A Toothful of Concepts: Towards a theory of weighted concept combination

Daniele Porello[1], Oliver Kutz[2], Guendalina Righetti[2], Nicolas Troquard[2], Pietro Galliani[2], and Claudio Masolo[1]

[1] Laboratory for Applied Ontology, ISTC-CNR, Trento, Italy
[2] KRDB Research Centre for Knowledge and Data, Free University of Bozen-Bolzano, Italy

Abstract. We introduce a family of operators to combine Description Logic concepts. They aim to characterise complex concepts that apply to instances that satisfy "enough" of the concept descriptions given. For instance, an individual might not have any tusks, but still be considered an elephant. To formalise the meaning of "enough", the operators take a list of weighted concepts as arguments, and a certain threshold to be met. We commence a study of the formal properties of these operators, and study some variations. The intended applications concern the representation of cognitive aspects of classification tasks: the interdependencies among the attributes that define a concept, the prototype of a concept, and the typicality of the instances.

CEUR-WS.org/Vol-2373/paper-24.pdf

1 Introduction

We begin the project of extending description logics to model cognitively relevant features of classification. We start from familiar Description Logic formalisms (in particular from $\mathcal{ALC}$), which is an important logical language to model concepts and concept combinations in knowledge representation. We introduce a family of operators which apply to sets of concept descriptions and return a composed concept whose instances are those that satisfy "enough" of the listed concept descriptions. To provide a meaning of "enough", the operator takes a list of weighted concepts as argument, as well as a threshold. The combined concept applies to every instance whose sum of the weights of the concepts it satisfies meets the threshold. Using a threshold, the presentation focuses on crisp categorisations. Although the framework of weighted concepts easily adapts to a many-valued setting, we do not admit degrees of classification here.

Depending on the base description logic used, the operators introduced might or might not extend the *extensional* expressivity of the concept language in the sense of increasing the expressive power to define new, previously undefinable concepts. However, they always allow for a more cognitively grounded modelling of the *intensional* aspects of classifications, which are concerned with how the parts of a concept definition contribute to the classification task overall. The operators also allow for more compact representations.

The approach to weighted logics that we follow here takes inspiration from the use of sets of weighted proposition for representing utility functions in [1]. Extensions of that approach to description logics have been developed in [2]. Two related articles are [3] and [4], where cognitive features of categorisation have been modelled by means of sets of weighted predicative formulas. The main difference in the present approach is that we study weighted combinations of concepts by explicitly introducing *syntactic operators* on concepts extending the basic concept languages, and investigating their logical properties.

The intended applications of this framework are inspired by the idea of providing a cognitively meaningful representation of classification tasks. Cognitive models of concepts and classification are usually grouped into the *prototype view*, the *exemplar view*, and the *knowledge view* also called *theory-theory* (see [5, 6]), but also Gärdenfors's *theory of conceptual spaces* [7] and Barsalou's *theory of frames* [8] enter this category. They hardly rely on logical representation of concepts, however. In this paper, we want to explore the possibility of extending logic-based representations of concepts to capture aspects of cognitive modelling.

In particular, we will see how the proposed operators allow for representing the *prototype* view of classification under a concept, cf. ([9]). Moreover, a number of cognitively relevant phenomena can be represented in this setting. For instance, the marginal contribution of the attributes entering the definition of a complex concept, the contextual dependence of a classification task on the available information, and the typicality of an instances, cf. [10]. Our work here is largely independent of the specifics of the concept language used; we will here use standard definitions and terminology from description logics [11], primarily working with the language $\mathcal{ALC}$.

2 Weighted concept combination

We introduce a class of m-ary operators, denoted by the symbol $\nabla\!\!\!\!\nabla$ (spoken 'tooth'), for combining concepts. Each operator works a follows: *i*) it takes a list of concept descriptions, *ii*) it associates a vector of weights to them, and *iii*) it returns a complex concept that applies to those instances that satisfy a certain combination of concepts, i.e., those instances for which, by summing up the weights of the satisfied concepts, a certain threshold is met.

The new logic is denoted by $\mathcal{ALC}_{\nabla\!\!\!\!\nabla^{\mathbb{R}}}$, where weights and thresholds range over real numbers $r \in \mathbb{R}$. In the following we will refer to the languages for brevity just as $\mathcal{ALC}_{\nabla\!\!\!\!\nabla}$. To define the extended language of $\mathcal{ALC}_{\nabla\!\!\!\!\nabla}$, we add combination operators as follows, which behave syntactically just like m-ary modalities. We assume a vector of m weights $\boldsymbol{w} \in \mathbb{R}^m$ and a threshold value $t \in \mathbb{R}$. Each pair $\boldsymbol{w}$, t specifies an operator: if $C_1, \ldots, C_m$ are concepts of $\mathcal{ALC}$, then $\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)$ is a concept of $\mathcal{ALC}_{\nabla\!\!\!\!\nabla}$. Note that in this basic definition, the possible nesting of the operator is excluded.[3]

[3] In a more fine-grained definition $\mathcal{ALC}^i_{\nabla\!\!\!\!\nabla^{\mathbb{K}}}$, $i \geq 0$, is the logic with i levels of allowed nesting and where weights and thresholds range over K; we will comment on this further below.

For $C'_i \in \mathcal{ALC}$, the set of $\mathcal{ALC}_{\nabla\!\!\!\!\nabla}$ concepts is then described by the grammar:

$$C ::= A \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \forall R.C \mid \exists R.C \mid \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C'_1, \ldots, C'_m)$$

The semantics of the operator is obtained by extending the definition of the semantics of $\mathcal{ALC}$ as follows. Let $I = (\Delta^I, \cdot^I)$ be an interpretation of $\mathcal{ALC}$. We define the *value* of a $d \in \Delta^I$ under a $\nabla\!\!\!\!\nabla$-concept $\mathtt{C} = \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)$ by setting:

$$v^I_{\mathtt{C}}(d) = \sum_{i \in \{1,\ldots,m\}} \{w_i \mid d \in C^I_i\} \tag{1}$$

The interpretation (i.e., the extension) of a $\nabla\!\!\!\!\nabla$-concept in $I = (\Delta^I, \cdot^I)$ is then:

$$(\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I = \{d \in \Delta^I \mid v^I_{\mathtt{C}}(d) \geq t\} \tag{2}$$

To better visualise the weights an operator associates to the concepts, we sometimes use the notation $\nabla\!\!\!\!\nabla^t((C_1, w_1), \ldots, (C_m, w_m))$ instead of $\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)$.

In the following examples, we will consider the value of an object name a (aka individual constant) wrt. a $\nabla\!\!\!\!\nabla$-concept for interpretations that satisfy a certain knowledge base $\mathcal{K}$ (i.e. a set of formulas).

Definition 1 (Weights relative to a knowledge base). *Let a be an object name of $\mathcal{ALC}$ and $\mathcal{K}$ an $\mathcal{ALC}$ knowledge base. We set*

$$v^{\mathcal{K}}_{\mathtt{C}}(a) := \sum_{i \in \{1,\ldots,m\}} \{w_i \mid \mathcal{K} \models C_i(a)\}$$

I.e., $v^{\mathcal{K}}_{\mathtt{C}}(a)$ gives the accumulated weight of those C_i that are entailed by $\mathcal{K}$ to satisfy a.

Note that for positive weights, a given name a and a fixed interpretation I such that $I \models \mathcal{K}$, we always have that $v^{\mathcal{K}}_{\mathtt{C}}(a) \leq v^I_{\mathtt{C}}(a^I)$.

Example 1. Consider the set of concepts $\mathcal{C} = \{\mathtt{Red}, \mathtt{Round}, \mathtt{Coloured}\}$ and the concept $\mathtt{C}$ defined by means of the $\nabla\!\!\!\!\nabla$ operator

$$\mathtt{C} = \nabla\!\!\!\!\nabla^t((\mathtt{Red} \sqcup \mathtt{Round}, w_1), (\exists \mathtt{above.Coloured}, w_2))$$

The definition of $\mathtt{C}$ means that the relevant information to establish the categorisation under $\mathtt{C}$ of an object is whether (*i*) it is red or round, and (*ii*) it is above a coloured thing.

Consider the following knowledge base $\mathcal{K} = \{\mathtt{Red}(a), \exists \mathtt{above.Blue}(a), \mathtt{Blue} \sqsubseteq \mathtt{Coloured}\}$, i.e., an agent knows that the object a is red and it is above a blue thing and that blue things are coloured, $\mathtt{Blue} \sqsubseteq \mathtt{Coloured}$.

The value of a returned by $v^{\mathcal{K}}_{\mathtt{C}}$ is computed as follows. Firstly, if a satisfies $\mathtt{Red}$, then a satisfies $\mathtt{Red} \sqcup \mathtt{Round}$, so the weight w_1 can be obtained. Moreover, since $\mathtt{Blue} \sqsubseteq \mathtt{Coloured} \in \mathcal{K}$ and a satisfies $\exists \mathtt{above.Blue}$, then a satisfies $\exists \mathtt{above.Coloured}$, so also the weight w_2 can be obtained. Thus, $v^{\mathcal{K}}_{\mathtt{C}}(a)$ is $w_1 + w_2$. If $w_1 + w_2 \geq t$, then a is classified under $\mathtt{C}$.

2.1 General properties of the $\nabla\!\!\!\nabla$ concept constructor

We discuss a few general properties of the $\nabla\!\!\!\nabla$ operators which allow for reasoning about combinations of concepts.

Firstly, we note that, for every possible choice of weights and thresholds, the operator is well-defined: the $\nabla\!\!\!\nabla$s of equivalent concepts return equivalent concepts, i.e. equivalence is a congruence for the tooth. For every I,

$$C_i^I = D_i^I \implies (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_i, \ldots, C_m))^I = (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, D_i, \ldots, C_m))^I \quad (3)$$

Proof. Assume an interpretation I such that $C_i^I = D_i^I$. Suppose that $d \in (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_i, \ldots, C_m))^I$, thus, by definition, $\sum\{w_i \mid d \in C_i^I\} \geq t$. Since D_i is equivalent to C_i, d is also in $(\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, D_i, \ldots, C_m))^I$. □

Consider now the following statement, which resembles the monotonicity condition of (normal) modal operators. The statement holds true whenever the weights are non-negative numbers, i.e. for $w_i \in \mathbb{R}_0^+$ we have:

$$C_i^I \subseteq D_i^I \implies (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_i, \ldots, C_m))^I \subseteq (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, D_i, \ldots, C_m))^I \quad (4)$$

Proof. We assume all weights are non-negative and establish 4. Assume that $C_i^I \subseteq D_i^I$. Suppose $d \in \nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_i, \ldots, C_m))^I$, then, by definition, $\sum\{w_i \mid d \in C_i\} \geq t$. We have two relevant cases: (i) $d \in C_i^I$, thus, by assumption, $d \in D_i^I$. Since the weight associated to D_i is the same as the weight associated to C_i the sum does not change. (ii) Suppose $d \notin C_i^I$ and $d \in D_i^I$. In this case, the sum now adds the weight associated to D_i. Since the weights are non-negative, the sum is increasing, thus $\sum\{w_i \mid d \in C_i\} + w_i$ is still greater than t. □

Extending on this result, only under certain conditions does the tooth operator fit in between the conjunction and the disjunction of the concepts. Namely,

$$t > 0 \implies (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I \subseteq (C_1 \sqcup \cdots \sqcup C_m)^I \quad (5)$$

Indeed, for any $d \notin (C_1 \sqcup \ldots \sqcup C_m)^I$ the value $v_{\mathtt{C}}^I(d)$ would be zero, and hence $d \notin \nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I$.

On the other hand,

$$t \leq \sum_i w_i \implies (C_1 \sqcap \cdots \sqcap C_m)^I \subseteq (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I \quad (6)$$

Indeed, for any $d \in (C_1 \sqcap \ldots \sqcap C_m)^I$ we have that $v_{\mathtt{C}}^I(d) = \sum_i w_i \geq t$ and hence that $d \in (\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I$.

Moreover, if the set of weights is restricted to non-negative numbers, $w_i \in \mathbb{R}_0^+$, then:

$$(\nabla\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I \subseteq (\nabla\!\!\!\nabla^t_{\boldsymbol{w}, w_{m+1}}(C_1, \ldots, C_m, C_{m+1}))^I \quad (7)$$

That is, by adding positive attributes to the definition of a concept, we cannot invalidate the categorisation of an instance under the concept.

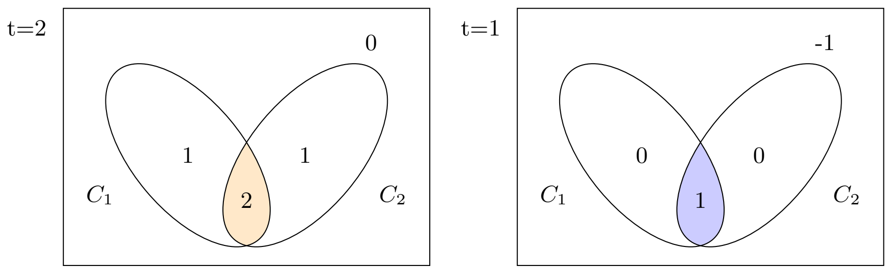

Fig. 1. Consider: $\nabla\!\!\!\!\nabla^2_{(1,1)}(C_1, C_2)$ (on the left) and $\nabla\!\!\!\!\nabla^1_{(1,1,-1)}(C_1, C_2, \top)$ (to the right).

The properties of weights affect the classification in the following way. Uniform permutations of weights and concepts arguments correspond to the same concept. For every permutation σ:

$$(\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I = (\nabla\!\!\!\!\nabla^t_{\sigma(\boldsymbol{w})}(\sigma(C_1, \ldots, C_m)))^I \tag{8}$$

When $C_1^I = C_2^I$ (in particular when $C_1 = C_2$):

$$(\nabla\!\!\!\!\nabla^t_{(w_1,\ldots,w_m)}(C_1, \ldots, C_m))^I = (\nabla\!\!\!\!\nabla^t_{(w_1+w_2,w_3,\ldots,w_m)}(C_1, C_3, \ldots, C_m))^I \tag{9}$$

Moreover, every positive transformation of weights and thresholds returns the same sets of entities. For every $k > 0$, we have that:

$$(\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m))^I = (\nabla\!\!\!\!\nabla^{k \cdot t}_{k \cdot \boldsymbol{w}}(C_1, \ldots, C_m))^I \tag{10}$$

For every k, we have that:

$$(\nabla\!\!\!\!\nabla^t_{(w_1,\ldots,w_m)}(C_1, \ldots, C_m))^I = (\nabla\!\!\!\!\nabla^{t+k}_{(w_1,\ldots,w_m,k)}(C_1, \ldots, C_m, \top))^I \tag{11}$$

One example of Eq. 11, which can be represented as shown in Figure 1, is the following:

$$(C_1 \sqcap C_2)^I = (\nabla\!\!\!\!\nabla^2_{(1,1)}(C_1, C_2))^I = (\nabla\!\!\!\!\nabla^1_{(1,1,-1)}(C_1, C_2, \top))^I$$

3 Expressivity and Definability

We discuss in this section the definability of concepts from the purely *extensional* point of view. Every (complex) concept C of $\mathcal{ALC}$ can be trivially represented by means of $\nabla\!\!\!\!\nabla^t_{(t)}(C)$. However, it is interesting to discuss whether we can provide a representation of C in terms of weights to be associated with the atomic concepts that define C. We first focus solely on the Boolean structure of complex concepts followed by a discussion of counting and maximisation.

3.1 Boolean operations and the $\mathbb{W}$

For instance, the Boolean operators can be expressed as special cases of $\mathbb{W}$ of atomic concepts:

- $(C_1 \sqcap C_2)^I = (\mathbb{W}^2_{(1,1)}(C_1, C_2))^I$
- $(C_1 \sqcup C_2)^I = (\mathbb{W}^1_{(1,1)}(C_1, C_2))^I$
- $(\neg C_1)^I = (\mathbb{W}^0_{(-1)}(C_1))^I$

More generally, some Boolean functions can be captured with $\mathbb{W}$, without having recourse to complex concepts as arguments. In the case of Boolean functions over two variables (i.e. atomic concepts) C_1 and C_2, we obtain the following:

$$
\begin{array}{llll}
\top & = \mathbb{W}^0_{(0,0)}(C_1, C_2) & \bot & = \mathbb{W}^1_{(0,0)}(C_1, C_2) \\
C_1 & = \mathbb{W}^1_{(1,0)}(C_1, C_2) & C_2 & = \mathbb{W}^1_{(0,1)}(C_1, C_2) \\
\neg C_1 & = \mathbb{W}^0_{(-1,0)}(C_1, C_2) & \neg C_2 & = \mathbb{W}^0_{(0,-1)}(C_1, C_2) \\
C_1 \sqcap C_2 & = \mathbb{W}^2_{(1,1)}(C_1, C_2) & \neg C_1 \sqcap C_2 & = \mathbb{W}^1_{(-1,1)}(C_1, C_2) \\
C_1 \sqcap \neg C_2 & = \mathbb{W}^1_{(1,-1)}(C_1, C_2) & \neg C_1 \sqcap \neg C_2 & = \mathbb{W}^0_{(-1,-1)}(C_1, C_2) \\
C_1 \sqcup C_2 & = \mathbb{W}^1_{(1,1)}(C_1, C_2) & \neg C_1 \sqcup C_2 & = \mathbb{W}^0_{(-1,1)}(C_1, C_2) \\
C_1 \sqcup \neg C_2 & = \mathbb{W}^0_{(1,-1)}(C_1, C_2) & \neg C_1 \sqcup \neg C_2 & = \mathbb{W}^{-1}_{(-1,-1)}(C_1, C_2)
\end{array}
$$

The operator $\mathbb{W}$ is thus, by itself, a functionally complete logical connective if we allow for nesting the operator. However, it is impossible to represent $(C_1 \sqcap \neg C_2) \sqcup (\neg C_1 \sqcap C_2)$ (the symmetric difference, XOR) and $(C_1 \sqcap C_2) \sqcup (\neg C_1 \sqcap \neg C_2)$ (both or none, the negation of XOR) without recursion (that is, nesting of the $\mathbb{W}$), complex concepts as arguments, or without the Boolean combination of more than one $\mathbb{W}$.

Indeed, suppose that the symmetric difference C_1 XOR C_2 is definable as an expression of the form $\mathbb{W}^t_{\boldsymbol{w}}(C_1, C_2, \top, \bot)$ for $\boldsymbol{w} = (w_1, w_2, w_\top, w_\bot)$. Then it can be easily verified that $w_1 > 0$, since for $d \notin C_1^I \cup C_2^I$ we must have that $v_{\mathsf{C}}^I(d) = w_\top < t$ while for $d \in C_1^I \setminus C_2^I$ we have that $v_{\mathsf{C}}^I(d) = w_\top + w_1 \geq t$. A similar argument shows that $w_2 > 0$ as well; but then, for $d \in C_1^I \cap C_2^I$ we have that $v_{\mathsf{C}}^I(d) = w_\top + w_1 + w_2 > t$ and hence $I, d \models \mathbb{W}^t_{\boldsymbol{w}}(C_1, C_2, \top, \bot)$.

Of course, one can use the following Boolean combinations, using the previous characterizations:

- $(C_1 \sqcap \neg C_2) \sqcup (\neg C_1 \sqcap C_2) = \mathbb{W}^1_{(1,-1)}(C_1, C_2) \sqcup \mathbb{W}^0_{(-1,1)}(C_1, C_2)$
- $(C_1 \sqcap C_2) \sqcup (\neg C_1 \sqcap \neg C_2) = \mathbb{W}^2_{(1,1)}(C_1, C_2) \sqcup \mathbb{W}^{-1}_{(-1,-1)}(C_1, C_2)$

With complex concept arguments, we have:

- $(C_1 \sqcap \neg C_2) \sqcup (\neg C_1 \sqcap C_2) = \mathbb{W}^1_{(1,1,-2)}(C_1, C_2, C_1 \sqcap C_2)$
- $(C_1 \sqcap C_2) \sqcup (\neg C_1 \sqcap \neg C_2) = \mathbb{W}^0_{(-1,-1,2)}(C_1, C_2, C_1 \sqcap C_2)$

With nesting, we could do:

- $(C_1 \sqcap \neg C_2) \sqcup (\neg C_1 \sqcap C_2) = \mathbb{W}^1_{(1,1,-2)}(C_1, C_2, \mathbb{W}^2_{(1,1)}(C_1, C_2))$

– $(C_1 \sqcap C_2) \sqcup (\neg C_1 \sqcap \neg C_2) = \nabla\!\!\!\!\nabla^0_{(-1,-1,2)}(C_1, C_2, \nabla\!\!\!\!\nabla^2_{(1,1)}(C_1, C_2))$

On the other hand, given any expression $\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)$, consider the set $\varGamma = \{U \subseteq \{1 \ldots n\} : \sum_{i \in U} w_i \geq t\}$ of all sets of indices that, if they were the only ones corresponding to concepts that are satisfied by an individual d in an interpretation I, would make it so that $d \in \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)^I$. Then we have at once that $\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)$ is logically equivalent to

$$\bigsqcup_{U \in \varGamma} \left(\bigsqcap_{i \in U} C_i \sqcap \bigsqcap_{i \in \{1 \ldots n\} \setminus U} \neg C_i \right)$$

Notice that this translation may lead to an exponentially longer formula.

3.2 Counting and Majority

The combination operators can be instantiated to capture, in a compact way, a large class of concept compositions. Suppose that `D` is a complex concept whose definition applies the concepts $C_1, \ldots, C_m$.

We can define a concept `D` such that d is in $\mathtt{D}^I$ iff *the majority of concepts* among $C_1 \ldots, C_m$ applies to d. In this case, we are assuming that all the m concepts have equal weight and d is classified under the majority iff it satisfies more than $\frac{m}{2}$ concepts. Equivalently, we associate weight 2 to each concept C_i and we set $t = m$: $\nabla\!\!\!\!\nabla^m((C_1, 2), \ldots, (C_m, 2))$. In a similar way, we can introduce operators that use a quota rule, i.e they apply to instances that satisfy at least a number $q \in \{1, \ldots, m\}$ of the concepts in the scope of $\nabla\!\!\!\!\nabla$.

Moreover, a preferential structure on the combined concepts can be rendered by means of (Borda) scores: $\nabla\!\!\!\!\nabla^t((C_1, s_1), \ldots, (C_m, s_m))$, where the weights are subject to the constraint $s_1 \geq \cdots \geq s_m$. This means that, e.g., satisfying C_1 is more important than C_2, and so on. By choosing the weights and the threshold in a suitable way, we can then express situations where satisfying the first l concepts is significant, or dominating for the classification. I.e., the combined weight of the first l concepts will suffice to classify under the concept, and if weights are set high enough, this condition can be turned into a necessary one.

Finally, it is possible to define the set of instances that at most reach a given threshold:

$$\nabla\!\!\!\!\nabla^{\leq t}((C_1, w_1), \ldots (C_m, w_m) \equiv \nabla\!\!\!\!\nabla^{-t}((C_1, -w_1), \ldots, (C_m, -w_m)) \tag{12}$$

and the concept of instances that exactly score a certain threshold value t:

$$\nabla\!\!\!\!\nabla^{=t}((C_1, w_1), \ldots (C_m, w_m) \equiv \nabla\!\!\!\!\nabla^{t}((C_1, w_1), \ldots, (C_m, w_m) \sqcap \nabla\!\!\!\!\nabla^{\leq t}((C_1, w_1), \ldots (C_m, w_m) \tag{13}$$

3.3 Maximisation and the Universal Modality

Finally, it is interesting to consider the set of entities that *maximally* satisfy a combination of concepts $C_1, \ldots C_m$ of $\mathcal{ALC}$. That is, we may define an operator with the following semantics:

$$\left(\nabla\!\!\!\!\nabla^{max}((C_1, w_1), \ldots (C_m, w_m))\right)^I = \{d \in \Delta \mid v^I_{\mathsf{C}}(d) \geq v^I_{\mathsf{C}}(d') \text{ for all } d' \in \Delta\} \tag{14}$$

Defining $\nabla\!\!\!\!\nabla^{max}$ in terms of $\nabla\!\!\!\!\nabla^t$ would require to use a *universal role*, which significantly increases the expressive power of $\mathcal{ALC}$ [12].

First of all, let us see how the $\nabla\!\!\!\!\nabla^{max}$ operator may be defined in terms of the universal modality. Given a tuple $\boldsymbol{w} = (w_1 \ldots w_m)$ of weights, let $S = \{\sum_i a_i w_i : a_i \in \{0, 1\}\}$ be the (finite) set of all possible scores that can be obtained by using these weights; and then, let $\epsilon = \min\{|v - v'| : v, v' \in S\}$ be the smallest possible distance between different scores. Then we have that

$$\nabla\!\!\!\!\nabla^{max}_{\boldsymbol{w}}(C_1 \ldots C_m) \equiv \bigsqcup_{t \in S} \left(\nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1 \ldots C_m) \sqcap \forall U.\neg \nabla\!\!\!\!\nabla^{t+\epsilon}_{\boldsymbol{w}}(C_1 \ldots C_m)\right) \tag{15}$$

Proof. Indeed, suppose that $I, d \models \nabla\!\!\!\!\nabla^{max}_{\boldsymbol{w}}(C_1 \ldots C_m)$ for some individual d of our domain.[4] Now let $t = v^I_{\mathsf{C}}(d)$: then, by definition, $t \geq v^I_{\mathsf{C}}(d')$ for all individuals $d' \in \Delta$. This implies that $I, d \models \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1 \ldots C_m)$ and that $I, d' \not\models \nabla\!\!\!\!\nabla^{t+\epsilon}_{\boldsymbol{w}}(C_1 \ldots C_m)$ for all $d' \in \Delta$, which implies at once that $I, d \models \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1 \ldots C_m) \sqcap \forall U.\neg \nabla\!\!\!\!\nabla^{t+\epsilon}_{\boldsymbol{w}}(C_1 \ldots C_m)$ as required.

Conversely, suppose that $I, d \models \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1 \ldots C_m) \sqcap \forall U.\neg \nabla\!\!\!\!\nabla^{t+\epsilon}_{\boldsymbol{w}}(C_1 \ldots C_m)$ for some $t \in S$. Then $v^I_{\mathsf{C}}(d) = t$, and for all individuals $d' \in \Delta$ we have that $I, d' \not\models \nabla\!\!\!\!\nabla^{t+\epsilon}_{\boldsymbol{w}}(C_1 \ldots C_m)$, which by the definition of ϵ implies at once that $v^I_{\mathsf{C}}(d') \leq t$. Thus $I, d \models \nabla\!\!\!\!\nabla^{max}_{\boldsymbol{w}}(C_1 \ldots C_m)$, and this concludes the proof. □

Conversely, given the $\nabla\!\!\!\!\nabla^{max}$ operator it is possible to define the universal modality as

$$\forall U.C \equiv \left(\nabla\!\!\!\!\nabla^{max}_{(-1)}(C)\right) \sqcap C \tag{16}$$

Proof. Indeed, suppose that $I, d \models \forall U.C$. This means that $I, d' \models C$ for all individuals $d' \in \Delta$; and, therefore, $v^I_{(C,-1)}(d') = -1$ for all $d' \in \Delta$. In particular, we thus have that $I, d \models C$; and moreover, $v^I_{(C,-1)}(d) = -1 = v^I_{(C,-1)}(d')$ for all $d' \in \Delta$, and so $I, d \models \nabla\!\!\!\!\nabla^{max}_{(-1)}(C)$. So $I, d \models (\nabla\!\!\!\!\nabla^{max}_{(-1)}(C)) \sqcap C$, as required.

Conversely, if $I, d \models (\nabla\!\!\!\!\nabla^{max}_{(-1)}(C)) \sqcap C$ then first of all we have that $I, d \models C$ and that therefore $v^I_{(C,-1)}(d) = -1$. Now take any other individual $d' \in \Delta$. We state that $I, d' \models C$ as well; indeed, if instead $I, d' \not\models C$ we would have that $v^I_{(C,-1)}(d') = 0 > v^I_{(C,-1)}(d)$, which is impossible since by assumption $I, d \models (\nabla\!\!\!\!\nabla^{max}_{(-1)}(C))$. Thus $C^I = \Delta$ as required. □

We leave a detailed study of the $\nabla\!\!\!\!\nabla^{max}$ operator and similar extensions for future work.

[4] This is a mild abuse of notation for $d \in \nabla\!\!\!\!\nabla^{max}_{\boldsymbol{w}}(C_1 \ldots C_m)^I$. In what follows, we will freely write expressions like $I, d \models C$ with the intended meaning of $d \in C^I$.

4 Modelling with the Tooth Operator

We see now how the $\nabla\!\!\!\!\nabla$ operators allow for representing fine-grained dependencies among the attributes that define a concept. We may call this feature *intensional expressivity*. We study situations where an agent who has knowledge represented by $\mathcal{K}$ performs the task of classifying an object a under the concept $\mathtt{C} = \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)$. That is, we focus on how $\mathcal{K} \models \nabla\!\!\!\!\nabla^t_{\boldsymbol{w}}(C_1, \ldots, C_m)\ (a)$ and we discuss how the pieces of knowledge in $\mathcal{K}$ may contribute to satisfy the attributes occurring in $\mathtt{C}$.

We say that the function $v^{\mathcal{K}}_{\mathtt{C}}$ is *additive* on the information in $\mathcal{K} = \{\phi_1, \ldots, \phi_l\}$ iff for every individual a, $v^{\mathcal{K}}_{\mathtt{C}}(a) = \sum\{v^{\{\phi_i\}}_{\mathtt{C}}(a) \text{ for } \phi_i \in \mathcal{K}\}$. In this case, the satisfaction of each ϕ_i contributes independently of the other formulas in $\mathcal{K}$. We say that the function $v^{\mathcal{K}}_{\mathtt{C}}$ is *super-additive* on $\mathcal{K}$ iff for every a, $v^{\mathcal{K}}_{\mathtt{C}}(a) \geq \sum\{v^{\{\phi_i\}}_{\mathtt{C}}(a) \text{ for } \phi_i \in \mathcal{K}\}$ and *sub-additive* iff $v^{\mathcal{K}}_{\mathtt{C}}(a) \leq \sum\{v^{\{\phi_i\}}_{\mathtt{C}}(a) \text{ for } \phi_i \in \mathcal{K}\}$. Super-additivity represents positive synergies between the information provided by $\mathcal{K}$, whereas sub-additivity expresses negative synergies. We illustrate this by means of the following examples.

Example 2. Suppose that the concept *elephant* $\mathtt{E}$ is defined by four attributes:

$$\mathtt{E} = \nabla\!\!\!\!\nabla^t((\mathtt{Large}, w_1), (\mathtt{Heavy}, w_2), (\mathtt{hasTrunk}, w_3), (\mathtt{Grey}, w_4))$$

This definition entails that each of the attributes in $\mathtt{E}$ contributes independently of the others to the classification of an object as an elephant.

Consider a knowledge base $\mathcal{K} = \{\mathtt{Large}(a), \mathtt{Heavy}(a), \mathtt{hastrunk}(a), \mathtt{Grey}(a)\}$. Then, $v^{\mathcal{K}}_{\mathtt{E}}(a) = v^{\mathtt{Large}(a)}_{\mathtt{E}}(a) + v^{\mathtt{Heavy}(a)}_{\mathtt{E}}(a) + v^{\mathtt{Grey}(a)}_{\mathtt{E}}(a) + v^{\mathtt{hasTrunk}(a)}_{\mathtt{E}}(a)$, that is, $v^{\mathcal{K}}_{\mathtt{E}}$ is an additive wrt. the formulas in $\mathcal{K}$.

By exploiting complex concept descriptions in the scope of $\nabla\!\!\!\!\nabla$, we can enable positive and negative synergies among the attributes.

Example 3. Suppose now we redefine the concept *elephant* as follows, where we suppose that $w_5 > w_1 + w_3$, $w_6 > w_2 + w_3$, and $w_7 > w_3 + w_4$ and each W_i is positive.

$$\begin{aligned}\mathtt{E}' = &\nabla\!\!\!\!\nabla^t((\mathtt{Large}, w_1), (\mathtt{Heavy}, w_2), (\mathtt{hasTrunk}, w_3), (\mathtt{Grey}, w_4),\\ &(\mathtt{Large} \sqcap \mathtt{hasTrunk}, w_5), (\mathtt{Heavy} \sqcap \mathtt{hasTrunk}, w_6), (\mathtt{Grey} \sqcap \mathtt{hasTrunk}, w_7))\end{aligned}$$

In this case, the relevance of the pieces of information in $\mathcal{K} = \{\mathtt{Large}(a), \mathtt{Heavy}(a), \mathtt{hasTrunk}(a), \mathtt{Grey}(a)\}$ outweighs the sum of the values of each attribute. That is, $v^{\mathcal{K}}_{\mathtt{E}'}(a) \geq v^{\mathtt{Large}(a)}_{\mathtt{E}'}(a) + v^{\mathtt{Heavy}(a)}_{\mathtt{E}'}(a) + v^{\mathtt{Grey}(a)}_{\mathtt{E}'}(a) + v^{\mathtt{hasTrunk}(a)}_{\mathtt{E}'}(a)$.

The meaning of this representation is that adding "having a trunk" significantly increases the importance of the combination of attributes for classifying an elephant. Accordingly the value of $v^{\mathcal{K}}_{\mathtt{E}'}$ is in this case super-additive on $\mathcal{K}$.

In the case of sub-additive functions, the combination of two or more attributes may lower the salience for the classification. The combination of the use of conjunctions of concepts and negative weights allows for modelling how certain attributes may decrease the salience of the combination of other attributes.

Example 4. Suppose that we want to classify an individual according to the disease that she may suffer. For instance, the concept of flu may be represented as follows:

$$\begin{aligned}\texttt{FLU} =& \nabla\!\!\!\!\nabla((\texttt{Fever} \sqcap \texttt{Nausea}, w_1), (\texttt{Fever} \sqcap \texttt{Spots}, -w_2), (\texttt{Nausea} \sqcap \texttt{Spots}, -\texttt{w}_3),\\ &(\texttt{Fever}, w_4), (\texttt{Nausea}, w_5))\end{aligned}$$

In this case, the combination of fever and nausea is highly significant for the diagnosis, whereas adding the symptom 'spots' significantly decreases the reliability of the classification under FLU, because it is a strong indication of chickenpox. So, $w_1 > w_4 + w_5$ and w_2 and w_3 are both greater than $w_1 + w_4 + w_5$. That is, the function $v^{\mathcal{K}}_{\texttt{FLU}}$ is sub-additive on $\mathcal{K} = \{\texttt{Fever}(a), \texttt{Nausea}(a), \texttt{Spots}(a)\}$.

Complex concepts occurring as arguments of $\nabla\!\!\!\!\nabla$ provide a way to express many types of dependencies among the weights of the attributes. A comprehensive study of this type of expressivity is left for a future work and requires rephrasing the results in [1] to the case of DLs.

5 Applications: prototypes, typicality, and similarity

We illustrate how the $\nabla\!\!\!\!\nabla$ operators can represent the cognitive approach to concepts based on *prototypes*. In particular, we follow [13] where a "prototype is a prestored representation of the usual properties associated with the concept's instances"[13, p.487]. A *prototype* is represented in terms of a set of *attributes* (e.g., colour or weight) and a set of *values* for each attribute (e.g., red and blue for the colour-attribute). The relevance of an attribute for classifying an object—e.g., the relevance of colour for classifying apples—is represented by its *diagnosticity* (a numerical value), while the *salience* of an attribute-value represents its *typicality* (also a numerical value)—e.g., how frequent is for apple to have a red colour.

Let us define a prototype $\pi_{\texttt{C}}$ for a concept C as (where Q^j_i is the i-th value of the j-th attribute, s^j_i is the salience of Q^j_i wrt. C, d^j is the diagnosticity of the j-th attribute wrt. C):

$$\pi_{\texttt{C}} = \{(Q^1_1, s^1_1{\cdot}d^1), \ldots, (Q^1_r, s^1_r{\cdot}d^1), \ldots, (Q^n_1, s^n_1{\cdot}d^n), \ldots, (Q^n_m, s^n_m{\cdot}d^n)\}.$$

Note that here we weighted the salience of each attribute-value with the diagnosticity of the attribute, but more elaborate strategies exist. Furthermore, the attributes within a single dimension are assumed as mutually exclusive, e.g., if something is red, then it is not of any other colour, formally $Q^j_i \cap Q^j_k = \emptyset$ for $i \neq k$. Assume now to know some attribute-values of a given object a. The classification of a under the concept C is usually done by leveraging on a (usually metric) function that establishes, on the basis of the matching of features, how similar the object a and the prototype of C are.

In our setting, we can introduce a concept C by using a $\nabla\!\!\!\!\nabla$ operator that directly considers the attribute-values, the saliences, and the diagnosticities in the prototype $\pi_{\mathtt{C}}$ (where we assume $Q_i^j \sqcap Q_k^j \equiv \bot$ for $i \neq k$):

$$\mathtt{C} = \nabla\!\!\!\!\nabla^t((Q_1^1, s_1^1{\cdot}d_1), \ldots, (Q_r^1, s_r^1{\cdot}d_1), \ldots, (Q_1^n, s_1^n{\cdot}d_n), \ldots, (Q_m^n, s_m^n{\cdot}d_n))$$

The classification under C applies then to the objects that have enough features in common with the prototype to exceed the threshold t, i.e., they are *close enough* with respect to the prototype. We can also individuate the prototypical instances of C as the objects (if they exist) that satisfy all the Q_j^i in C.

Note that the object-prototype similarity is here simply rendered by summing up all the weights of the matching Q_j^i. The discussion of more sophisticated choices to measure the distance to the prototype is left for future work. Moreover, in our setting, we can enable synergies between attributes in a super-additive or in a sub-additive fashion. However, to define prototypes in this rich setting, we will use the $\nabla\!\!\!\!\nabla^{max}$ operator, or approximate it by carefully selecting the threshold.

This setting allows for defining a *typicality operator*, as in [10]. The most typical instances of a concept can be defined as those instances that maximise the sum of weights in $\boldsymbol{w}$. Notice that the instances can be ordered in terms of typicality with respect to a concept C, by means of the values $v_{\mathtt{C}}^I$ (cf. [10]). Thus, in principle, we can enable comparisons between instances concerning their typicality wrt. a concept and introduce various degrees of typicality.

Finally, the $\nabla\!\!\!\!\nabla^{=t}$ operator may be used to define similarity of instances with respect to a (number of) concept(s), that is, those instances that are not distinguishable in terms of the complex concept.

6 Conclusion and future work

We introduced a class of operators for defining complex concepts that weigh the role of the defining attributes. We presented a few general properties of these operators, and we started investigating their expressivity. Pinpointing the exact expressivity of various combinations of DLs and $\nabla\!\!\!\!\nabla$-operators as well as studying succinctness effects is part of future work.

We further illustrated how the operators may be applied to render cognitively meaningful mechanisms for classification. Future work will be dedicated to properly investigate the logical properties of the operators and their natural extensions, and to apply them to describe salient cognitive features of concepts such as concept combinations and blending [14]. Another line of research will deepen the comparison with the formal studies on typicality (e.g. [10]), work on threshold concepts [15], relaxed query answering [16], and the relation and combination with similarity frameworks based on a notion of distance (e.g. [17–19]).

References

1. J. Uckelman, Y. Chevaleyre, U. Endriss, and J. Lang. Representing utility functions via weighted goals. *Mathemathical Logic Quartely*, 55(4):341–361, 2009.
2. A. Ragone, T. D. Noia, F. M. Donini, E. D. Sciascio, and M. P. Wellman. Weighted description logics preference formulas for multiattribute negotiation. In *Scalable Uncertainty Management, Third International Conference, SUM 2009, Washington, DC, USA, September 28-30, 2009. Proceedings*, pages 193–205, 2009.
3. C. Masolo and D. Porello. Representing concepts by weighted formulas. In *Formal Ontology in Information Systems - Proceedings of the 10th International Conference, FOIS 2018, Cape Town, South Africa, 19-21 September 2018*, pages 55–68, 2018.
4. C. Masolo and D. Porello. Understanding predication in conceptual spaces. In R. Ferrario and W. Kuhn, editors, *Formal Ontology in Information Systems: Proceedings of the 9th International Conference (FOIS 2016)*, pages 139–152. IOS Press, 2016.
5. E. Margolis and S. Laurence, editors. *Concepts: Core Readings.* MIT Press, 1999.
6. G. L. Murphy. *The Big Book of Concepts.* MIT press, 2002.
7. P. Gärdenfors. *Conceptual spaces - The geometry of thought.* MIT Press, 2000.
8. L. W. Barsalou. Frames, concepts, and conceptual fields. In A. Leherer and E. F. Kittay, editors, *Frames, Fields, and Contrasts - New Essays in Semantic and Lexical Organization*, chapter 1, pages 21–74. Lawrence Erlbaum Associates, Inc, 1992.
9. E. Rosch. Principles of categorization. In E. Margolis and S. Laurence, editors, *Concepts: Core Readings*, volume 189, chapter 8, pages 189–206. MIT press, 1999.
10. A. Lieto and G. L. Pozzato. A description logic of typicality for conceptual combination. In *Foundations of Intelligent Systems - 24th International Symposium, ISMIS 2018, Limassol, Cyprus, October 29-31, 2018, Proceedings*, pages 189–199, 2018.
11. F. Baader, I. Horrocks, C. Lutz, and U. Sattler. *An Introduction to Description Logic.* Cambridge University Press, 2017.
12. E. Hemaspaandra. The price of universality. *Notre Dame Journal of Formal Logic*, 37:174–203, 1996.
13. E. E. Smith, D. N. Osherson, L. J. Rips, and M. Keane. Combining prototypes: A selective modification model. *Cognitive Science*, 12(4):485–527, 1988.
14. M. Eppe, E. Maclean, R. Confalonieri, O. Kutz, M. Schorlemmer, E. Plaza, and K.-U. Kühnberger. A computational framework for conceptual blending. *Artificial Intelligence*, 256:105–129, 2018.
15. F. Baader, G. Brewka, and O. F. Gil. Adding threshold concepts to the description logic $\mathcal{EL}$. In C. Lutz and S. Ranise, editors, *Frontiers of Combining Systems*, pages 33–48, Cham, 2015. Springer International Publishing.
16. A. Ecke, R. Peñaloza, and A.-Y. Turhan. Similarity-based relaxed instance queries. *Journal of Applied Logic*, 13(4, Part 1):480–508, 2015.
17. R. Confalonieri, O. Kutz, P. Galliani, R. Peñaloza, D. Porello, M. Schorlemmer, and N. Troquard. Coherence, similarity, and concept generalisation. In *Proc. of DL*, volume 1879. CEUR-WS, 2017.
18. M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev. A Logic for Concepts and Similarity. *J. of Logic and Computation*, 17(3):415–452, 2007.
19. J. Hois and O. Kutz. Counterparts in Language and Space—Similarity and $\mathcal{S}$-Connection. In C. Eschenbach and M. Grüninger, editors, *Formal Ontology in Information Systems (FOIS 2008)*, pages 266–279. IOS Press, 2008.