary, in DISPONTE probabilities can be associated to concept inclusions as well as to ABox facts. In [32] a nonmonotonic procedure for reasoning about *surprising* scenarios in DLs has been proposed. In this approach, the 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*. The main difference with the logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$ is that, in this proposal, the degree of expectedness is not intended as a proportion, but it is essentially a rank used in order to define a notion of a syntactical extension of the ABox by means of typicality assertions about individuals. Degrees of expectedness are then 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.

In future research we aim at extending our approaches to more expressive DLs, such as those underlying the standard OWL language. Starting from the work of [14], applying the logic with the typicality operator and the rational closure to $\mathcal{SHIQ}$, we intend to study whether and how our extensions, and in particular the logic $\mathcal{ALC}\mathbf{T}^{\mathrm{P}}$, could provide an alternative solution to the problem of the 'all or nothing' behavior of rational closure with respect to property inheritance.

Concerning the application to the combination of prototypes of concepts, we envision for the novel framework proposed different areas of application where the problem of concept combination is crucial. Such areas range from computational creativity and generative AI, to intelligent recommendations of novel categories to the new area of autonomic computing concerning the problem of the automatic generation of novel knowledge in a cognitive artificial agent, starting from an initial commonsense knowledge base [8, 25–28]. Finally, as a short-term future work, we aim at investigating the use of both formalisms for the generation of metaphors obtained via common-sense conceptual combination.

References

[1] F. Baader and B Hollunder. Priorities on defaults with prerequisites, and their application in treating specificity in terminological default logic. *Journal of Automated Reasoning (JAR)* **15**, 41–68, 1995. https://doi.org/10.1007/BF00881830.

[2] F. Baader, D. Calvanese and D. L McGuinness (eds). *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003.

[3] P. A. Bonatti, C. Lutz and F. Wolter. The complexity of circumscription in dls. *Journal of Artificial Intelligence Research (JAIR)* **35**, 717–73, 2009. https://doi.org/10.1613/jair.2763.

[4] P. A. Bonatti, M. Faella and L. Sauro. Defeasible inclusions in low-complexity dls. *Journal of Artificial Intelligence Research (JAIR)* **42**, 719–64, 2011.

[5] P. A. Bonatti, M. Faella and I. Petrova. A new semantics for overriding in description logics. *Artificial Intelligence* **222**, 1–48, 2015. https://doi.org/10.1016/j.artint.2014.12.010.

[6] G. Casini and U. Straccia. Rational closure for defeasible description logics. In *Proc. of JELIA, LNAI*, T. Janhunen and I. Niemelä, eds, vol. 6341, pp. 77–90. Springer, 2010. https://doi.org/10.1007/978-3-642-15675-5_9

[7] G. Casini and U. Straccia. Defeasible inheritance-based description logics. *Journal of Artificial Intelligence Research (JAIR)* **48**, 415–73, 2013. https://doi.org/10.1613/jair.4062.

[8] E. Chiodino, D. Di Luccio, A. Lieto, A. Messina, G. L. Pozzato and D. Rubinetti. A knowledge-based system for the dynamic generation and classification of novel contents in multimedia broadcasting. In *Proceedings of ECAI* 2020, De Giacomo *et al.*, eds, Volume 325 of Frontiers in Artificial Intelligence and Applications, pp. 680–687. IOS Press, 2020.

[9] F. M. Donini, D. Nardi and R. Rosati. Description logics of minimal knowledge and negation as failure. *ACM Transactions on Computational Logics (ToCL)* **3**, 177–225, 2002. https://doi.org/10.1145/505372.505373.

[10] M. Frixione and A. Lieto. Representing concepts in formal ontologies. Compositionality vs. typicality effects. *Logic and Logical Philosophy* **21**, 391–414, 2012.

[11] M. Frixione and A Lieto. Towards an extended model of conceptual representations in formal ontologies: a typicality-based proposal. *Journal of Universal Computer Science* **20**, 257–76, 2014.

[12] L. Giordano, V. Gliozzi, N. Olivetti and G. L. Pozzato. ALC+T: a preferential extension of description logics. *Fundamenta Informaticae* **96**, 341–72, 2009. https://doi.org/10.3233/FI-2009-182.

[13] L. Giordano, V. Gliozzi, N. Olivetti and G. L. Pozzato. A Nonmonotonic description logic for reasoning about typicality. *Artificial Intelligence* **195**, 165–202, 2013. https://doi.org/10.1016/j.artint.2012.10.004.

[14] L. Giordano, V. Gliozzi, N. Olivetti and G. L. Pozzato. 2014. Rational closure in $\mathcal{SHIQ}$. In *Proc. of Description Logics, CEUR Workshop Proceedings*, M. Bienvenu *et al.*, eds, Vol. 1193, pp. 543–55. CEUR-WS.org.

[15] L. Giordano, V. Gliozzi and N. Olivetti. Semantic characterization of rational closure: from propositional logic to description logics. *Artificial Intelligence* **226**, 1–33, 2015. https://doi.org/10.1016/j.artint.2015.05.001.

[16] L. Giordano, V. Gliozzi, A. Lieto, N. Olivetti and G. L. Pozzato. Reasoning about typicality and probabilities in preferential description logics. In *Applications and Practices in Ontology Design, Extraction, and Reasoning*, Studies on the Semantic Web 49, 145-161, IOS Press 2020, ISBN 978-1-64368-142-9, 2020.

[17] J. A Hampton. Conceptual combinations and fuzzy logic. *Concepts and Fuzzy Logic* **209**, 2011. https://doi.org/10.7551/mitpress/8842.003.0010.

[18] S. Kraus, D. Lehmann and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* **44**, 167–207, 1990. https://doi.org/10.1016/0004-3702(90)90101-5.

[19] J. Lawry and Y. Tang. Uncertainty modelling for vague concepts: a prototype theory approach. *Artificial Intelligence* **173**, 1539–58, 2009. https://doi.org/10.1016/j.artint.2009.07.006.

[20] D. J. Lehmann. Another perspective on default reasoning. *Annals of Mathematics and Artificial Intelligence* **15**, 61–82, 1995. https://doi.org/10.1007/BF01535841.

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

[22] A Lieto. *Cognitive Design for Artificial Minds*. Routledge (2021), https://doi.org/10.4324/9781315460536

[23] A. Lieto and G. L. Pozzato. A description logic of typicality for conceptual combination. In *Proceedings of the International Symposium on Methodologies for Intelligent Systems (ISMIS)*, M. Ceci *et al.*, eds, LNAI 11177, pp. 189–199. Springer, 2018, https://doi.org/10.1007/978-3-030-01851-1_19

[24] A. Lieto and G. L. Pozzato. A description logic framework for commonsense conceptual combination integrating typicality, probabilities and cognitive heuristics. *Journal of Experimental & Theoretical Artificial Intelligence* **32**, 769–804, 2020. https://doi.org/10.1080/0952813X.2019.1672799.

[25] A. Lieto, F. Perrone and G. L. Pozzato. Beyond subgoaling: a dynamic knowledge generation framework for creative problem solving in cognitive architectures. *Cognitive Systems Research* **58**, 305–16, 2019. https://doi.org/10.1016/j.cogsys.2019.08.005.

[26] A. Lieto, G. L. Pozzato, S. Zoia, V. Patti and R. Damiano. A commonsense reasoning framework for explanatory emotion attribution, generation and re-classification. *Knowledge-Based Systems* **227**, 107166, 2021. https://doi.org/10.1016/j.knosys.2021.107166.

[27] A. Lieto, G. L. Pozzato, M. Striani, S. Zoia and R. Damiano. Degari 2.0: a diversity-seeking, explainable, and affective art recommender for social inclusion. *Cognitive Systems Research* **77**, 1–17, 2023. https://doi.org/10.1016/j.cogsys.2022.10.001.

[28] A. Lieto, M. Striani, C. Gena, E. Dolza, A. M. Marras, G. L. Pozzato and R. Damiano. A sensemaking system for grouping and suggesting stories from multiple affective viewpoints in museums. *Human–Computer Interaction* **39**, 109–43, 2024.

[29] T. Lukasiewicz. Expressive probabilistic description logics. *Artificial Intelligence* **172**, 852–83, 2008. https://doi.org/10.1016/j.artint.2007.10.017.

[30] E. Machery and L. G. Lederer. *Simple Heuristics for Concept Combination,* 2012.

[31] D. N. Osherson and E. E. Smith. On the adequacy of prototype theory as a theory of concepts. *Cognition* **9**, 35–58, 1981. https://doi.org/10.1016/0010-0277(81)90013-5.

[32] G. L. Pozzato. Reasoning about plausible scenarios in Description Logics of typicality. *Intelligenza Artificiale* **11**, 25–45, 2017. https://doi.org/10.3233/IA-170105.

[33] G. L. Pozzato. Reasoning in description logics with typicalities and probabilities of exceptions. In *Proc. of ECSQARU, LNAI*, A. Antonucci, L. Cholvy and O. Papini eds, vol. 10369, pp. 409–420. Springer, 2017, https://doi.org/10.1007/978-3-319-61581-3_37

[34] G. Pozzato. Typicalities and probabilities of exceptions in nonmotonic description logics. *International Journal of Approximate Reasoning* **107**, 81–100, 2019. https://doi.org/10.1016/j.ijar.2019.02.003.

[35] F. Riguzzi, E. Bellodi and E. Lamma. Probabilistic description logics under the distribution semantics. *Semantic Web* **6**, 477–501, 2015. https://doi.org/10.3233/SW-140154.

[36] F. Riguzzi, E. Bellodi, E. Lamma and R. Zese. Reasoning with probabilistic ontologies. In *Proc. of IJCAI*, Q. Yang and M. Wooldridge, eds, pp. 4310–6. AAAI Press, 2015.

[37] E. Rosch. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* **104**, 192–233, 1975. https://doi.org/10.1037/0096-3445.104.3.192.

[38] J. F. Sowa. Semantic networks. *Encyclopedia of Artificial Intelligence*, Vol. 2, 1493–511, 1992.

[39] U. Straccia. Default inheritance reasoning in hybrid kl-one-style logics. In *Proc. of IJCAI'93*, pp. 676–81. Morgan Kaufmann, 1993.

Received 10 June 2024