Creative Concept Generation by Combining Description Logic of Typicality, Probabilities and Cognitive Heuristics

Antonio Lieto* and Gian Luca Pozzato**

*Dipartimento di Informatica, Università di Torino, and ICAR-CNR (Palermo), Italy
*Dipartimento di Informatica, Università di Torino, Italy
{antonio.lieto,gianluca.pozzato}@unito.it

Abstract. We propose a nonmonotonic Description Logic of typicality as a tool for the generation and the exploration of novel creative concepts, that could be useful in many applicative scenarios, ranging from video games to the creation of new movie characters. In particular, our logic is able to deal with the phenomenon of prototypical concept combination, which has been shown to be problematic to model for other formalisms like fuzzy logic. The proposed logic relies on the logic of typicality $\mathcal{ALC} + \mathbf{T_R}$, whose semantics is based on a notion of rational closure, as well as on the distributed semantics of probabilistic Description Logics, and takes into account the insights coming from the heuristics used by humans for concept composition. Besides providing framework able to account for typicality-based concept combination, we also outline that reasoning in the proposed Description Logic is EXPTIME-complete as for the underlying $\mathcal{ALC}$.

1 Introduction

Inventing novel concepts by combining the typical knowledge of pre-existing ones is an important human creative ability. Dealing with this problem requires, from an AI perspective, the harmonization of two conflicting requirements that are hardly accommodated in symbolic systems: the need of a syntactic compositionality (typical of logical systems) and one concerning the exhibition of typicality effects [1]. According to a well-known argument [2], in fact, prototypical concepts are not compositional. The argument runs as follows: consider 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).

In this work we provide a framework able to account for this type of human-like concept combination in the scenario of typical concept invention. We exploit a nonmonotonic Description Logic (from now on DL) of typicality called $\mathbf{T}^{\mathsf{CL}}$ (typical compositional logic), that has been already shown able to capture well established examples in the literature of cognitive science concerning concept combination [3]. This logic combines two main ingredients. The first one relies on the DL of typicality $\mathcal{ALC} + \mathbf{T_R}$ introduced in [4]. In this logic, "typical" properties can be directly specified by means of a "typicality" operator $\mathbf{T}$ enriching the underlying DL, and a TBox can contain inclusions of the form $\mathbf{T}(C) \sqsubseteq D$ to represent that "typical Cs are also Ds". As a difference with standard DLs, in the logic $\mathcal{ALC} + \mathbf{T_R}$ one can consistently express exceptions and reason about defeasible inheritance as well. For instance, a knowledge base can consistently express that "normally, athletes are fit", whereas "sumo wrestlers usually are not fit" by $\mathbf{T}(\mathit{Athlete}) \sqsubseteq \mathit{Fit}$ and $\mathbf{T}(\mathit{SumoWrestler}) \sqsubseteq \neg \mathit{Fit}$, given that $\mathit{SumoWreslter} \sqsubseteq \mathit{Athlete}$. The semantics of the $\mathbf{T}$ operator is characterized by the properties of *rational logic* [5], recognized as the core properties of nonmonotonic reasoning. $\mathcal{ALC} + \mathbf{T_R}$ is characterized by a minimal model semantics corresponding to an extension to DLs of a notion of *rational closure* as defined in [5] for propositional logic: the idea is to adopt a preference relation among $\mathcal{ALC} + \mathbf{T_R}$ models, where intuitively a model is preferred to another one if it contains less exceptional elements, as well as a notion of *minimal entailment* restricted to models that are minimal with respect to such preference relation. As a consequence, $\mathbf{T}$ inherits well-established properties like *specificity* and *irrelevance*: in the example, the logic $\mathcal{ALC} + \mathbf{T_R}$ allows us to infer $\mathbf{T}(\mathit{Athlete} \sqcap \mathit{Bald}) \sqsubseteq \mathit{Fit}$ (being bald is irrelevant with respect to being fit) and, if one knows that Hiroyuki is a typical sumo wrestler, to infer that he is not fit, giving preference to the most specific information.

As a second ingredient, we consider a distributed semantics similar to the one of probabilistic DLs known as DISPONTE [6], allowing to label axioms with degrees representing probabilities, but restricted to typicality inclusions. The basic idea is to label inclusions $\mathbf{T}(C) \sqsubseteq D$ with a real number between 0.5 and 1, representing its probability[1]: such a number represents the probability of finding elements of C being also D, then we impose that it is at least 50% since we are only considering typicality properties. We assume that the axioms are independent from each other. The resulting knowledge base defines a probability distribution over *scenarios*: roughly speaking, a scenario is obtained by choosing, for each typicality inclusion, whether it is considered as true of false. In a slight extension of the above example, we could have the need of representing that both typicality inclusions about athletes and sumo wrestlers have a probability of 80%, whereas we also believe that athletes are usually young with a higher probability of 95%, with the following KB: (1) $\mathit{SumoWrestler} \sqsubseteq \mathit{Athlete}$; (2) $0.8\ ::\ \mathbf{T}(\mathit{Athlete}) \sqsubseteq \mathit{Fit}$; (3) $0.8\ ::\ \mathbf{T}(\mathit{SumoWrestler}) \sqsubseteq \neg \mathit{Fit}$; (4) $0.95\ ::\ \mathbf{T}(\mathit{Athlete}) \sqsubseteq \mathit{YoungPerson}$. We consider eight different scenarios, representing all possible combinations of typicality inclusion: as an example, $\{((2), 1), ((3), 0), ((4), 1)\}$ represents the scenario in which (2) and (4) hold, whereas (3) does not. We equip each scenario with a probability depending on those of the involved inclusions, then we restrict reasoning to scenarios whose probabilities belong to a given and fixed range.

As an additional element of the proposed formalization we employ a method inspired by cognitive semantics [7] for the identification of a dominance effect between the concepts to be combined: for every combination, we distinguish a HEAD, representing the stronger element of the combination, and a MODIFIER. The basic idea is: given a KB and two concepts C_H (HEAD) and C_M (MODIFIER) occurring in it, we consider only *some* scenarios in order to define a revised knowledge base, enriched

[1] Here, we focus on the proposal of the formalism itself, therefore the machinery for obtaining probabilities from an application domain will not be discussed.

by typical properties of the combined concept $C \sqsubseteq C_H \sqcap C_M$ (the heuristics for the scenario selections are detailed in Section 2, Definition 7).

In this work, we show that the proposed logic $\mathbf{T}^{\mathsf{CL}}$ is able to tackle the problem of composing prototypical concepts (by creating a novel, plausible, conceptual prototype) and we exploit the proposed formalism as a tool for the generation of novel creative concepts, that could be useful in many applicative scenarios. We also outline that the reasoning complexity of $\mathbf{T}^{\mathsf{CL}}$ is EXPTIME-complete, as in the standard $\mathcal{ALC}$ logic, witnessing that the proposed approach is essentially inexpensive.

2 A Logic for Concept Combination

The nonmonotonic Description Logic $\mathbf{T}^{\mathsf{CL}}$ combines the semantics based on the rational closure of $\mathcal{ALC} + \mathbf{T_R}$ [4] with the probabilistic DISPONTE semantics [6]. By taking inspiration from [8], we consider two types of properties associated to a given concept: rigid and typical. Rigid properties are those that hold under any circumstance, e.g. $C \sqsubseteq D$ (all Cs are Ds). Typical properties are represented by inclusions equipped by a probability. Additionally, we employ a cognitive heuristic for the identification of a dominance effect between the concepts to be combined, distinguishing between HEAD and MODIFIER[2].

The language of $\mathbf{T}^{\mathsf{CL}}$ 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, whose meaning is that "normally, Cs are also D with probability p".

Definition 1 (Language of $\mathbf{T}^{\mathsf{CL}}$). *We consider an alphabet* $\mathtt{C}$ *of concept names,* $\mathtt{R}$ *of role names, and* $\mathtt{O}$ *of individual constants. 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 of the inclusion;*
- $\mathcal{A}$ *is the ABox, i.e. a finite set of formulas of the form either* $C(a)$ *or* $R(a, b)$*, where* $a, b \in \mathtt{O}$*.*

It is worth noticing that we avoid typicality inclusions with degree 1. Indeed, an inclusion $1 :: \mathbf{T}(C) \sqsubseteq D$ would mean that it is a certain property, that we represent with $C \sqsubseteq D \in \mathcal{R}$. Also, observe that we only allow typicality inclusions equipped with probabilities $p > 0.5$. Indeed, the very notion of typicality derives from the one of probability distribution, in particular typical properties attributed to entities are those characterizing the majority of instances involved. Moreover, in our effort of integrating two different semantics – DISPONTE and typicality logic – the choice of having probabilities higher than 0.5 for typicality inclusions seems to be the only compliant with both formalisms. In fact, despite the DISPONTE semantics allows to assign also low

[2] Here we assume that some methods for the automatic assignment of the HEAD/MODIFER pairs are/may be available and focus on the discussion of the reasoning part.

probabilities/degrees of belief to standard inclusions, in the logic $\mathbf{T}^{\textsf{CL}}$ it would be misleading to also allow low probabilities for typicality inclusions. Please, note that this is not a limitation of the expressivity of the logic $\mathbf{T}^{\textsf{CL}}$: we can in fact represent properties not holding for typical members of a category, for instance if one need to represent that typical students are not married, we can have that $0.8 \ :: \ \mathbf{T}(Student) \sqsubseteq \neg Married$.

Following from the DISPONTE semantics, each axiom is independent from each others. This this allows us to deal with conflicting typical properties equipped with different probabilities.

A model $\mathcal{M}$ of $\mathbf{T}^{\textsf{CL}}$ extends standard $\mathcal{ALC}$ models by a preference relation among domain elements as in the logic of typicality [4]. In this respect, $x < y$ means that x is "more normal" than y, and that the typical members of a concept C are the minimal elements of C with respect to this relation. An element $x \in \Delta^{\mathcal{I}}$ is a *typical instance* of some concept C if $x \in C^{\mathcal{I}}$ and there is no element in $C^{\mathcal{I}}$ *more typical* than x.

Definition 2 (Model). *A model $\mathcal{M}$ is any structure $\langle \Delta^{\mathcal{I}}, <, .^{\mathcal{I}} \rangle$ where: (i) $\Delta^{\mathcal{I}}$ is a non empty set of items called the domain; (ii) $<$ 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 $\Delta^{\mathcal{I}}$; (iii) $.^{\mathcal{I}}$ is the extension function that maps each concept C to $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$, and each role R to $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. For concepts of $\mathcal{ALC}$, $C^{\mathcal{I}}$ is defined as usual. For the $\mathbf{T}$ operator, we have $(\mathbf{T}(C))^{\mathcal{I}} = Min_<(C^{\mathcal{I}})$, where $Min_<(C^{\mathcal{I}}) = \{x \in C^{\mathcal{I}} \mid \nexists y \in C^{\mathcal{I}}$ 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 [4]: the rank of x is the length of the longest chain $x_0 < \ldots < x$ 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 $k_{\mathcal{M}}(x) < k_{\mathcal{M}}(y)$.

Definition 3 (Model satisfying a KB). *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: (i) $\mathcal{M}$ satisfies $\mathcal{R}$ if, for all $C \sqsubseteq D \in \mathcal{R}$, we have $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$; (ii) $\mathcal{M}$ satisfies $\mathcal{T}$ if, for all $q \ :: \ \mathbf{T}(C) \sqsubseteq D \in \mathcal{T}$, we have that $\mathbf{T}(C)^{\mathcal{I}} \subseteq D^{\mathcal{I}}$, i.e. $Min_<(C^{\mathcal{I}}) \subseteq D^{\mathcal{I}}$; (iii) $\mathcal{M}$ satisfies $\mathcal{A}$ if, for each assertion $F \in \mathcal{A}$, if $F = C(a)$ then $a^{\mathcal{I}} \in C^{\mathcal{I}}$, otherwise if $F = R(a, b)$ then $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$.*

Even if the typicality operator $\mathbf{T}$ itself is nonmonotonic (i.e. $\mathbf{T}(C) \sqsubseteq E$ does not imply $\mathbf{T}(C \sqcap D) \sqsubseteq E$), what is inferred from a KB can still be inferred from any KB' with KB $\subseteq$ KB', i.e. the resulting logic is monotonic. In order to perform useful nonmonotonic inferences, in [4] 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 corresponds to a notion of *rational closure* on top of $\mathcal{ALC} + \mathbf{T_R}$. Such a notion is a natural extension of the rational closure construction provided in [5] for the propositional logic. This nonmonotonic semantics 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. This is needed when reasoning about the rank of the concepts: it is important to have them all represented. A query F is minimally entailed from a KB if it holds in all minimal canonical models of KB. In [4] it is shown that query entailment in the nonmonotonic $\mathcal{ALC} + \mathbf{T_R}$ is in EXPTIME.

Definition 4 (Entailment). *Let $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a KB and let F be either $C \sqsubseteq D$ (C could be $\mathbf{T}(C')$) or $C(a)$ or $R(a, b)$. We say that F follows from $\mathcal{K}$ if, for all minimal $\mathcal{M}$ satisfying $\mathcal{K}$, then also $\mathcal{M}$ satisfies F.*

Let us now define the notion of *scenario* of the composition of concepts. 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 5 (Atomic choice). *Given $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{E_1 = q_1 \ :: \mathbf{T}(C_1) \sqsubseteq D_1, \ldots, E_n = q_n \ :: \mathbf{T}(C_n) \sqsubseteq D_n\}$ we define (E_i, k_i) an* atomic choice *for some $i \in \{1, 2, \ldots, n\}$, where $k_i \in \{0, 1\}$.*

Definition 6 (Selection). *Given $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{E_1 = q_1 \ :: \mathbf{T}(C_1) \sqsubseteq D_1, \ldots, E_n = q_n \ :: \mathbf{T}(C_n) \sqsubseteq D_n\}$ 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 $i = 1, 2, \ldots, n$. 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 7 (Scenario). *Given $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T} = \{E_1 = q_1 \ :: \mathbf{T}(C_1) \sqsubseteq D_1, \ldots, E_n = q_n \ :: \mathbf{T}(C_n) \sqsubseteq D_n\}$ 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 *when it admits a model in the logic $\mathbf{T}^{\mathsf{CL}}$.*

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

Given a KB $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ and given two concepts C_H and C_M occurring in $\mathcal{K}$, our logic allows to define the compound concept C as the combination of the HEAD C_H and the MODIFIER C_M, where $C \sqsubseteq C_H \sqcap C_M$ and the typical properties of the form $\mathbf{T}(C) \sqsubseteq D$ to ascribe to the concept C are obtained in the set of scenarios that:

1. are consistent;
2. are not trivial, i.e. those with the highest probability, in the sense that the scenarios considering *all* properties that can be consistently ascribed to C are discarded;
3. are those giving preference to the typical properties of the HEAD C_H (with respect to those of the MODIFIER C_M) with the highest probability. Notice that, in case of conflicting properties like D and $\neg D$, given two scenarios w_1 and w_2, both

belonging to the set of consistent scenarios with the highest probability and such that an inclusion $p_1 :: \mathbf{T}(C_H) \sqsubseteq D$ belongs to w_1 whereas $p_2 :: \mathbf{T}(C_M) \sqsubseteq \neg D$ belongs to w_2, the scenario w_2 is discarded in favor of w_1.

In order to select the wanted scenarios we apply points 1, 2, and 3 above to blocks of scenarios with the same probability, in decreasing order starting from the highest one. More in detail, we first discard all the inconsistent scenarios, then we consider the remaining (consistent) ones in decreasing order by their probabilities. We then consider the blocks of scenarios with the same probability, and we proceed as follows:

- we discard those considered as *trivial*, consistently inheriting all (or most of) the properties from the starting concepts to be combined;
- among the remaining ones, we discard those inheriting properties from the MODIFIER in conflict with properties inherited from the HEAD in another scenario of the same block (i.e., with the same probability);
- if the set of scenarios of the current block is empty, i.e. all the scenarios have been discarder either because trivial or because preferring the MODIFIER, we repeat the procedure by considering the block of scenarios, all having the immediately lower probability;
- the set of remaining scenarios are those selected by the logic $\mathbf{T}^{\mathsf{CL}}$.

The knowledge base obtained as the result of combining concepts C_H and C_M into the compound concept C is called C*-revised* knowledge base:

$$\mathcal{K}_C = \langle \mathcal{R}, \mathcal{T} \cup \{p : \mathbf{T}(C) \sqsubseteq D\}, \mathcal{A} \rangle,$$

for all D such that $\mathbf{T}(C) \sqsubseteq D$ is entailed in w. The probability p is defined as follows: if D is a property inherited either from the HEAD (or from both the HEAD and the MODIFIER), then p corresponds to the probability of such inclusion of the HEAD in the initial knowledge base, i.e. $p : \mathbf{T}(C_H) \sqsubseteq D \in \mathcal{T}$; otherwise, p corresponds to the probability of such inclusion of a MODIFIER in the initial knowledge base, i.e. $p : \mathbf{T}(C_M) \sqsubseteq D \in \mathcal{T}$. Notice that, since the C*-revised* knowledge base is still in the language of the $\mathbf{T}^{\mathsf{CL}}$ logic, we can iteratively repeat the same procedure in order to combine not only atomic concepts, but also compound concepts. We leave a detailed analysis of this topic for future works.

We conclude the section by showing that reasoning in $\mathbf{T}^{\mathsf{CL}}$ remains in the same complexity class of standard $\mathcal{ALC}$. For the completeness, let n be the size of KB, then the number of typicality inclusions is $O(n)$. It is straightforward to observe that we have an exponential number of different scenarios, for each one we need to check whether the resulting KB is consistent in $\mathcal{ALC} + \mathbf{T_R}$ which is EXPTIME-complete. Hardness immediately follows from the fact that $\mathbf{T}^{\mathsf{CL}}$ extends $\mathcal{ALC} + \mathbf{T_R}$. In [3] we have shown that reasoning in $\mathbf{T}^{\mathsf{CL}}$ in the revised knowledge is EXPTIME-complete.

3 Artificial Prototypes Composition and Concept Invention

In this section we exploit the logic $\mathbf{T}^{\mathsf{CL}}$ to show both i) how it allows to automatically generate novel, plausible, prototypical concepts by composing two initial prototypes and ii) how it can be used as a generative tool in the field of computational creativity (with applications in the so called creative industry). In detail, we first show how our logic can model the generation of a quite complex concept recently introduced in the field of narratology, i.e. that one of the ANTI-HERO (a role invented by narratologists to generate new story lines), by combining the typical properties of the concepts HERO and VILLAIN. Of course, the specific domain of the example is not relevant here; our goal is showing how $\mathbf{T}^{\text{CL}}$ can model this kind of prototypical concept composition (a crucial aspect of human concept invention) that, on the other hand, has been proven to be problematic for other kinds of logics (e.g fuzzy logic, [2, 9]). We then show how the same machinery can be used as a creative support tool to generate a new type of villain for a video game or a movie.

3.1 The Anti Hero

We will take into account the concepts of HERO, ANTI-HERO and VILLAIN extracted by the common sense descriptions coming from the TvTropes repository (https://tvtropes.org). In such online repository, typical descriptions of character roles are provided. They can be useful for practitioners of the narrative field in order to design their own character according to the main assets presented in such schemas. In particular, Tropes can be seen as devices and conventions that a writer can reasonably rely on as being present in the audience members minds and expectations. Regarding the HERO, TvTropes identifies the following relevant representative features: e.g. the fact that it is characterized by his/her fights against the VILLAIN of a story, the fact that his/her actions are necessarily guided by general goals to be achieved in the interest of the collectivity, the fact that they fight against the VILLAIN in a fair way and so on. Examples of such Trope are: Superman, Flash Gordon etc.. The ANTI-HERO, on the other hand, is described as characterized by the fact of sharing most of its typical traits with the HERO (e.g. the fact that it is the protagonist of a plot fighting against the VILLAIN of the story); however, his/her moves are not guided by a general spirit of sacrifice for the collectivity but, rather, they are usually based on some personal motivations that, incidentally and/or indirectly, coincide with the needs of the collectivity. Furthermore the ANTI-HERO may also act in a not fair way in order to achieve the desired goal. A classical example of such trope is Batman, whose moves are guided by his desire of revenge. Finally the VILLAIN is represented as a classic negative role in a plot and is characterized as the main opponent of the protagonist/HERO. In addition to this classical contraposition, TvTropes also reports some physical elements characterizing such role from a visual point of view. For example: the characters of this Trope are usually physically endowed with some demoniac cues (e.g. they have the "eyes of fire"). Finally, they are guided by negative moral values. Examples of such role can be easily taken from the classical literature to the modern comics. Some representative exemplars are Cruella de Vil in Disneys filmic saga or Voldemort in Harry Potter.

Let us now exploit our logic $\mathbf{T}^{\text{CL}}$ in order to define a prototype of ANTI-HERO. First of all, we define a knowledge base describing both rigid and typical properties of concepts HERO and VILLAIN, then we rely on the logic $\mathbf{T}^{\text{CL}}$ in order to formalize an $AntiHero$-revised knowledge base.

Let $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a KB, where the ABox $\mathcal{A}$ is empty. Concerning rigid properties, let $\mathcal{R}$ be as follows:

R1 $\mathit{Hero} \sqsubseteq \exists \mathit{hasOpponent}.\mathit{Villain}$
R2 $\mathit{Villain} \sqsubseteq \forall \mathit{fightsFor}.\mathit{PersonalGoal}$
R3 $\mathit{Villain} \sqsubseteq \mathit{WithNegativeMoralValues}$
R4 $\mathit{CollectiveGoal} \sqcap \mathit{PersonalGoal} \sqsubseteq \bot$
R5 $\mathit{WithPositiveMoralValues} \sqcap \mathit{WithNegativeMoralValues} \sqsubseteq \bot$
R6 $\mathit{AngelicIconicity} \sqcap \mathit{DemoniacIconicity} \sqsubseteq \bot$

Prototypical properties of villains and heroes are described in $\mathcal{T}$ is as follows:

T1 $0.95 \ :: \ \mathbf{T}(\mathit{Hero}) \sqsubseteq \mathit{Protagonist}$
T2 $0.85 \ :: \ \mathbf{T}(\mathit{Hero}) \sqsubseteq \exists \mathit{fightsFor}.\mathit{CollectiveGoal}$
T3 $0.9 \ :: \ \mathbf{T}(\mathit{Hero}) \sqsubseteq \mathit{WithPositiveMoralValues}$
T4 $0.6 \ :: \ \mathbf{T}(\mathit{Hero}) \sqsubseteq \mathit{AngelicIconicity}$
T5 $0.75 \ :: \ \mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{DemoniacIconicity}$
T6 $0.8 \ :: \ \mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Implulsive}$
T7 $0.75 \ :: \ \mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Protagonist}$

We make use of the logic $\mathbf{T}^{\text{CL}}$ in order to build the compound concept $\mathit{AntiHero}$ as the result of the combination of concepts Hero and $\mathit{Villain}$. Differently from what the natural language seems to suggest, we consider this compound concept by assuming that the HEAD is $\mathit{Villain}$ (since the ANTI-HERO shares more typical traits with this concept than with the HERO concept).

First of all, we have that the compound concepts inherits all the rigid properties of both its components (if not contradictory), therefore in the logic $\mathbf{T}^{\text{CL}}$ we have that:

(i) $\mathit{AntiHero} \sqsubseteq \exists \mathit{hasOpponent}.\mathit{Villain}$
(ii) $\mathit{AntiHero} \sqsubseteq \forall \mathit{fightsFor}.\mathit{PersonalGoal}$
(iii) $\mathit{AntiHero} \sqsubseteq \mathit{WithNegativeMoralValues}$

For the typical properties, we consider all the $2^7 = 256$ different scenarios obtained from all possible selections about inclusion in $\mathcal{T}$. Some of them are inconsistent, namely those including either axiom T2 or axiom T3, since they would ascribe properties in contrast with inherited rigid properties of (ii) and (iii): rigid properties impose that an anti hero has negative moral values, and all his goals are personal, therefore he is an atypical hero in those respects (T2 states that typical heroes fights also for some collective goals, whereas T3 states that normally heroes have positive moral values). Also scenarios containing both axioms T4 and T5 are inconsistent, due to the fact that the concepts $\mathit{AngelicIconicity}$ and $\mathit{DemoniacIconicity}$ are disjoint (formalized by R6).

Let us consider the remaining, consistent scenarios: the one having the highest probability considers all the properties of both concepts by excluding only $\mathit{AngelicIconicity}$, that is to say the one with the lowest probability between the two properties in conflict. In $\mathbf{T}^{\text{CL}}$ this scenario is discarded since it is the most trivial one. When we consider scenarios less trivial, i.e., more surprising scenarios (we analyze scenarios in decreasing order of probability), we discard the scenario with probability 0.13%, which includes T4, associated to the MODIFIER, rather than T5, associated to the HEAD, allowing to conclude, in a counter intuitive way, that typical anti heroes have an angelic iconicity rather than a demoniac one.

Next scenarios, sharing the same probability (0.09%), are as follows:

T1 0.95 :: $\mathbf{T}(\mathit{Hero}) \sqsubseteq \mathit{Protagonist}$ T5 0.75 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{DemoniacIconicity}$ T6 0.8 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Implulsive}$	T1 0.95 :: $\mathbf{T}(\mathit{Hero}) \sqsubseteq \mathit{Protagonist}$ T6 0.8 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Implulsive}$ T7 0.75 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Protagonist}$

According to the logic $\mathbf{T}^{\mathsf{CL}}$, both are adequate and represent the outcome of the whole heuristic procedures adopted in $\mathbf{T}^{\mathsf{CL}}$. Probably, in this case, it could be more useful to opt for the solution on the left allowing to inherit a further property (i.e. $\mathit{DemoniacIconicity}$) for the generated prototypical Anti-Hero. However, we remain agnostic about the selection of the final options provided by $\mathbf{T}^{\mathsf{CL}}$. This choice can be plausibly left to human decision makers and based on their own goals.

A final element that is worth noticing in $\mathbf{T}^{\mathsf{CL}}$ is the following: in our logic, adding a new inclusion, e.g. $\mathbf{T}(\mathit{AntiHero}) \sqsubseteq \mathit{Brave}$, would not be problematic. That is to say that our formalism is able to tackle the phenomenon of prototypical *attributes emergence* for the new compound concept, widely described in the Cognitive Science literature [7]. In the next subsection we show how $\mathbf{T}^{\mathsf{CL}}$ can be used to invent novel concepts.

3.2 The Villain Chair

Let us assume to generate a novel concept obtained as the combination of concepts $\mathit{Villain}$ (as HEAD) and Chair (as MODIFIER). Let $\mathcal{K} = \langle \mathcal{R}, \mathcal{T}, \emptyset \rangle$ be as follows:

R1 $\mathit{Villain} \sqsubseteq \exists \mathit{fightsFor}.\mathit{PersonalGoal}$
R2 $\mathit{Villain} \sqsubseteq \mathit{Animate}$
R3 $\mathit{Villain} \sqsubseteq \mathit{WithNegativeMoralValues}$
R4 $\mathit{Chair} \sqsubseteq \exists \mathit{hasComponent}.\mathit{SupportingSeatComponent}$
R5 $\mathit{Chair} \sqsubseteq \exists \mathit{hasComponent}.\mathit{Seat}$
R6 $\mathit{CollectiveGoal} \sqcap \mathit{PersonalGoal} \sqsubseteq \bot$

and $\mathcal{T}$ is as follows:

T1 0.9 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{DemoniacIconicity}$
T2 0.75 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \exists \mathit{hasOpponent}.\mathit{Hero}$
T3 0.75 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Protagonist}$
T4 0.8 :: $\mathbf{T}(\mathit{Villain}) \sqsubseteq \mathit{Impulsive}$
T5 0.95 :: $\mathbf{T}(\mathit{Chair}) \sqsubseteq \neg \mathit{Animate}$
T6 0.95 :: $\mathbf{T}(\mathit{Chair}) \sqsubseteq \exists \mathit{hasComponent}.\mathit{Back}$
T7 0.65 :: $\mathbf{T}(\mathit{Chair}) \sqsubseteq \exists \mathit{madeOf}.\mathit{Wood}$
T8 0.8 :: $\mathbf{T}(\mathit{Chair}) \sqsubseteq \mathit{Comfortable}$
T9 0.7 :: $\mathbf{T}(\mathit{Chair}) \sqsubseteq \mathit{Inflammable}$

We consider the 512 scenarios, from which we discard the inconsistent ones, namely those including T5: indeed, since R2 imposes that villains are animate, in the underlying $\mathcal{ALC} + \mathbf{T_R}$ we conclude that $\mathit{Villain} \sqcap \mathit{Chair} \sqsubseteq \mathit{Animate}$, therefore all scenarios including T5, imposing that $\mathit{Villain} \sqcap \mathit{Chair} \sqsubseteq \neg \mathit{Animate}$ are inconsistent. We also discard the most obvious scenario including all the typicality inclusions of $\mathcal{R}$, having probability of 14%, and the ones containing all the inclusions related to the HEAD. The first suitable scenarios are those have probability 4.67% and contain all properties coming from the MODIFIER and three out of four properties coming from the HEAD. Such scenarios define two alternative revised knowledge bases: one containing T2 and not T3, the other one containing T3 and not T2. These scenarios are the preferred ones selected by the logic $\mathbf{T}^{\mathsf{CL}}$.

However, in this application setting, we could imagine to use our framework as a creativity support tool and thus considering alternative - more surprising - scenarios by adding additional constraints. For example, we could impose that the compound concept should inherit exactly six properties. In this case, we would get that the scenario having the highest probability (3.2%) is the one including all the properties of the HEAD, namely T1, T2, T3 and T4, and two out of four properties of the MODIFIER, namely T6 and T8. Due to its triviality, this scenario is discarded, in favor of the following more creative scenarios, with probability 2.51%, obtained by excluding T7 of the MODIFIER and one out of four properties of the HEAD:

T1 $0.9 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{DemoniacIconicity}$
T2 $0.75 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \exists \mathit{hasOpponent}.\mathit{Hero}$
T4 $0.8 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Impulsive}$
T6 $0.95 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \exists \mathit{hasComponent}.\mathit{Back}$
T8 $0.8 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Comfortable}$
T9 $0.7 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Inflammable}$

T1 $0.9 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{DemoniacIconicity}$
T3 $0.75 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Protagonist}$
T4 $0.8 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Impulsive}$
T6 $0.95 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \exists \mathit{hasComponent}.\mathit{Back}$
T8 $0.8 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Comfortable}$
T9 $0.7 \ :: \ \mathbf{T}(\mathit{Villain} \sqcap \mathit{Chair}) \sqsubseteq \mathit{Inflammable}$

---

4 Related and Future Works

In this work we have considered a nonmonotonic Description Logic $\mathbf{T}^{\mathsf{CL}}$, extending the DL of typicality $\mathcal{ALC} + \mathbf{T_R}$ with a DISPONTE semantics, in order to deal with the generation of novel creative concepts. This logic enjoys good computational properties, since entailment in it remains `ExpTime` as the underlying monotonic $\mathcal{ALC}$, and is able to take into account the concept combination of prototypical properties. To this aim, the logic $\mathbf{T}^{\mathsf{CL}}$ allows to have inclusions of the form $p \ :: \ \mathbf{T}(C) \sqsubseteq D$, representing that, with a probability p, typical Cs are also Ds. Then, several different scenarios –

having different probabilities – are described by including or not such inclusions, and prototypical properties of combinations of concepts are obtained by restricting reasoning services to scenarios having suitable probabilities, excluding "trivial" ones with the highest probabilities.

In AI, several approaches have been proposed to deal with the problem of prototypical concept composition in a human-like fashion. The authors of [10] present a detailed analysis of the limits of the set-theoretic approaches, the fuzzy logics (whose limitations was already shown in [11]), the vector-space models and quantum probability approaches proposed to model this phenomenon. In addition, they propose to use hierarchical conceptual spaces [12] to model the phenomenon in a way that accurately reflects how humans exploit their creativity in conjunctive concept combination. While we agree with the authors with the comments moved to the described approaches, we showed that our logic can equally model, in a cognitively compliant-way, the composition of prototypes by using a nonmonotonic formalism whose complexity remains in the same class of standard monotonic $\mathcal{ALC}$. Other attempts similar to the one proposed here concerns the conceptual blending: a task where the obtained concept is *entirely novel* and has no strong association with the two base concepts (while in concept combination the compound concept is always a subset of the base concepts, for details see [13]). In [14] the authors propose a mechanism for conceptual blending based on the DL $\mathcal{EL}^{++}$. They construct the generic space of two concepts by introducing an upward refinement operator that is used for finding common generalizations of $\mathcal{EL}^{++}$ concepts. However, differently from us, what they call prototypes are expressed in the standard monotonic DL, which does not allow to reason about typicality and defeasible inheritance. More recently, a different approach is proposed in [15], where the authors see the problem of concept blending as a nonmonotonic search problem and propose to use Answer Set Programming (ASP) to deal with this search problem in a nonmonotonic way. In a related work [16], the author extends the logic of typicality $\mathcal{ALC} + \mathbf{T_R}$ by means of probabilities equipping typicality inclusions of the form $\mathbf{T}(C) \sqsubseteq_p D$, whose intuitive meaning is that, "normally, Cs are Ds and we have a probability of $1 - p$ of having exceptional Cs not being Ds". Probabilities of exceptions are then used in order to reason about plausible scenarios, obtained by selecting only *some* typicality assumptions and whose probabilities belong to a given and fixed range. As a difference with the logic $\mathbf{T}^{\textsf{CL}}$, all typicality assumptions are systematically taken into account: as a consequence, one cannot exploit such a DL for capturing compositionality, since it is not possible to block inheritance of prototypical properties in concept combination. The logic $\mathbf{T}^{\textsf{CL}}$ extends the work of [16] in that it does not systematically take into account all typicality assumptions. As a consequence, $\mathbf{T}^{\textsf{CL}}$ allows to block inheritance of prototypical properties in concept combination. The same criticism applies also to the approach proposed in [17], where $\mathcal{ALC} + \mathbf{T_R}$ is extended by inclusions of the form $\mathbf{T}(C) \sqsubseteq_d D$, where d is a *degree of expectedness*, used to define a preference relation among extended ABoxes: entailment of queries is then restricted to ABoxes that are minimal with respect to such preference relations and that represent surprising scenarios. Also in this case, however, the resulting logic does not allow to define scenarios containing only some inclusions, since all of them are systematically considered. Similarly, probabilistic DLs [18] themselves cannot be employed as a framework for dealing with the combination of concepts, since these logics are not able to represent and reason about prototypical properties.

In future research we aim at extending our approach to more expressive DLs, such as those underlying the standard OWL language. Starting from the work of [19], applying the logic with the typicality operator and the rational closure to $\mathcal{SHIQ}$, we intend to study whether and how $\mathbf{T}^{\textsf{CL}}$ could provide an alternative solution to the problem of the "all or nothing" behavior of rational closure with respect to property inheritance. We also aim at implementing efficient reasoners for $\mathbf{T}^{\textsf{CL}}$, relying on the prover RAT-OWL [20] which allows to reason in the nonmonotonic logic $\mathcal{ALC} + \mathbf{T_R}$ underlying our approach. The knowledge base obtained by adding typicality inclusions of a compound concept remains in the language of the logic $\mathbf{T}^{\textsf{CL}}$, allowing to iterate the process in order to define concepts as combinations of – not necessarily atomic – existing concepts. We plan to investigate this opportunity in future works.

Acknowledgements

This work has been partially supported by the project "ExceptionOWL: Nonmonotonic Extensions of Description Logics and OWL for defeasible inheritance with exceptions", Università di Torino and Compagnia di San Paolo, call 2014 "Excellent (young) PI".

References

1. Frixione, M., Lieto, A.: Representing concepts in formal ontologies: Compositionality vs. typicality effects. Logic and Logical Philosophy **21**(4) (2012) 391–414
2. Osherson, D.N., Smith, E.E.: On the adequacy of prototype theory as a theory of concepts. Cognition **9**(1) (1981) 35–58
3. Lieto, A., Pozzato, G.L.: A description logic of typicality for conceptual combination. In: Proceedings of the 24th International Symposium on Methodologies for Intelligent Systems ISMIS2018. (2018)
4. Giordano, L., Gliozzi, V., Olivetti, N., Pozzato, G.L.: Semantic characterization of Rational Closure: from Propositional Logic to Description Logics. Artificial Intelligence **226** (2015) 1–33
5. Lehmann, D., Magidor, M.: What does a conditional knowledge base entail? Artificial Intelligence **55**(1) (1992) 1–60
6. Riguzzi, F., Bellodi, E., Lamma, E., Zese, R.: Reasoning with probabilistic ontologies. In Yang, Q., Wooldridge, M., eds.: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015, AAAI Press (2015) 4310–4316
7. Hampton, J.A.: Inheritance of attributes in natural concept conjunctions. Memory & Cognition **15**(1) (1987) 55–71
8. Lieto, A., Minieri, A., Piana, A., Radicioni, D.P.: A knowledge-based system for prototypical reasoning. Connection Science (27) (2015) 137–152
9. Hampton, J.A.: Conceptual combinations and fuzzy logic. Concepts and fuzzy logic **209** (2011)
10. Lewis, M., Lawry, J.: Hierarchical conceptual spaces for concept combination. Artificial Intelligence **237** (2016) 204–227

11. Smith, E.E., Osherson, D.N.: Conceptual combination with prototype concepts. Cognitive science **8**(4) (1984) 337–361
12. Gärdenfors, P.: The Geometry of Meaning: Semantics Based on Conceptual Spaces. MIT Press (2014)
13. Nagai, Y., Taura, T.: Formal description of concept-synthesizing process for creative design. Design computing and cognition06 (2006) 443–460
14. Confalonieri, R., Schorlemmer, M., Kutz, O., Peñaloza, R., Plaza, E., Eppe, M.: Conceptual blending in EL++. In Lenzerini, M., Peñaloza, R., eds.: Proceedings of the 29th International Workshop on Description Logics, Cape Town, South Africa, April 22-25, 2016. Volume 1577 of CEUR Workshop Proceedings., CEUR-WS.org (2016)
15. Eppe, M., Maclean, E., Confalonieri, R., Kutz, O., Schorlemmer, M., Plaza, E., Kühnberger, K.U.: A computational framework for conceptual blending. Artificial Intelligence **256** (2018) 105–129
16. Pozzato, G.L.: Reasoning in description logics with typicalities and probabilities of exceptions. In Antonucci, A., Cholvy, L., Papini, O., eds.: Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 14th European Conference, ECSQARU 2017, Lugano, Switzerland, July 10-14, 2017, Proceedings. Volume 10369 of Lecture Notes in Computer Science., Springer (2017) 409–420
17. Pozzato, G.L.: Reasoning about plausible scenarios in description logics of typicality. Intelligenza Artificiale **11**(1) (2017) 25–45
18. Riguzzi, F., Bellodi, E., Lamma, E., Zese, R.: Probabilistic description logics under the distribution semantics. Semantic Web **6**(5) (2015) 477–501
19. Giordano, L., Gliozzi, V., Olivetti, N., Pozzato, G.L.: Rational closure in $\mathcal{SHIQ}$. In: DL 2014, 27th International Workshop on Description Logics. Volume 1193 of CEUR Workshop Proceedings., CEUR-WS.org (2014) 543–555
20. Giordano, L., Gliozzi, V., Pozzato, G.L., Renzulli, R.: An efficient reasoner for description logics of typicality and rational closure. In Artale, A., Glimm, B., Kontchakov, R., eds.: Proceedings of the 30th International Workshop on Description Logics, Montpellier, France, July 18-21, 2017. Volume 1879 of CEUR Workshop Proceedings., CEUR-WS.org (2017)