In a Nutshell: Perceptron Connectives in Knowledge Representation (Extended Abstract)

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

[1] Free University of Bozen-Bolzano, Bolzano, Italy
[2] University of Genova, Genova, Italy.

Weighted Threshold Operators are n-ary logical operators which compute a weighted sum of their arguments and verify whether it reaches a certain threshold. These operators have been extensively studied in the context of circuit complexity theory (see e.g. [16]), and they are also known in the neural network community under the alternative name of *perceptrons* (see e.g. [6]).[1]

In [12], threshold operators were studied in the context of Knowledge Representation, focusing in particular on Description Logics (DLs) (see [5] for an introduction to DL). Adding threshold operators to DL is not hard. In brief, if $C_1 \ldots C_n$ are concept expressions, $w_1 \ldots w_n \in \mathbb{R}$ are weights, and $t \in \mathbb{R}$ is a threshold, we can introduce a new concept $\nabla\!\!\!\nabla^t(C_1 : w_1 \ldots C_n : w_n)$ to, semantically, designate those individuals d such that $\sum\{w_i : C_i \text{ applies to } d\} \geq t$. In the context of DL and concept representation, such threshold ("Tooth") expressions are natural and useful: they provide a simple way to describe classes of individuals that satisfy "enough" of a certain set of desiderata. In this brief abstract, we summarise the basic complexity and usability results obtained in [10].

Consider as an example the *Felony Score Sheet*[2] used in the State of Florida, in which various aspects of a crime are assigned points, and a threshold must be reached to decide compulsory imprisonment. For example, possession of cocaine corresponds to 16 points if it is the primary offense and to 2.4 points otherwise, a victim injury describable as "moderate" corresponds to 18 points, and a failure to appear for a criminal proceeding results in 4 points. Imprisonment is compulsory if the total is greater than 44 points and not compulsory otherwise. A knowledge base describing the laws of Florida would need to represent this score sheet as part of its definition of the **CompulsoryImprisonment** concept, e.g. as:

$$\nabla\!\!\!\nabla^{44}(\mathbf{CocainePrimary} : 16, \mathbf{ModerateInjuries} : 18, \ldots).$$

While it would be possible to also describe it (or any other Boolean function) in terms of more ordinary logical connectives (e.g. by a DNF expression), a definition in terms of Tooth expressions is far simpler and more readable. As such, the definition is more transparent and more explainable. We refer the

[1] Under the modern understanding of the term, a 'Perceptron' may have an activation function different from the Step Function.

[2] http://www.dc.state.fl.us/pub/scoresheet/cpc_manual.pdf (accessed: June 2021)

reader to [12,8] for an in-depth analysis of the properties of this operator. In these papers you can also find extensive discussions of related work such as [3,4,11,12]. A particularly interesting connection, not yet analysed in [10], is the embeddability (shown in [9]) of $\mathcal{ALC}^{\mathbb{W}}$ into the logic $\mathcal{ALCSCC}$ (see [2,1] for definitions) that allows the expression of rich cardinality constraints. Although the worst-case complexity of $\mathcal{ALCSCC}$ remains the same as $\mathcal{ALC}$, $\mathcal{ALC}^{\mathbb{W}}$ has the advantage that one can perform, via an efficient polynomial encoding into $\mathcal{ALC}$, practical reasoning by using available services.

Having Tooth expressions in a language of knowledge representation also has notable advantages from a cognitive point of view and from the practical point of view of knowledge acquisition. First, in psychology and cognitive science, the combination of two or more concepts has a more subtle semantics than set theoretic operations [15]. As shown in [14], Tooth operators can be used to represent these new concepts more faithfully regarding the way in which humans think of them and combine them. Second, as illustrated in [8], since a Tooth expression is simply a linear classification model, it is possible to use standard linear classification algorithms (such as the Perceptron Algorithm, Logistic Regression, or Linear SVM) to learn its weights and its threshold given a set of assertions about individuals (ABox). Two basic questions, however, need to be answered in order to assess the viability of this proposed addition to the language(s) of DL:

1. Given a DL $\mathcal{L}$, let $\mathcal{L}(\mathbb{W})$ be the logic obtained by adding threshold operators to it. How does the inference problem for $\mathcal{L}(\mathbb{W})$ compare to that for $\mathcal{L}$? More specifically: let $\mathcal{K}$ be a $\mathcal{L}(\mathbb{W})$-knowledge base and let ϕ be a $\mathcal{L}(\mathbb{W})$ axiom. Can we reduce the problem of whether $\mathcal{K} \models \phi$ (that is, of whether every interpretation that satisfies $\mathcal{K}$ satisfies ϕ) to the problem of whether $\mathcal{K}_0 \models \phi_0$ for some $\mathcal{K}_0, \phi_0 \in \mathcal{L}$ with an at-most-polynomial overhead?
2. Can we find examples in which simple threshold expressions can be used to express, more shortly and readably than (but roughly as accurately as) alternative approaches, non-trivial concepts derived from real data? If so, this would validate the claim that such expressions are well-suited for representing complex concepts in a readable way [12,14].

The answers obtained in [10] are summarised as follows. Firsty, the inference problem for $\mathcal{L}(\mathbb{W})$ is indeed not computationally harder than that of $\mathcal{L}$ (as long as $\mathcal{L}$ contains the Boolean connectives, e.g. $\mathcal{ALC}$). In a nutshell, this is shown by describing the digital circuits computing the sums and comparisons necessary for the evaluation of all Tooth expressions in terms of DL axioms (using new atomic concepts to describe the states of the internal nodes of the circuits). Secondly, we considered a small but non-trivial machine learning problem—attempting to learn the Molecular Function Gene Ontology [7] annotations of yeast proteins given their Cellular Component and Biological Process ones—and showed that even very simple threshold expressions can represent possible solutions that are roughly as accurate as the (much less easily interpretable) models learned by state-of-the-art learning algorithms. In our view, these two results strongly validate the usefulness of threshold expressions in the context of knowledge representation in general and DL specifically. In particular, it becomes straightforward to import concepts learnt from data into an existing ontology.

Future work on perceptron operators is currently being considered in a number of directions, including adding 'counting capabilities' to the language [9] as well as using perceptron operators to address compositionality and typicality effects in concept combination [13].

References

1. Baader, F.: A new description logic with set constraints and cardinality constraints on role successors. In: Dixon, C., Finger, M. (eds.) Frontiers of Combining Systems - 11th International Symposium, FroCoS 2017, Brasília, Brazil, September 27-29, 2017, Proceedings. LNCS, vol. 10483, pp. 43–59. Springer (2017)
2. Baader, F., Bednarczyk, B., Rudolph, S.: Satisfiability and query answering in description logics with global and local cardinality constraints. In: ECAI 2020 - 24th European Conference on Artificial Intelligence. Frontiers in Artificial Intelligence and Applications, vol. 325, pp. 616–623. IOS Press (2020)
3. Baader, F., Brewka, G., Gil, O.F.: Adding threshold concepts to the description logic $\mathcal{EL}$. In: Lutz, C., Ranise, S. (eds.) Frontiers of Combining Systems. pp. 33–48. Springer International Publishing, Cham (2015)
4. Baader, F., Ecke, A.: Reasoning with prototypes in the description logic $\mathcal{ALC}$ using weighted tree automata. In: Dediu, A.H., Janoušek, J., Martín-Vide, C., Truthe, B. (eds.) Language and Automata Theory and Applications. pp. 63–75. Springer International Publishing, Cham (2016)
5. Baader, F., Horrocks, I., Lutz, C., Sattler, U.: Introduction to Description Logic. Cambridge University Press (2017)
6. Bishop, C.M.: Pattern recognition and machine learning. Springer Science+ Business Media (2006)
7. Consortium, G.O.: The Gene Ontology (GO) database and informatics resource. Nucleic acids research **32**(suppl_1), D258–D261 (2004)
8. Galliani, P., Kutz, O., Porello, D., Righetti, G., Troquard, N.: On knowledge dependence in weighted description logic. In: Proc. of the 5th Global Conference on Artificial Intelligence (GCAI 2019). pp. 17–19 (2019)
9. Galliani, P., Kutz, O., Troquard, N.: Perceptron Operators That Count. In: Homola, M., Ryzhikov, V., Schmidt, R. (eds.) Proceedings of the 34th International Workshop on Description Logics (DL 2021). CEUR Workshop Proceedings, CEUR-WS.org (2021)
10. Galliani, P., Righetti, G., Kutz, O., Porello, D., Troquard, N.: Perceptron Connectives in Knowledge Representation. In: Keet, M., Dumontier, M. (eds.) Proceedings of 22nd International Conference on Knowledge Engineering and Knowledge Management (EKAW 2020). pp. 183–193. Springer (2020)
11. Murphy, G.: The Big Book of Concepts. MA: The MIT Press, Cambridge (2002)
12. Porello, D., Kutz, O., Righetti, G., Troquard, N., Galliani, P., Masolo, C.: A toothful of concepts: Towards a theory of weighted concept combination. In: Proceedings of the 32nd International Workshop on Description Logics. vol. 2373. CEUR-WS (2019), http://ceur-ws.org/Vol-2373/paper-24.pdf
13. Righetti, G., Masolo, C., Troquard, N., Kutz, O., Porello, D.: Concept Combination in Weighted Logic. In: Sanfilippo, E.M., Kutz, O., Hahmann, T., Hoehndorf,

R., Masolo, C., Vita, R., et al. (eds.) Proceedings of the Joint Ontology Workshops, CAOS workshop, co-located with the Bolzano Summer of Knowledge (BOSK 2021), Virtual & Bozen-Bolzano, Italy, September 10–18, 2021. CEUR Workshop Proceedings, CEUR-WS.org (2021)

14. Righetti, G., Porello, D., Kutz, O., Troquard, N., Masolo, C.: 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)
15. Righetti, G., Porello, D., Troquard, N., Kutz, O., Hedblom, M., Galliani, P.: Asymmetric Hybrids: Dialogues for Computational Concept Combination. In: Brodaric, B., Neuhaus, F. (eds.) 12th International Conference on Formal Ontology in Information Systems - FOIS 2021. Frontiers in Artificial Intelligence and Applications, IOS Press (2021)
16. Vollmer, H.: Introduction to circuit complexity: a uniform approach. Springer Science & Business Media (2013)