The Moving Apple

An image-schematic investigation into the Leuven Concept Database

Guendalina Righetti[1], Oliver Kutz[1]

[1]KRDB Research Centre for Knowledge and Data, Free University of Bozen-Bolzano, Italy

Abstract

Image schemas have been recognized as fundamental cognitive structures that precede language and serve as foundational elements for conceptualization. Extensive research has investigated their involvement in metaphor construction, analogy formation, and conceptual blending. This paper initiates the project of analysing their role in the representation of concepts provided by Prototype Theory. To tackle this problem, we analyse the Leuven Concept Database which is collecting (common sense) information over the features exhibited by 15 concepts. Specifically, our focus lies in investigating the image schematic content embedded within these features and exploring how this content varies across different conceptual domains.

Keywords

Image Schema, Prototype Theory, Feature, Leuven Concept Database

1. Introduction

In Knowledge Representation (KR), concepts are often assumed as purely extensional, namely as the set of entities they identify. Experimental findings in Cognitive and Experimental Psychology have long since demonstrated that concepts have much more subtle semantics than what is expressible by simply assuming a purely extensional view of concepts.

Different, alternative, cognitive theories of concepts have been proposed in the literature to provide a working definition of what concepts are and how they are represented in the human mind. Among these theories, the Prototype theory is one of the most formally developed [1, 2]. Maybe for this reason, it is also the one from which many formal systems take their inspiration when trying to provide more adequate modelling of human concepts in Logic and KR [3, 4, 5, 6, 7, 8, 9].

The Prototype theory has its roots in Wittgenstein's idea of family resemblance (1953) and the experimental evidence provided by Eleanor Rosch (1975, 1978). Prototype Theory has sometimes been read as a "single best example" model: that is, as proposing that each category could be represented as its best, or ideal, member, which reifies all the characteristics commonly associated with the category. More frequently, Prototype Theory is understood by its proponents as a *summary representation* [11], that is, a unified representation that describes the category as a whole.

The Seventh Image Schema Day (ISD7), September 2, 2023, Rhodes, Greece
guendalina.righetti@unibz.it (G. Righetti); oliver.kutz@unibz.it (O. Kutz)

One of its central notions is then the one of *feature*. While the notion of concept is well discussed in the literature, both philosophical and psychological [12], the notion of *feature*, is much less developed and possibly even more controversial. It is mostly taken for granted in many cognitive theories on concepts, so an operational definition of what constitutes a feature for a concept, and what is the concrete difference between features and concepts is lacking in the literature.

Notwithstanding this, Prototype Theory suggests representing concepts as a list of features, meant as the characteristics, attributes, or properties usually found across the instances of the concepts. For instance, a *Pet* may be described as an animal, furry, having an owner: all of these descriptions may be considered features of the concept *Pet*. Moreover, according to the Prototype Theory, these features have different 'importance' in the description of the concept. The importance is often expressed through 'weights' (i.e. numbers) these features are associated with. For example, the concept *Pet* could associate a higher importance weight with the feature "having fur" and a lower weight with "having scales". Nevertheless, the list of features can include (intuitively) contradictory features, allowing in this way the representation of the variability of the category (let us consider again the iguana pet example). Similarly, this explains many of the so-called *typicality effect*: exemplars sharing more features with the prototype would be more typical, and exemplars sharing a few numbers of features would instead be atypical.

While the Prototype view is widely accepted as a prominent theory for representing concepts, it is not without its share of criticisms, also because of the naive use of the notion of feature.

According to Lakoff [13, p. 133] "[...] the properties that are relevant for the characterization of human categories are not objectively existing properties that are 'out there' in the world. Rather they are 'interactional properties,' what we understand as properties by virtue of our interactive functioning in our environment." Moreover, (ibid. p.135) "Prototype effects are real, but superficial. They may arise from a variety of sources. It is important not to confuse prototype effects with the structure of the category as given by cognitive models."

Proponents of the Prototype theory argue, contra this kind of critique, that it is important to distinguish between the theory itself and the operational methodology employed to support it [14]. Operationally, the prototype representation is assessed by utilising experimental psychology's methods, and more precisely through *feature generation tasks*: in a controlled experiment, people have to list all the features a category shows and that are common to the category as a whole. These features can then be collected and evaluated according to their importance (in a way that will be made clear below, see Section 2.1) to produce a prototype representation of the concept of interest.

However, according to Hampton: "It is not supposed for example that the mind contains a list of attributes in the verbal form in which they are generated. Clearly meaning has to be grounded [...] in experiential sub-symbolic levels of cognition, so it is unhelpful for a psychological model to give the meaning of one word simply in terms of others unless there is a primitive base of terms that are defined non-verbally."[14].

This is nothing but another variant of the Symbol Grounding Problem, which in AI refers to the issue of connecting symbols, such as words or abstract representations, with their corresponding meanings or referents in the real world. In the context of prototypes' representation, the problem lies in bridging the gap between the symbolic representations (reified in the list of features which describe the concepts) and the actual sensory experiences or physical entities they are meant to represent. Hampton's reference to "experiential sub-symbolic levels of cognition" points here to embodied experience as a lower-level source of the representation which is presupposed by Prototype Theory (and mainstream cognitive theories of concepts in general).

Image Schemas refer to recurring patterns of perceptual and embodied experiences that shape our understanding of the world. As such, they are often been pointed to as a possible way of solving the Symbol Grounding Problem [15]. An interesting question is whether we can find evidence of image-schematic information coded into the features that are normally employed in prototype representations of concepts.

To start tackling this question, in this paper we analyse the content of the Leuven Concept Database [16]. The Leuven concept database (LCD) is a repository of information curated by a team of psychologists from the University of Leuven. It collects the features associated with 15 concepts with the purpose of providing evidence on how individuals perceive and categorize the world around them. The goal of our analysis will be to study whether the features gathered into the database presuppose or appeal to image-schematic notions, and, when the answer is yes, to study which Image Schemas are mostly involved. As the LCD covers different ontological domains (artefacts, animals, activities), we will also analyse whether there is a different distribution of such conceptual primitives across the different domains.

While there exist works aiming at automatically extracting Image Schemas from texts [17, 18, 19], we will carry out this preliminary analysis manually in this paper.

2. Background

2.1. The Leuven Concept Database

The Leuven Concept Database (LCD) is a large dataset that links sets of features to concepts (or category labels) and exemplars (or lexical entries). It was compiled by the ConCat group at the University of Leuven between 2004 and 2008. This database consists of 15 categories (with a total of 420 associated exemplars) and 500 associated features. Specifically, it covers the animal domain: birds, fish, insects, mammals, reptiles, and amphibians. It also includes information on the artefact domain: musical instruments, tools, vehicles, clothing, kitchen utensils, and weapons. Additionally, there are categories for fruits and vegetables, as well as activities, such as professions and sports. The data collection involved the participation of at least a thousand students.

Although the material was collected in Dutch, an English translation is provided to facilitate further experimental and modelling approaches.

The studies conducted at the University of Leuven contribute to the ongoing debate between the Prototype Theory [1] and the Exemplar Theory [20]. These studies involve a series of experiments that aim to investigate various aspects of each theory. Specifically, we are interested in the studies related to a feature-generation task, in which participants were asked to generate lists of features associated with the 15 category labels mentioned earlier.

The participants' responses to the feature generation task were manually compiled and adjusted with minimal stemming into 15 distinct Excel tables (one for each category). The frequency of feature production was recorded as an indirect measure of their importance.

Moreover, participants were directly asked to rate the importance of each feature in defining the concept it was associated with, on a rating scale ranging from +3 (very important feature) to -3 (very unimportant feature). The ratings assigned to each feature by the participants of the experiments were collected in each table of the database, leading to a representation of concepts very close to the one endorsed by proponents of the Prototype Theory.

2.2. Previous Work: Ontological Analysis of the Leuven Concept Database

In previous work [21], we presented an in-depth case study of the Leuven concept database (LCD) to transform the common-sense knowledge therein into a format that could be effectively utilised for practical applications. More specifically, the goal of the paper was to translate the LCD into OWL (the most widespread language for authoring ontologies [22]) and to use the ontologies built upon it in a concrete computational implementation for concept combination [23, 24]. To achieve this level of formalisation, we explored a hybrid approach that involved analyzing the syntactic structure of the LCD entries along with conducting semantic and ontological analyses. More specifically, we employed notions from Foundational Ontology (FO): foundational or upper ontologies establish the precise definitions of highly general terms, such as object, event, property, quality, relation, process, etc. FO define the thus fundamental categories that are applicable across multiple domains and are inherently embedded within common sense knowledge. These high-level, ontological distinctions, were used, in our case study, both to disambiguate the meaning of some of the features described in the LCD and to inform some of the formalisation choices in the process of translation into OWL.

In the above-mentioned paper [21], we individuated 7 *modes of disambiguation* which could guide our translation of the LCD: the distinction between Rigid and Non-Rigid properties [25], capturing the difference between properties which are always inherently true of one class, and properties which may change over time; Mereology, namely the theory of part-whole relations; Constitution, capturing the features expressing information on material substances; Quality, helping the formalisation of quality spaces such as size, shape, colour, etc; Action and Ability, ascribing abilities to agents; Functionality, often employed in the artefact domain to capture objects' purposes and affordances; and finally, we took into account other facets of core knowledge that are grounded in cognition, such as spatiotemporal relations. Among them, we also discussed the Image Schema CONTAINMENT, but we did not carry out a thorough analysis of the other Image Schemas involved in the content of the Database.

In this paper, we aim to go a step further and analyse the features contained in the LCD by studying the image-schematic information they codify. As we already carried out an ontological analysis of the LCD, this will also allow for some consideration of the relationship between Image Schemas and ontological distinctions previously identified.

3. Unveiling the role of Image Schemas in the Leuven Concept Database

Image schemas are primitive, cognitive structures that arise from our interactions with the world, shaping our understanding of reality since childhood. Image schemas are derived from our sensory and motor interactions and provide the building blocks for abstract reasoning and conceptual knowledge. These recurring patterns of perception and embodied experiences are thus foundational elements for conceptualization. Despite a long line of research trying to give a comprehensive classification, there is still no common agreement on the list of Image Schemas [26]. In this paper, we manually analyse the content of the LCD by considering the list of image schemas proposed by Hedblom [15] (CONTAINMENT, SOURCE_PATH_GOAL, CYCLE, CONTACT, SUPPORT, LINK, VERTICALITY, SCALING, NEAR_FAR, BLOCKAGE, CAUSED_MOVEMENT, SELF_MOVEMENT, ATTRACTION). In addition to this list, we also included the PART_WHOLE and the OBJECT Image Schemas.

3.1. Analysis: Criteria and Overview

As mentioned above, the LCD contains 500 features, distributed among 15 concepts.

Nevertheless, some features generated during the experiments posed challenges in their interpretation within the context of defining the concept. Some of the features were either blatantly false when applied to the entire class or connected to a semantic context completely unrelated to the experiment's intended purpose (e.g., a *Fish* was described as "a constellation"). Furthermore, certain features captured language and societal biases, such as a *Profession* being perceived as "different for men and women," and a *Kitchen Utensil* being labelled as "especially used by women," while a *Tool* was considered "primarily used by men."

As previously mentioned, all generated features underwent an evaluation to assess their relevance to the respective class (see again Section 2.1). Following what was done in [21], and inspired by the concept of *salience* in Prototype theory, to exclude possibly controversial features, the means of subjects' judgments were calculated. Entries falling strictly below the threshold of 0 were excluded. This evaluation procedure resulted in the exclusion of 103 features, (accounting for approximately 20% of the total features).

As a result, we retained a total of 397 features that spanned 15 different categories. We proceeded to manually annotate all of these features, identifying and documenting any Image Schemas from the previously mentioned list we were able to identify as hidden in the expression of the feature. Figure 3.1 shows an example for the concept Bird.

Out of these 397 features, we successfully categorized 228 features based on image-schematic information. In other words, approximately 40% of the features defied classification within the chosen Image Schemas.

This percentage is however not equally distributed across the different entries of the database. As already mentioned, we can distinguish between 4 macro-areas, or domains, in the database: the Animal domain (including Bird, Mammal, Fish, Insect, and Reptile, for a total of 141 features); the Artefact domain (including Vehicle, Kitchen Utensil, Clothing, Tool, Musical Instrument, Weapon, for a total of 155 features); the Plant domain (including Vegetable, Fruit, for a total 53 features); Activity domain (including Profession, Sport, for a total of 48 features).

In the case of the Animal domain, we were able to categorise into image-schematic terms 72% of the features. For the Artefact domain, the percentage is lower, and it goes to 54%, and, similarly in the Plant domain is 58%. The most difficult domain resulted to be that of Activities, where the percentage of classified features drop to 27%.

DUTCH	ENGLISH	ImageSchema													MEAN	PF
heeft veren	has feathers	PART-WHOLE	3	3	2	2	3	3	3	3	3	3	3	3	2,833	20
heeft een snavel	has a bill	PART-WHOLE	3	3	2	3	3	3	3	3	3	3	3	1	2,75	15
heeft vleugels	has wings	PART-WHOLE	3	3	2	2	3	2	3	3	3	3	3	3	2,75	9
kan vliegen	can fly	SELF-MOVEMENT	2	3	2	2	3	3	3	3	3	3	3	3	2,75	20
bouwt nesten	builds nests	SOURCE-PATH-GOAL?/SUPPORT?CONTAINMENT?	3	3	3	2	2	2	2	3	3	3	3	3	2,666	16
heeft twee vleugels	has two wings	PART-WHOLE	3	3	3	3	3	3	0	3	3	3	3	2	2,666	3
legt eieren	lays eggs	CYCLE?	3	3	2	2	3	1	3	3	3	3	3	2	2,583	20
			-	-	-	-	-	-	-	-	-	-	-	-	-	
heeft twee poten	has two paws	PART-WHOLE	3	3	3	2	3	0	3	3	3	3	3	1	2,5	3
heeft een bek	has a beak	PART-WHOLE	3	3	1	3	3	3	2	2	2	3	3	2	2,5	6
is een dier	is an animal		3	3	3	2	3	-1	3	3	3	3	3	1	2,416	3
fladdert	flutters	SELF-MOVEMENT	2	2	2	3	2	2	3	3	3	2	-1	3	2,166	2
eet wormen	eats worms	CONTAINMENT? SELF-MOVEMENT	3	2	2	2	3	3	-2	0	2	2	2	1	1,666	8
eet kleine dieren	eats small animals	CONTAINMENT? SELF-MOVEMENT	2	1	2	0	3	1	2	1	2	1	1	0	1,333	2
fluit	sings (whistles)	SELF-MOVEMENT? SOURCE-PATH-GOAL?	2	2	1	2	1	-1	-3	3	1	3	2	3	1,333	7
tsjilpt	chirps	SELF-MOVEMENT? SOURCE-PATH-GOAL?	1	2	1	1	2	3	-3	1	2	2	2	2	1,333	6
leeft in het wild	lives in the wild	IN/SELF-MOVEMENT	2	2	0	1	1	-2	-3	3	2	3	3	1	1,083	3
is een trekvogel	is a migratory bird	SELF-MOVEMENT/CYCLE/SOURCE-PATH-GOAL	2	1	2	2	-1	3	-3	3	2	1	1	-1	1	4
vind je in bomen	can be found in trees	IN	1	2	1	2	3	1	0	2	1	2	-2	-1	1	5
heeft poten	has legs	PART-WHOLE	0	1	-3	2	3	-1	1	1	1	1	1	2	0,75	3
in kooi	lives in a cage	IN	1	2	1	2	1	1	-3	0	1	2	1	-1	0,666	2
heeft luchtzakken	has air sacs	PART-WHOLE	-3	2	-3	1	1	1	3	-3	3	3	2	1	0,666	3
kraaloogjes	has beady eyes	PART-WHOLE	3	3	1	3	1	-2	-3	-3	1	2	2	-2	0,5	2
zingt	sings	-	1	1	1	2	2	-3	-3	-3	1	3	1	2	0,416	3
kan lopen	can walk	SELF-MOVEMENT	-1	1	-2	-2	3	0	-1	-1	0	1	1	1	0	2
kraait 's morgens	in the morning		1	1	-2	2	-2	3	-3	1	1	1	-3	0	0	2
kan zich voortplanten	is able to reproduce		-3	1	-2	-1	3	0	0	1	-3	-1	-3	2	-0,5	4
wordt soms als huisdi	is sometimes kept as a pet		-3	0	-1	1	1	-1	0	-2	-3	-2	-2	-2	-1,16	2
wordt soms door mer	is sometimes eaten by man		-3	0	-1	1	0	0	0	-3	-3	0	-3	-3	-1,25	2

Figure 1: Example, for the concept Bird, of the annotation carried out in our analysis.

3.2. Detailed Analysis

We can now consider how the selected Image Schemas are distributed across the Database. Clearly, as complex features are involved in the description of the concepts, often the classification did not satisfy a unique Image Schema, but rather a *collection* (or profile) of Image Schemas [27]. We will present the percentages for each Image Schema individually in the following.

1. CONTAINMENT represents the connection between the interior and exterior of an entity, along with the boundary that separates them [28]. In our analysis we counted as CONTAINMENT also its dynamic counterparts IN and OUT. Based on our analysis, we found that approximately 14% of the features in the Database can be categorized as involving a CONTAINMENT relationship. Breaking it down by domains, there were 24 out of 141 features in the Animal domain (e.g. Birds 'eat worms', Fishes 'contain omega 3'), 16 out of 155 in the Artefact domain (e.g. Kitchen Utensils are 'used to store food', Clothes are 'found in a clothing cupboard', Weapons 'contain bullets'), 12 out of 53 in the Plant domain (e.g. Fruit 'contains seeds'), and 4 out of 48 in the Activity domain (e.g. Sport 'is practised indoor') that were classified based on this primitive.
2. SOURCE_PATH_GOAL pertains to the movement from a source and directed towards a goal. Around 11% of the features in the database were classified as implying it, mostly in the Artefact domain: 30 features out of 155 involved SOURCE_PATH_GOAL (many of them in the Vehicle category: e.g. with a Vehicle 'you can transport things', or a Vehicle 'can be used to get from one place to another'). In the other domains, the number of

features captured by this Image Schema drops: only 4 out of 141 features in the Animal domain (e.g. Reptiles ‘descend from dinosaurs); 5 out of 53 in the Plant domain (Fruit is ‘used for making fruit salad’); 7 out of 48 in the Activities domain (mostly taking a metaphorical reading), resulting in being the most frequent Image Schema identified in this domain (e.g. Sport ‘makes you sweat’, a Profession ‘earns you money’).

3. CYCLE is the Image Schema of recurring patterns (such as seasons or daily cycles). For this reason, it is maybe not surprising that it was found to be especially frequent in the Plant domain (13 out of 53 features, in cases such as Fruit ‘is seasonal’ and ’rots after a while’, or Vegetables ‘grow in the garden’). Globally, around 9% of the features were classified as involving a CYCLE.
4. CONTACT in its easiest variant refers to physical contact between two objects, but may also involve abstract contact among entities [15]. Only around 4% of the features in the Database involved a CONTACT relationship (e.g. Clothing ‘is used to cover yourself’, a Mammal ‘breastfeeds its babies’, Vegetables ‘grow in the ground’), and none of them was in the Activity domain.
5. SUPPORT, the Image Schema of physical upholding, follows a similar pattern to that of CONTACT, as these two Image Schemas were frequently grouped together in classification. In fact, it was found in around 4% of the features of the Database (e.g. Reptiles ‘live on land’, Musical Instruments ‘can be played on’, Fruit ‘grows on Tree’, in Sports, ‘supporters encourage the players’).
6. LINK comprises at least two entities that are connected either physically or metaphorically, with a particular attention on the bond holding between them. We identified this Image Schema only 3 times in the database: Mammals ‘breastfeed [their] babies’, a Tool is ’is used to make something with’ (there is then a link between it what you produce with it), and your Profession ‘has strong influence upon the rest of your life’.
7. SCALING involves the variation in object size, encompassing the dynamic transformation of objects as they either grow or shrink in size. It was found in around 6% of features, and it was especially frequent in the Plant domain, where many features involved the growing of Fruits and Vegetables.
8. BLOCKAGE, the Image Schema where a force is hindered or redirected by an obstacle (whether in a physical or metaphorical sense), was identified only 6 times in the LCD, and all of them in the Artefact domain. More specifically, it was found 5 times in the concept Clothing: e.g. clothes ‘protects against the cold’; and once in the category of Weapons, which are ‘used to protect you’.
9. CAUSED_MOVEMENT is the image schema which describes the transfer of movement from one object into another. Its presence in the LCD was again quite limited, as it classified only 6 of the features. Not surprisingly, it was involved in the description of the concept ’Vehicle’. Similarly, Insects are said to ‘carry over diseases’, and, more metaphorically, Fruit ‘needs a lot of sun to grow’.
10. SELF_MOVEMENT, the Image Schema related to the notion of agency, classified almost 6% of the features, rather obviously only in the Animal (Birds ‘flutter’, Fishes ‘can swim’, etc) and the Plant (Fruit ’grows on trees’, etc.) domains. Interestingly, around 15 features in the Artefact domain seem to imply the negation of the SELF_MOVEMENT schema (e.g. a Vehicle ‘is operated’).

11. PART_WHOLE is the image schema involving mereological relationships. As such, it involved almost 15% of the features, across the domains of Animal, Plant and Artefact. More precisely, we classified as PART_WHOLE 37 out of 141 features in the Animal domain (e.g. Insects 'has an exoskeleton', Reptiles 'have scales'), 17 out of 155 in the Artefact domain (e.g. a Musical Instrument 'has buttons'), 4 out of 53 in the Plant domain (e.g. Fruit 'contains seeds').
12. OBJECT is a more controversial image schema, which is sometimes considered to be only subsidiary to other primitives (such as CONTAINMENT or SUPPORT) [29]. We identified it in 5 of the features of the LCD, maybe not surprisingly all of them in the Artefact domain (e.g. a Weapon is a 'Tool').

We did not find incarnations of VERTICALITY, NEAR_FAR, and ATTRACTION in the Leuven Concept Database, although a number of features may imply some of them (see below).

3.3. Discussion

The analysis presented above has shown that the role of Image Schemas varies across the different domains. For instance, we have noticed how their role is predominant in describing the Animal domain, whereas most of the features describing the Activities escaped a classification in terms of image schematic information. There exist different possible explanations for this. First, the Activities domain (hence the concepts of Profession and Sport) is arguably more complex to describe than the Animal one, as it is more difficult to instantiate at the basic level the involved categories, and this leads to the use of more complex, or at least more inventive, features in describing them. Identifying the Image Schemas involved in such descriptions may thus require a more in-depth metaphorical analysis of the involved terminology. Second, and aside from these practical difficulties, this effect may also reflect the different amount of background, contextual and extra-bodily knowledge which is employed when describing concepts such as Sports and Professions (e.g. a Profession 'requires a certain education', or a Sport 'requires motivation').

In other cases, the features employed to describe the concepts imply the reference to embodied experience but are of difficult classification with the list of image schema identified for the task. Examples are features referring to sounds (Vehicles 'produce noise'), colour (Vegetables 'are green'), or even emotions (Sport 'is fun', or Mosquitos 'make an extremely irritating sound'). Although, in many cases, there is not an obvious direct connection to primitive spatial structures, there is however space for a less straightforward reading of some of the selected Image Schemas, which may allow for an interpretation of some of the above-mentioned features [26]. VERTICALITY, for instance, can allow a metaphorical interpretation of emotions: good emotions are UP in the scale, while bad emotions drag you DOWN [30].

Another interesting point put in evidence by our analysis is the possibility to identify *negation* of Image Schemas among the features of a concept. An example is the negation of the Image Schema SELF-MOVEMENT in cases such as Vehicles 'are operated' or Mammals 'can't fly' (poor bats). The possibility of considering *negation* of Image Schemas paves the way for interesting lines of research, although may sound controversial, as the negation is not easy to identify as a conceptual primitive.

In conclusion, it is noteworthy to establish a connection between the current study and the ontological analysis conducted in prior research (Righetti, 2022). Several Image Schemas chosen for our investigation align closely with the ontological "modes of disambiguation" proposed in the same work (Righetti, 2022). For instance, the PART_WHOLE schema is easily associated with the mode of 'Mereology', providing a clear and evident link. Additionally, the relationship between the 'Functionality' mode and the SOURCE_PATH_GOAL schema might be less apparent, but they often run parallel in our analysis of the Leuven Concept Database. A different story is about the mode 'Quality', encompassing colour, taste and sounds in our ontological analysis, for which was often much more difficult (if not impossible) to identify a corresponding Image Schema.

4. Conclusion

In conclusion, this paper aimed to address the question of whether the features contained in the Leuven Concept Database (LCD) are associated with image-schematic notions. The purpose was to start tackling the question of whether and in what way we can ground the notion of feature, as used in Prototype Theory, into bodily experience and conceptual primitives such as Image Schemas. One immediate application is to have a better handle on which kinds of features are essential in formal approaches to giving prototype definitions, as e.g. [31], and specifically, what is the contribution of image-schematic features in such definitions.

According to our analysis, around 60% of the features gathered in the database can be described in terms of image-schematic information. Moreover, a number of additional features which escaped our classification was still traceable to bodily experiences (e.g. emotions, sound experiences, etc).

Although there is empirical evidence of high agreement when people have to identify image schematic information within text or images [32], clearly this is only a preliminary analysis which will require further study and validation. Consider the feature: Fruit 'grows on trees'. How many Image Schemas does it imply? How many would agree on the fact that the *growing* of an apple would imply the SCALING schema? Possibly many. How many would agree on the fact that *growing* implies the SELF_MOVEMENT schema? Possibly less. The features contained in the LCD can be subject to various interpretations, and the level of strictness in these interpretations may yield slightly different outcomes wrt our analysis. Conducting a more thorough analysis that addresses these concerns is an avenue for future research.

References

[1] E. Rosch, Cognitive representations of semantic categories, Journal of Experimental Psychology:General 104 (1975) 192–233.

[2] E. Rosch, Principles of categorization, Rosch, E., Lloyd, B. B.. Hillsdale, NJ: Erlbaum., 1978, pp. 27–48.

[3] F. Baader, A. Ecke, Reasoning with prototypes in the description logic alc alc using weighted tree automata, in: Language and Automata Theory and Applications: 10th International

Conference, LATA 2016, Prague, Czech Republic, March 14-18, 2016, Proceedings 10, Springer, 2016, pp. 63–75.
[4] D. Porello, O. Kutz, G. Righetti, N. Troquard, P. Galliani, C. Masolo, A toothful of concepts: Towards a theory of weighted concept combination, in: M. Simkus, G. E. Weddell (Eds.), Proceedings of the 32nd International Workshop on Description Logics, Oslo, Norway, June 18-21, 2019, volume 2373 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019. URL: http://ceur-ws.org/Vol-2373/paper-24.pdf.
[5] G. Righetti, D. Porello, O. Kutz, N. Troquard, C. Masolo, Pink panthers and toothless tigers: three problems in classification, in: A. Cangelosi, A. Lieto (Eds.), Proceedings of the 7th International Workshop on Artificial Intelligence and Cognition, Manchester, UK, September 10-11, 2019, volume 2483 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 39–53. URL: http://ceur-ws.org/Vol-2483/paper4.pdf.
[6] G. Righetti, P. Galliani, C. Masolo, Concept Combination in Weighted DL, in: S. A. Gaggl, M. V. Martinez, M. Ortiz (Eds.), Logics in Artificial Intelligence - 18th European Conference, JELIA 2023, Dresden, September 20-22, 2023, Proceedings, volume 14281 of *Lecture Notes in Computer Science*, Springer, 2023.
[7] P. Galliani, O. Kutz, N. Troquard, Succinctness and complexity of ALC with counting perceptrons, in: P. Marquis, T. C. Son, G. Kern-Isberner (Eds.), Proceedings of the 20th International Conference on Principles of Knowledge Representation and Reasoning, KR 2023, Rhodes, Greece, September 2-8, 2023, 2023, pp. 291–300. URL: https://doi.org/10.24963/kr.2023/29. doi:10.24963/kr.2023/29.
[8] A. Lieto, 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, 2018, pp. 189–199.
[9] K. Britz, J. Heidema, T. A. Meyer, Semantic preferential subsumption, in: KR, AAAI Press, 2008, pp. 476–484.
[10] L. Wittgenstein, Philosophical Investigations, Oxford: Blackwell, 1953.
[11] G. L. Murphy, The Big Book of Concepts, Cambridge, MA: The MIT Press, 2002.
[12] E. Margolis, S. Laurence, Concepts (2014). URL: https://plato.stanford.edu/archives/spr2014/entries/concepts/.
[13] G. Lakoff, Cognitive models and prototype theory, The cognitive linguistics reader (2007) 130–167.
[14] J. A. Hampton, The Routledge Handbook of Semantics, Abingdon, UK: Routledge, 2015, pp. 125–141.
[15] M. M. Hedblom, Image schemas and concept invention: cognitive, logical, and linguistic investigations, Springer Nature, 2020.
[16] S. De Deyne, S. Verheyen, E. Ameel, W. Vanpaemel, M. J. Dry, W. Voorspoels, G. Storms, Exemplar by feature applicability matrices and other dutch normative data for semantic concepts, Behavior research methods 40 (2008) 1030–1048.
[17] L. Wachowiak, D. Gromann, Systematic analysis of image schemas in natural language through explainable multilingual neural language processing, in: Proceedings of the 29th International Conference on Computational Linguistics, 2022, pp. 5571–5581.
[18] D. Gromann, M. M. Hedblom, Body-mind-language: Multilingual knowledge extraction based on embodied cognition., in: AIC, 2017, pp. 20–33.

[19] S. De Giorgis, A. Gangemi, D. Gromann, Imageschemanet: A framester graph for embodied commonsense knowledge, Semantic Web (????) 1–25.

[20] R. M. Nosofsky, Similarity, frequency, and category representations., Journal of Experimental Psychology: learning, memory, and cognition 14 (1988) 54.

[21] G. Righetti, O. Kutz, D. Porello, N. Troquard, A game of essence and serendipity: Superb owls vs. cooking-woodpeckers, in: M. M. Hedblom, A. A. Kantosalo, R. Confalonieri, O. Kutz, T. Veale (Eds.), Proceedings of the 13th International Conference on Computational Creativity, Bozen-Bolzano, Italy, June 27 - July 1, 2022, Association for Computational Creativity (ACC), 2022, pp. 300–309. URL: http://computationalcreativity.net/iccc22/papers/ICCC-2022_paper_88.pdf.

[22] G. Antoniou, F. v. Harmelen, Web ontology language: Owl, Handbook on ontologies (2009) 91–110.

[23] G. Righetti, D. Porello, N. Troquard, O. Kutz, M. Hedblom, P. Galliani, Asymmetric hybrids: Dialogues for computational concept combination, in: Formal Ontology in Information Systems: Proceedings of the 12th International Conference (FOIS 2021), IOS Press, 2021.

[24] G. Righetti, D. Porello, N. Troquard, O. Kutz, M. M. Hedblom, P. Galliani, Asymmetric hybrids: Dialogues for computational concept combination (extended abstract), in: L. D. Raedt (Ed.), Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022, ijcai.org, 2022, pp. 5329–5333. URL: https://doi.org/10.24963/ijcai.2022/745. doi:10.24963/ijcai.2022/745.

[25] N. Guarino, C. A. Welty, An overview of ontoclean, Handbook on ontologies (2009) 201–220.

[26] J. M. Mandler, C. P. Cánovas, On defining image schemas, Language and cognition 6 (2014) 510–532.

[27] M. M. Hedblom, O. Kutz, R. Peñaloza, G. Guizzardi, Image schema combinations and complex events, KI-Künstliche Intelligenz 33 (2019) 279–291.

[28] M. M. Hedblom, D. Gromann, O. Kutz, In, out and through: Formalising some dynamic aspects of the image schema containment, in: Proceedings of the 33rd Annual ACM Symposium on Applied Computing, 2018, pp. 918–925.

[29] F. Santibáñez, The object image-schema and other dependent schemas, Atlantis (2002) 183–201.

[30] M. M. Hedblom, O. Kutz, Hidden Meaning: Using Image Schema Violations to Expose Hidden Semantical Structures in Metaphors, in: Proceedings of the Joint Ontology Workshops (JOWO 2019), Episode V: The Styrian Autumn of Ontology, volume 2518, CEUR-WS, 2019.

[31] G. Sacco, L. Bozzato, O. Kutz, Introducing Weighted Prototypes in Description Logics for Defeasible Reasoning, in: Proceedings of the 38th Italian Conference on Computational Logic (CILC 2023), CEUR-WS, 2023.

[32] M. M. Hedblom, O. Kutz, Conceptual puzzle pieces: An image schema experiment on object conceptualisation, in: Modeling and Using Context: 11th International and Interdisciplinary Conference, CONTEXT 2019, Trento, Italy, November 20–22, 2019, Proceedings 11, Springer, 2019, pp. 98–111.