Concepts, Logic and Cognitive Adequacy

Supervisor: **Oliver Kutz**
Free University of Bozen-Bolzano

Guendalina Righetti
PhD student 34th cycle

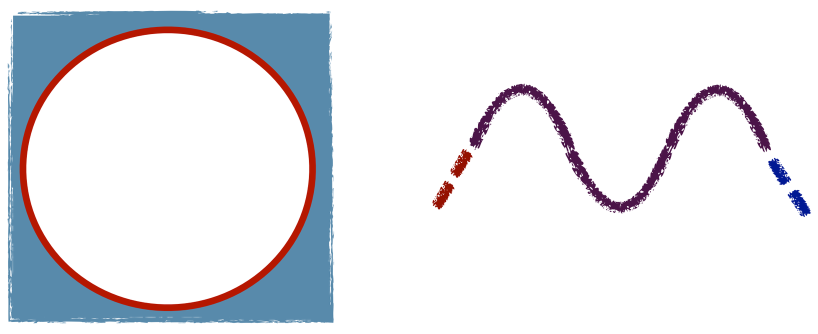
Co-supervisor: **Daniele Porello**
University of Genoa

Logic-based definitions: specify the extension of a concept by means of a logical (open) formula, in terms of necessary and sufficient conditions

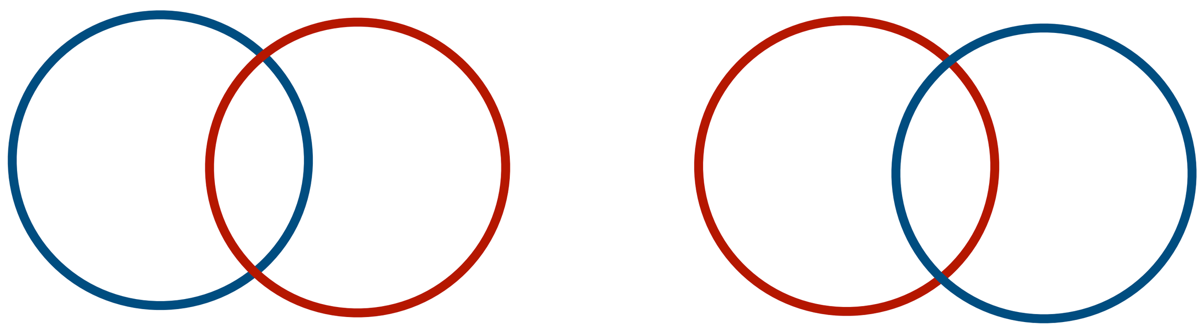
Logic ≀ Cognitive Science

Cognitive definitions: deviating from the logic-based view to provide adequate theories of human conceptualisation (e.g. Prototype Theory, Exemplar Theory, Theory Theory)

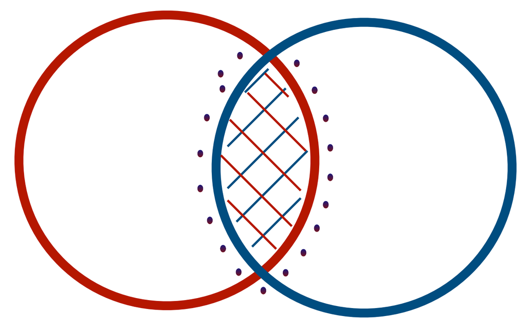
Typicality



Distinction between Head and Modifier



Overextension in conjunction



[1]

Weighted Description Logic [2, 3]

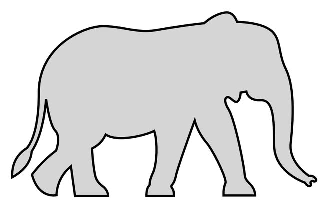
General idea, inspired by Prototype Theory: use weighted formulas, by introducing on top of $\mathcal{ALC}$ a class of n-ary operators $\nabla\!\!\!\nabla$ (spoken Tooth).

Each **Tooth Operator:**

1) Takes a list of concept descriptions **2)** Associates a weight to each of them

3) Returns a complex concept that applies to those instances that satisfy a combination of concepts, meeting a given threshold

$$\nabla\!\!\!\nabla^6((\text{Animal}, 5), (\text{Large}, 1), (\text{Heavy}, 1), (\exists \text{hasPart}\,.\,\text{Trunk}, 2), (\text{Grey}, 1))$$



$$(\nabla\!\!\!\nabla^t((C_1, w_1), \ldots, (C_m, w_m)))^I = \{d \in \Delta^I \mid v_c^I(d) \geq t\}$$

Where $v_c^I(d) = \sum_{i \in \{1,\ldots,m\}} = \{w_i \mid d \in C_i^I\}$

Interesting properties:

- there exists a link between tooth-expressions and linear classifiers: tooth-operators behave like perceptrons (a non-nested Tooth expression is a linear classification model)
- adding Tooth operators to any language, including the booleans, does not increase the expressivity and complexity of the language (Tooth expressions are equivalent to DNFs).

Combining weighted formulas [4,5]

How can we combine weighted formulas in a cognitively grounded way?

We provide an algorithm that exploits the power of logical inference but which is also inspired by the cognitive analysis on concept combination, and which makes use of the intensional definition offered by the Tooth operator.

To model Hampton's model of attribute inheritance [1], we assume a distinction between Head ($Fish$) and Modifier (Pet), and propose a 3 phases algorithm:

1. Features selection (step 1 and 2).
2. Weights assignment.
3. Threshold assignment.

A toy example:

$Fish \equiv \nabla\!\!\!\nabla^{10}((\forall breathesThrough\,.\,Gill, 3), (\forall\ livesIn\,.\,Water, 3), (\neg WarmBlooded, 3), (Grey, 0{,}9), (Scaly, 0{,}9), (\exists hasPart\,.\,Fin, 1))$

$Pet \equiv \nabla\!\!\!\nabla^{10}((CaredFor, 3), (\forall\ livesIn\,.\,House, 3), (Pretty, 3), (WarmBlooded, 0{,}9), (Furry, 0{,}9), (\exists provideCompanionshipTo\,.\,Human, 1))$

TBox:
A. $\forall livesIn\,.\,Water \sqcap \forall livesIn\,.\,House \sqsubseteq \forall livesIn\,.\,Aquarium$
B. $Furry \sqsubseteq WarmBlooded$ C. $Pretty \sqcap \forall breathesTrough\,.\,Gill \sqsubseteq \neg Grey$

Features selection:

Step 1: collects all the necessary features of the Head and all the *necessary* features of Modifier which are not impossible for the Head (i.e. the set $\bar{ft}(Modifier \circ Head)$).

$\bar{ft}(Pet \circ Fish) = \{\forall breathesThrough\,.\,Gill, \forall\ livesIn\,.\,Water, \neg WarmBlooded, CaredFor, \forall\ livesIn\,.\,House, Pretty\}$

Both according to the weights assignment and the TBox

Step 2: select the non necessary features excluding all the features of the Head (resp. Modifier), which are *impossible* for the concept Head (resp. Modifier) itself, once one adds all the necessary features of the Modifier (resp. Head) collected in $\bar{ft}(Pet \circ Fish)$.

$\text{ft}(Pet \circ Fish) = \{\forall breathesThrough\,.\,Gill, \forall\ livesIn\,.\,Water, CaredFor, \forall\ livesIn\,.\,House, Pretty, Scaly, \exists hasPart\,.\,Fin, \neg WarmBlooded, \exists provideCompanionshipTo\,.\,Human\}$

$Furry,\ WarmBlooded,\ Grey$ are excluded

Weights assignment: The weight associated to each feature of each component concept is divided by the score of the best exemplars of that concept.

Threshold assignment: The lower bound is the sum of the weights of the collected features minus the smaller weight of the (strongly) necessary features. The upper bound is the sum of the weights of the collected features, minus the bigger weight of the non necessary features.

$$\nabla\!\!\!\nabla^{1.49}((\forall breathesThrough\,.\,Gill, 0{,}25), (\forall\ livesIn\,.\,Water, 0{,}25), (\neg WarmBlooded, 0{,}25), (CaredFor, 0{,}25), (\forall\ livesIn\,.\,House, 0{,}25), (Pretty, 0{,}25), (Scaly, 0{,}07), (\exists hasPart\,.\,Fin, 0{,}08), (\exists provideCompanionshipTo\,.\,Human, 0{,}08))$$

$$(t \in (1.48, 1.65))$$

Evaluating Tooth interpretability [6, 7]

Are tooth-expressions more understandable than other equivalent formalisms?

We provide a user study to evaluate the interpretability of Tooth expressions and standard DNF.

$C_1 \equiv \nabla\!\!\!\nabla^1((A, -1), (B, 2), (C, 1))$

$DNF1$: $C_1 \equiv (\neg A \sqcap C) \sqcup B$

$DNF2$: $C_1 \equiv (A \sqcap B) \sqcup (\neg A \sqcap C) \sqcup (\neg A \sqcap B \sqcap \neg C)$

$DNF3$: $C_1 \equiv (A \sqcap B \sqcap C) \sqcup (\neg A \sqcap B \sqcap C) \sqcup (\neg A \sqcap B \sqcap \neg C) \sqcup (\neg A \sqcap \neg B \sqcap C) \sqcup (A \sqcap B \sqcap \neg C)$

We selected 24 concept definitions and designed two distinct questionnaires, one for the DNFs and one for Tooth expressions.

Task 1: Given the formula C_1, if an instance i is $\neg A$, B and $\neg C$, then i is an instance of C_1.

Task 2: Given the formula C_1, having B is necessary for being classified as C_1.

Subjects:
- 25 students with a background in philosophy
- 33 students with a background in computer science

	Task 1		Task 2	
Measures	Tooth	DNF	Tooth	DNF
%Correct Responses	0,90 (0,29)	0,91 (0,28)	0,83 (0,37)	0,87 (0,34)
Time (sec)	29,87 (20,72)	46,78 (58,90)	19,78 (19,78)	28,67 (28,78)
Confidence	5,65 (1,51)	5,74 (1,32)	5,82 (1,49)	5,70 (1,32)
Understandability	5,55 (1,44)	5,80 (1,24)	5,81 (1,43)	5,79 (1,24)

- The difference in accuracy (correct responses) is not statistically significant
- Subjects were faster when using the Tooth operator
- In the second task, subjects were more confident when using the Tooth operator, which was also perceived as more understandable

	Computer Science		Philosophy	
Measures	Tooth	DNF	Tooth	DNF
%Correct Responses	0,88 (0,31)	0,90 (0,32)	0,90 (0,34)	0,86 (0,30)
Time	25,23 (17,79)	37,29 (55,29)	24,80 (16,77)	36,39 (28,06)
Confidence	5,73 (1,71)	5,98 (1,29)	5,88 (1,15)	5,44 (1,28)
Understandability	5,61 (1,65)	6,11 (1,17)	5,84 (1,10)	5,43 (1,20)

- Tooth yielded faster responses in both groups
- Computer scientist were more confident with DNFs and perceived them as more understandable
- Philosopher found Tooth operators more understandable and were more confident with their answers when using Tooth operators

Tooth operators:
- are generally faster to use, leading to a lower time of response;
- are perceived as more understandable than DNFs in tasks that benefit from a more compact representation of knowledge;
- result in better performances (also in perceived understandability) for users who are not familiar with logic (philosophers).

References

1. J. Hampton: *Inheritance of attributes in natural concept conjunctions*. Memory & Cognition. 1987 Jan;15(1):55-71.
2. D. Porello, O. Kutz, G. Righetti, N. Troquard, P. Galliani, C. Masolo: *A Toothful of Concepts: Towards a Theory of Weighted Concept Combination.* DL 2019.
3. P. Galliani, G. Righetti, O. Kutz, D. Porello and N. Troquard: *Perceptron Connectives in Knowledge Representation*. EKAW 2020: 183-193
4. G. Righetti, D. Porello, O. Kutz, N. Troquard, C. Masolo: *Pink panthers and toothless tigers: three problems in classification.* AIC 2019: 39-53.
5. G. Righetti, C. Masolo , D. Porello, O. Kutz, N. Troquard: *Combining concepts in Weighted Logic*. CAOS 2021
6. G. Righetti, D. Porello and R. Confalonieri. Evaluating the Interpretability of Tooth Expressions (Extended Abstract). DL 2022.
7. G. Righetti, D. Porello and R. Confalonieri. Evaluating the Interpretability of Threshold Operators. EKAW 2022.

guendalina.righetti@stud-inf.unibz.it

unibz