Weighted Description Logic for Classification Problems[*]

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

[1] CORE Conceptual and Cognitive Modelling Group
KRDB Research Centre for Knowledge and Data
Free University of Bozen-Bolzano, Italy
`firstname.lastname@unibz.it`

[2] LOA Laboratory for Applied Ontology, ISTC-CNR
Trento, Italy
`daniele.porello@cnr.it`

Abstract

We summarise the definitions and some technical aspects of a familiy of description logics that introduce concept constructors (so-called tooth-operators) which, under various constraints, accumulate weights of concepts. We demonstrate how these operators can be fruitfully and elegantly applied to a number of cognitively motivated classification problems which are difficult to handle with the standard expressive means of description logics.

1 Introduction

Many aspect of how humans form and combine concepts are notoriously difficult to capture formally. In Description Logics (DLs), concepts are normally represented in terms of sets (extensions) and by means of definitions, such that, if an item satisfies the definition of a concept, it will be considered an instance of the concept's extension, otherwise it will not. Moreover, concept combination is usually intended in a compositional fashion, that is, the meaning of a complex concept functionally depends on the meaning of the primitive concepts (or predicates) that compose it. Unfortunately, empirical evidence in psychology and cognitive science have shown that most concepts often lack a precise definition, that categorisation presents a certain degree of indeterminacy, and that it is subject to context dependency. Moreover, a number of cognitive phenomena linked to concept combination are rarely taken into account in logic-based approaches. Overextension in conjunction [3], situational effects [4], dominance effects [5], and attributes emergence [6] are just a subset of the phenomena that hardly reconcile with compositionality of the extensions: the sole sets of entities that fall under two concepts, without additional information, do not determine the set of entities that fall under the combined concept. How to integrate these phenomena in a logic-based framework is still an open challenge. We here summarise a formal analysis of the work of Hampton [6, 5, 7, 3] as developed in [1], where an abstract setting to represent three particular concept formation aspects is given: **overextension**, when a concept classifies more than what is constrained

---

[*]This short paper summarises [1] and partly [2].

by its parts, **underextension**, when it classifies less, and **dominance**, when some parts of a concept dominate the classification question. Building on the DL $\mathcal{ALC}$, we introduce a family of operators which apply to finite lists of concept descriptions and return a composed concept whose instances are those that satisfy 'enough' of the listed concept descriptions. The motivation for this work is to bridge traditional techniques of knowledge representation with cognitively grounded views of human classification by contributing to the project of extending DLs to model cognitively relevant features of classification [2].

2   Weighted concept combination: The $\mathbb{W}$-operator logic

We introduce a class of m-ary operators, denoted by $\mathbb{W}$ (spoken 'tooth'), for combining concepts in $\mathcal{ALC}$. Each operator works as follow: 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 enough of a concept description, where 'enough' is modeled by means of a threshold. We assume a set of atomic concepts N_C and a set of atomic roles N_R. For $C, C_1, \ldots, C_m \in N_C$, $R \in N_R$, $t \in \mathbb{R}$, $m \geq 1$, $\vec{w} \in \mathbb{R}^m$, the set of $\mathcal{ALC}_\mathbb{W}$ concepts is described by the grammar:

$$C \quad ::= \quad A \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \forall R.C \mid \exists R.C \mid \mathbb{W}^t_{\vec{w}}(C_1, \ldots, C_m)$$

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 $\mathbb{W}$-concept $\mathtt{C} = \mathbb{W}^t_{\vec{w}}(C_1, \ldots, C_m)$ 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 $\mathbb{W}$-concept in $I = (\Delta^I, \cdot^I)$ is then:

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

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

Having defined the basic $\mathbb{W}$-operator, a number of derived operations are easily obtained. An initial technical study of $\mathcal{ALC}_\mathbb{W}$ was undertaken in [2], and the problem of 'knowledge dependence' was studied in [8]. For the discussion in this paper, it is of interest to consider the set of entities that *maximally* satisfy a combination of concepts $C_1, \ldots C_m$ of $\mathcal{ALC}$. The semantics of $\mathbb{W}^{max}$ is as follows:

$$\left(\mathbb{W}^{max}((C_1, w_1), \ldots (C_m, w_m))\right)^I = \{d \in \Delta \mid v^I_\mathtt{C}(d) \geq v^I_\mathtt{C}(d') \text{ for all } d' \in \Delta\} \tag{3}$$

Defining $\mathbb{W}^{max}$ in terms of $\mathbb{W}^t$ would require to use a *universal role*, which significantly increases the expressive power of $\mathcal{ALC}$ [9], as shown in [2].

3   Three Problems in Classification

3.1   Overextension

In the 80s, Hampton started a series of experiments to evaluate how people interpret conjunctive phrases such as *sports which are also games* or *games which are also sports* ([5]). The classical semantic interpretation of these cases would interpret their meaning in terms of conjunction: the extension of the two phrases would correspond to the intersection of the set of *games* and the set of *sports*. However,

experimental results showed something different. In his experiments, Hampton first provided a list of activities and asked people to evaluate which items in the list were *sports* and which were *games* (with the possibility of overlapping). After some weeks, people were asked again to evaluate which items in the list could have been considered *sports which are also games* and, conversely, *games which are also sports*, and the results were interpreted to verify if people followed a rule corresponding to an intersection or conjunctive interpretation. The answer is no: a number of *sports* (*games*) that were not considered *games* (*sports*), according to the first classification, were considered as members of the conjunction in the second part of the experiments. This phenomenon is called overextension in conjunction, but the problem of non-extensional combination applies equally to disjunctions (and negation).

The tooth operator can capture some of these insights. Firstly, we can lower the threshold for the classification under C or under D. That is, in the combination of C and D the importance of the list of features defining C and D is relaxed. So, more instances shall fall under either concepts, returning a larger intersection. Given $C = \mathbb{W}^s((P_1, w_1), \ldots, (P_m, w_m))$ and $D = \mathbb{W}^t((Q_1, v_1), \ldots, (Q_n, v_n))$, a family of *overextending compositions* of C and D can be introduced:

$$C \sqcap^{O} D = \mathbb{W}^x((P_1, w_1), \ldots, (P_m, w_m)) \sqcap \mathbb{W}^y((Q_1, v_1), \ldots, (Q_n, v_n)) \tag{4}$$

where $s \geq x$ or $t \geq y$ or both. The effect of lowering the threshold for classifying an entity under the composing concepts has the consequence of *raising* the importance of a (number of) features for the composed concept (whether it is their salience or their diagnosticity). Globally lowering the threshold amounts to assuming no information about which specific feature is important in the combination, whereas directly raising the weight of a number of features entails knowing that those are the important ones, when combining the concepts. Thus, a second strategy for combining concepts that overextends the intersection is by raising the weights of some of the features in C and D.

3.2 Underextension

Overextension's specular, or dual, phenomenon is that of underextension, that is, when considered in the context of conjunction, the phenomenon in which objects categorised as instances of two categories are not classified as belonging to the (underextending) conjunction of the categories. Actually, when related to underextension, Hampton's experiments were mostly directed to examine the disjunction of common sense concepts, and to verify the consistency of people's judgements with respect to the set union rule [7]. So, in that context, underextension would happen when an item categorised under a concept is not categorised as an item of the disjunction of that concept and another concept. For instance, an item could be categorised as a *game*, but then could be not considered as an instance of *game or sport* (maybe because it was a very atypical *game*). Here, we are dealing with another intuition. The idea is that underextension in conjunctions would mean to identify particularly typical members of the intersection of two concepts. As a particular case, we can then define a prototype concept of the regular intersection of the concepts. Similarly to overextension, we can treat the case of underextension, that is, by raising the threshold required for classifying the composing concepts. One interpretation of this case is that when combining two concepts, only instances that are more typical of the two concepts can enter the combination of the two concepts. By means of the $\mathbb{W}^{max}$, we can define underextension operators as follows. Given $C = \mathbb{W}^s((P_1, w_1), \ldots, (P_m, w_m))$ and $D = \mathbb{W}^t((Q_1, v_1), \ldots, (Q_n, v_n))$, we set:

$$C \sqcap^{U} D = \mathbb{W}^{max}((P_1, w_1), \ldots, (P_m, w_m)) \sqcap \mathbb{W}^{max}((Q_1, v_1), \ldots, (Q_n, v_n)) \tag{5}$$

Another way of introducing underextension is to constrain the intersection of C and D by imposing that the instances must maximise the combination of the features of C and D: $C \sqcap D \quad \sqcap \mathbb{W}^{max}((P_1, w_1), \ldots, (P_m, w_m), (Q_1, w_1), \ldots, (Q_n, w_n))$.

This definition is non-equivalent to (5) only in the case of mutually incompatible attributes defining C and D. We leave this aspect for future work.

3.3 Dominance

Hampton's experiments showed another interesting phenomenon, namely the dominance effect. In the simplest case of the composition of just two concepts, one of the two concepts 'weighs more', so to speak, than the other in the definition of the conjunction. More precisely, the dominance effect was analysed with respect to typicality. Hampton's aim was in fact to predict the typicality of an item in a conjunction of two classes on the basis of the typicality of the item in each of the conjuncts. In his analyses, he discovered that the weights of the two constituent concepts in the determination of an item's typicality were usually not equal. On the contrary, dominance of a concept with respect to a second was frequently observed. For instance, the weight of the concept *sports* was almost twice the weight of *games* in determining the typicality of items in both of the concepts *sports which are also games* and *games which are also sports* [6]. Following these insights, one can distinguish between a concept *Head* (*sports* in the example), and a concept *Modifier* (e.g. *games*).

Given $C = \mathbb{W}^s((P_1, w_1), \ldots, (P_m, w_m))$ and $D = \mathbb{W}^t((Q_1, v_1), \ldots, (Q_n, v_n))$, we present three strategies to model *dominance operators*. A first strategy is to define dominance as:

$$C \sqcap_1^D D = \mathbb{W}^{s'}\big((P_1, w_1), \ldots, (P_m, w_m)\big) \sqcap \mathbb{W}^{t'}\big((Q_1, v_1), \ldots, (Q_n, v_n)\big) \tag{6}$$

where $s' \geq s$ and $t' \leq t$. In this case, globally raising the threshold of the dominant concept amounts to include the typical instances of C into the composition, whereas globally lowering the threshold of the non-dominant concept includes possibly non-D instances. This means that a *sport that is also a game* must be primarily a *sport*.

A second strategy is to unpack (or flatten) the definitions of C and D and raise the weights of the most salient features $\varnothing \neq S(C) \subseteq \{1, \ldots, m\}$ of the dominant concept C for the composition, whilst including some required elements $R(D) \subseteq \{1, \ldots, n\}$ of D to compute a new minimal threshold. We may set:

$$C \sqcap_2^D D = \mathbb{W}^{\pi(C,D)}\big((P_1, w_1^\pi), \ldots, (P_m, w_m^\pi), (Q_1, v_1^\pi), \ldots, (Q_n, v_n^\pi)\big) \tag{7}$$

where π is a function that re-assigns weights such that $\pi(w_i) = w_i^\pi > w_i$ for all $i \in S(C)$ (and $\pi(x) = x$ otherwise) and $\pi(C, D) = \Sigma_{i \in S}\pi(w_i) + \Sigma_{j \in R}\pi(v_j)$ is a new threshold that accumulates the inflated weights for the salient features of C and the required weights from D.

A third strategy to define a dominant concept would be to collect those instances that maximise the sum of weights of the dominant concept (using e.g. the $\mathbb{W}^{max}$ operator of [2]).

Notice that these types of compositions are generally non-commutative, as the first combined concept (the head) is more important than the second.

4 Future Work

Future work will be dedicated to further investigate the logical properties of the operators and their natural extensions, and to employ them in a systematic fashion to the modelling of salient cognitive features of concepts, concept combinations in general, and conceptual blending in particular [10, 11].

Moreover, the landscape of over/under-extension and dominance should be analysed in a more fine-grained manner, considering interdependencies, special cases, and the introduction of further (conjunctive) classification problems. Eventually, we will aim to justify the formal fine-theory of conjuctive (and disjunctive etc.) classification in a detailed empirical study.

Further lines of research will deepen the comparison with the formal studies on typicality (e.g. [12, 13, 14]), the relation and combination with similarity frameworks based on a notion of distance (e.g. [15, 16, 17]), and the connections to linear classification models as begun in [8].

References

[1] G. Righetti, D. Porello, O. Kutz, N. Troquard, and C. Masolo. Pink panthers and toothless tigers: Three problems in classification. In *Proc. of the 5th International Workshop on Artificial Intelligence and Cognition*, Manchester, September 10–11, 2019.

[2] D. Porello, O. Kutz, G. Righetti, N. Troquard, P. Galliani, and C. Masolo. A Toothful of Concepts: Towards a theory of weighted concept combination. In *Proceedings of the 32th International Workshop on Description Logics (DL 2019)*, volume 2373. CEUR-WS, 2019.

[3] J. A. Hampton. Compositionality and concepts. In J. A. Hampton and Y. Winter, editors, *Compositionality and Concepts in Linguistics and Psychology*, pages 95–121. Springer International Publishing, Cham, 2017.

[4] 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.

[5] J. A. Hampton. Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(1):12–32, 1988.

[6] J. A. Hampton. Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, 15(1):55–71, 1987.

[7] J. A. Hampton. Disjunction of natural concepts. *Memory & Cognition*, 16(6):579–591, Nov 1988.

[8] P. Galliani, O. Kutz, D. Porello, G. Righetti, and N. Troquard. On Knowledge Dependence in Weighted Description Logic. In *Proc. of the 5th Global Conference on Artificial Intelligence (GCAI 2019)*, 17–19 September 2019, Bolzano, Italy, 2019.

[9] E. Hemaspaandra. The price of universality. *Notre Dame Journal of Formal Logic*, 37:174–203, 1996.

[10] 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.

[11] R. Confalonieri and O. Kutz. Blending under Deconstruction: The roles of logic, ontology, and cognition in computational concept invention. *Annals of Mathematics and Artificial Intelligence*, 2019.

[12] K. Britz, J. Heidema, and T. Meyer. Modelling object typicality in description logics. In *Australasian Joint Conference on Artificial Intelligence*, pages 506–516. Springer, 2009.

[13] L. Giordano, V. Gliozzi, N. Olivetti, and G. L. Pozzato. A non-monotonic description logic for reasoning about typicality. *Artificial Intelligence*, 195:165–202, 2013.

[14] 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.

[15] R. Confalonieri, O. Kutz, N. Troquard, P. Galliani, D. Porello, R. Peñaloza, and M. Schorlemmer. Coherence, Similarity, and Concept Generalisation. In R. K. Alessandro Artale, Birte Glimm, editor, *Proceedings of the 30th International Workshop on Description Logics (DL 2017)*, volume 1879. CEUR-WS, 2017.

[16] 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.

[17] 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.