Representing Concepts by Weighted Formulas

Claudio MASOLO [a,1], Daniele PORELLO [b]
[a] *Laboratory for Applied Ontology (ISTC-CNR)*
[b] *Free University of Bozen/Bolzano*

Abstract. A concept is traditionally defined via the necessary and sufficient conditions that clearly determine its extension. By contrast, cognitive views of concepts intend to account for empirical data that show that categorisation under a concept presents typicality effects and a certain degree of indeterminacy. We propose a formal language to compactly represent concepts by leveraging on weighted logical formulas. In this way, we can model the possible synergies among the qualities that are relevant for categorising an object under a concept. We show that our proposal can account for a number of views of concepts such as the prototype theory and the exemplar theory. Moreover, we show how the proposed model can overcome some limitations of cognitive views.

Keywords. logical theories of concepts; prototype theory; exemplar theory; conceptual spaces; weighted formulas

1. Introduction

The nature of concepts has been longly debated and analysed both in philosophy and cognitive science (see [1,2] for good introductions). As summarised in [1], it is still controversial whether concepts are mental representations or abstract entities, whether they are objects or cognitive/behavioural abilities. In this paper we focus on how concepts can be formally modelled, that is, on the languages for representing and reasoning about concepts and categorisation. Our analysis is mostly independent of the exact nature of concepts, however, following [3], our underlying intuition is that concepts are a kind of representational device that enables cognitive agents to classify objects.

A formal model of concepts and categorisation is fundamental not only in the scope of cognitive science but also for establishing a motivated connection between experimental sciences and knowledge representation or, more generally, artificial intelligence. In these fields, the initial attempts to explicitly consider the connection between concepts and experimental data about how human agents actually perform classifications—e.g., the theory of frames developed by Marvin Minsky [4]—have been almost abandoned because of the lack of a clear and formal characterisation of concepts. Rigorous logical approaches usually represent concepts by formal definitions, i.e., sets of necessary and sufficient conditions for membership. This view, termed here the *classical view*, can be traced back to one of the founding fathers of analytical philosophy, namely to the

[1] Corresponding Author: via alla cascata 56/c, Trento, Italy, 38123; E-mail: masolo@loa.istc.cnr.it

seminal work of Frege [5], where concepts are defined by logical combinations of basic predicates. In this view, categorisation reduces to the satisfaction of all the conditions stated in the definition. An object is then either categorised or not under a concept, there is indeed no uncertainty, and all the instances of a concept are treated equally. By contrast, empirical experiments in psychology and cognitive science show that everyday concepts often lack a precise definition and that categorisation presents typicality effects and a certain degree of indeterminacy. Fuzzy, probabilistic, and non-monotonic logics modified the classical view to account for the typicality and the indeterminacy of the categorisation, but they introduced some technical difficulties that are still debated (e.g., a solid semantics, a grounded choice of logical operators). Moreover, although there is evidence that cognitive agents do not conform to classical logic, there is no evidence that they do conform to fuzzy, probabilistic, or non-monotonic logics instead. The limitations of the classical view, even when extended to non-classical logics, pushed cognitive scientists to develop alternative models to match the empirical data. These models are usually grouped into the *prototype view*, the *exemplar view*, and the *knowledge view* also called *theory-theory* (see [1,2]), but also Gärdenfors's *theory of conceptual spaces* [3] and Barsalou's *theory of frames* [6] enter this category.

A *prototype* is "a prestored representation of the usual properties associated with the concept's instances" [7, p.487] that is usually given in terms of an attribute-value model. For example, in [7] a prototype is represented by a set of attributes—e.g., colour, shape, taste—with associated *diagnosticity* values that represent "how useful the attribute is in discriminating instances of the concept from instances of contrasting concepts" [7, p.487]. For each attribute, a prototype contains a list of values—e.g., red, green, brown for the colour-attribute—weighted by their *salience*, which is a measure of typicality. By listing alternative values for a given attribute (together with their salience), a prototype does not describe a single best example of a concept, rather it accounts for the variability of the instances of such concept. For example, the apple prototype allows for several colours, although the most typical apple may be red. A prototype provides then a *summary representation* of a concept, a unified description of the concept as a whole "rather than separate representations for each member or for different classes of members" [2, p.42]. The degree of categorisation under a concept corresponds to the degree of similarity between (the representation of) the object and the prototype of the concept. This degree of similarity is computed by means of a rule that applies to all concepts, e.g., Tversky's contrast rule [8], which considers the contrast between shared and distinctive qualities.

The exemplar view rejects the idea of having a summary representation describing all the instances of a concept in an unified way. Instead, it grounds categorisation on the known exemplars. I.e., one must access the memory and compute the degree of similarity between the object to be categorised and the known exemplars (that have been already categorised). The degree of categorisation under a concept is then computed on the basis of these degrees of similarity and of the categorisation of the known exemplars. Differently from the prototype view, even after having seen several apples, there is no general description of what apples are. Furthermore, to categorise an object as an apple, the exemplars of apples are not enough, categorisation is a comparative process that necessarily involves all the known exemplars.

In the knowledge view, categorisation is intended as a reasoning process that uses and is consistent with general knowledge about the domain of interest. For reasons of space, we cannot enter the details, however note that the uniform treatment of the categorisation process and of general knowledge about the world is still a challenge for cognitive models.

Given the criticisms to the classical view, cognitive scientists maintain a quite reluctant attitude towards logic-based approaches. This attitude is even stronger among the followers of the exemplar view that undermines the very idea of a summary description of a concept. However, logical approaches offer—via the inference mechanism and the set of explicit constraints—a general framework to take into account how categorisation and general knowledge about the world interact.

In this paper, we try to re-analyse the limitations and the strengths of the classical view, by exploring a different approach to the formal specifications of concepts that accounts for typicality effects and categorisation indeterminacy. We expand the classical view by using weighted logical formulas to model the relevance and the possible synergies among the features that characterise a concept. To do that, we take advantage of the logical approaches to compactly represent functions developed in Economics and Multiagent Systems (cf. [9]). In particular, we adapt the techniques of [10,11]—originally devoted to compactly represent utility functions over combinations of goods—to represent concepts.

The remainder of the paper is organised as follows. In Section 2, we discuss the different views of concepts and we define a general framework to compare them. Section 3 introduces the formal language to represent concepts based on weighted formulas and assesses its expressivity. Section 4 shows how the prototype view and the exemplar view can be represented by using weighted formulas. Section 5 concludes the paper by indicating a number of viable extensions of the use of weighted formulas.

2. Preliminary considerations

In the classical view, complex concepts are defined in terms of logical combinations of basic concepts represented by primitive and unstructured predicates. The distinction between complex and basic concepts is however not exclusive of the classical view. We have seen that the prototype view embraces an attribute-value model, i.e., a prototype is characterised by several attributes that can assume different (weighted) values, which usually represent the sensory or perceptual capacities of subjects. For instance, red and blue are values on the colour-attribute, while 1kg and 2kg are both weights. The values of the attributes play here the role of basic concepts, applied to the objects to be categorised, but they are partitioned by attributes. Furthermore, values are usually considered holistic and, the ones belonging to the same attribute, mutually exclusive: an object cannot simultaneously weight 1kg and 2kg or be red and blue (even when it has a red part and a blue part). These assumptions encode knowledge about attributes and values that in the classic view must be explicitly introduced. The exemplar view—in particular the *Generalized Context Model* (GCM) introduced in [12]—as well as Gardenfor's conceptual spaces still assume that stimuli are decomposed along several dimensions, that roughly correspond to attributes.[2] In addition, the dimensions are endowed with a *metric*,

[2] Conceptual spaces are more sophisticated. They assume that *spaces* are decomposed into *domains* (e.g., the colour domain) that, in their turn, correspond to maximal sets of integral *dimensions* (e.g., {*hue*, *chromaticness*, *brightness*} for the colour-domain). Here we ignore this distinction and consider only unstructured attributes.

a distance relation—usually determined using *multidimensional scaling* [13] on the basis of similarity judgements—defined between the points (i.e, the counterparts of values) in these dimensions. It is then possible, for instance, to order weights or to express the fact that orange is closer to red than to blue. Furthermore, a metric on the whole space can be defined (several possibilities exist) on the basis of the metrics on its dimensions. The metrics of the dimensions encode quite sophisticated knowledge about the values of the attributes that, as we will see, partially explains the expressive power of these approaches.

In order to generalise from specific views of concepts, following [14] and [15], we consider an n-dimensional space $S^n = \Delta_1 \times \cdots \times \Delta_n$ where the Δ_i-dimensions correspond to attributes—e.g., the three dimensional $\Delta_{colour} \times \Delta_{shape} \times \Delta_{taste}$ space. The points in S^n are n-tuple of values, e.g., $\langle red, round, sweet \rangle$. In this framework, the categorisation under a concept `C` can then be seen as a function $f_{\texttt{C}} : S^n \to \mathbb{R}$ that associates to each point $p \in S^n$ the degree of categorisation (under `C`) of the object (represented by) p. In the classical view, there are indeed only two 'degrees' of categorisation, i.e., $f_{\texttt{C}} : S^n \to \{0,1\}$. Therefore, it is possible to define the extension of the concept `C` by $ext(\texttt{C}) = \{p \in S^n \,|\, f_{\texttt{C}}(p) = 1\}$.

This general framework can be used to compare the previously discussed views of concepts. As we have seen, a first element of comparison concerns the available knowledge about the values of the attributes. To comply with this framework, the classical view needs to be tuned by setting a correspondence between the basic predicates of the logical language and the values of the attributes, and by partitioning the basic predicates into the required dimensions. The prototype view embraces the attribute-value model but does not commit to metric dimensions. This also holds for some implementations of the exemplar view. Differently, conceptual spaces and GCM explicitly ground similarity on the metrics. A second and more important element of comparison concerns the way in which the categorisation function is specified. Note that the space S^n has a combinatorial nature, as it includes all the possible combinations of values. This combinatorial aspect is particularly delicate for cognitive scientists, who often work under the assumption that "the task of category systems is to provide maximum information with the least cognitive effort" [16, p.190]. Thus, in cognitive terms, the plain extensional representation of the functions $f_{\texttt{C}}$—i.e., the list of the degrees of categorisation for all the points in S^n—is not plausible; one needs in fact a more economic specification. In the prototype view, the $f_{\texttt{C}}$ is described by the prototype of `C` (also containing salience and typicality information) and by a categorisation rule, that is shared by all the concepts, e.g., the contrast rule, that computes the degree of categorisation on the basis of the shared and distinctive qualities between an object and a prototype. In the exemplar view, usually the whole set of known exemplars (together with their categorisation) is demanded to describe any $f_{\texttt{C}}$. The degrees of similarity between the exemplars and the object to be categorised can be computed following the technique of the prototype view, or by relying on the metric. Once these degrees of similarity are calculated, they can be aggregated into a single degree of categorisation under `C`. The representation of $f_{\texttt{C}}$ is then less compact than the previous one, where prototypes offer summary descriptions of concepts that abstract from the multitude of the known exemplars.

It is important to note that these cognitive views, and in particular the exemplar view, tend to establish a one-to-one correspondence between the points in S^n and the representations of objects. This entails that objects are reduced to clusters of qualities.

In [17] we criticised this hypothesis from an ontological perspective, however we claim that the identification between points and objects is problematic also from a cognitive or epistemological perspective. Firstly, if objects (exemplars) are perceived and memorised (together with their categorisation), then it seems quite implausible to presuppose that the precise values for all the attributes in S^n are known and stored in memory. This observation applies in particular to the attributes that have a very low or null salience. For instance, one can categorise an object as a melon even without knowing its temperature or just by guessing (with a given degree of tolerance) its weight. Second, categorisations seem revisable after the acquirement of new information. For instance, something that appeared to be a melon, after an accurate inspection, reveals to be a plastic model of a melon. For these reasons, we shall embrace a more liberal approach that allows for partial information about objects (an epistemic state about such object) to enter the input of the categorisation function. The value of the categorisation function expresses then the degree of categorisation of an object under a concept *given the available information*. On the one hand, this assumption seems to be consistent with the empirical data about the individuation and the re-identification of objects. For instance, Pylyshyn [18] supports the idea that the initial individuation and the tracking of objects is not conceptual, rather it is based on a lower level mechanism, which is built into the visual system, that allows for creating and updating an *object file* [19]. An object file groups and maintains the information available on an object that is acquired and updated through time. On the other hand, partial information allows for a non-monotonic categorisation mechanism, an important aspect that will be analysed in Section 3.

Let us now go back to the classical view and, by putting aside, for a moment, the degree of categorisation, consider the extension of a concept $\texttt{C}$ ($ext(\texttt{C})$) previously defined. In first-order logic, one can associate to $ext(\texttt{C})$ the formula $\texttt{C}(x) \equiv \bigvee_{\langle q_1,\ldots,q_n\rangle \in ext(\texttt{C})}(Q_1(x) \wedge \ldots \wedge Q_n(x))$ (where Q_i is the basic predicate that corresponds to the value q_i). That is, an object is classified under $\texttt{C}$ if and only if it satisfies one of the *conjunctions* of qualities of the formula associated to $ext(\texttt{C})$. Firstly, note that, differently from what suggested by the traditional arguments addressed to criticise the classical view, this definition does not necessarily presuppose *essential* qualities of $\texttt{C}$: a core set of qualities shared by all the instances of $\texttt{C}$ is not needed. The Wittgesteinian idea of family resemblance is then not ruled out by the classical view, by the mere logical phrasing. Secondly, although the previous formulation is in general not economic, however, in some cases, it can be compacted. For instance, suppose that $ext(\texttt{C}) = \{\langle crimson, 1kg\rangle, \langle scarlet, 1kg\rangle, \langle magenta, 1kg\rangle\}$. In this case, one may define $Red(x) \equiv Crimson(x) \vee Magenta(x) \vee Scarlet(x)$ and $\texttt{C}(x) \equiv Red(x) \wedge 1Kg(x)$. Thirdly, and more importantly, in the logical approach one can take advantage of the logical language to define constraints about how an object matches the specification. This is particularly relevant in the presence of knowledge about the domain. For instance, if we know that in our domain a complex combination of (possibly negated) qualities implies another quality, then we can use this information to classify an object under $\texttt{C}$, although we do not *explicitly* know the combinations of values listed in $ext(\texttt{C})$. The prototype and the exemplar views do not use this powerful inferential mechanism, whose adjunction would enable the integration of these views with the knowledge view.[3]

[3] In the framework of conceptual spaces, *correlations* are used to introduce constraints on the possible combinations of attribute-values. However, the expressive power of these correlations and the way they are represented are not explicitly stated.

In the remainder of this paper, we propose a uniform framework to model the different views of concepts that we discussed. Moreover, the proposed formalisation will allow us to indicate the limitations and to suggest possible extensions of those approaches.

3. A representation of concepts

We introduce a standard predicative language $\mathscr{L}$ with a set of individual constants $\mathbf{C}$, a set of individual variables $\mathbf{V}$, and sets of predicates $\mathbf{R}$. The set of atomic formulas $Atom$ is defined by $R(t_1,\dots,t_n) \in Atom$ iff $R \in \mathbf{R}$ and $t_1,\dots,t_n \in \mathbf{C} \cup \mathbf{V}$ (where n is the arity of R). The definition of formulas is by induction as follows:

$$\mathscr{L} := \psi \in Atom \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi \to \phi \mid \forall x \phi(x) \mid \exists x \phi(x).$$

Let $\mathscr{F}$ be a finite set of closed formulas of $\mathscr{L}$. An *epistemic state* $\mathscr{E}$ is just a subset of $\mathscr{F}$. We say that a closed formula ϕ of $\mathscr{L}$ holds in the epistemic state $\mathscr{E}$ iff $\mathscr{E} \vdash \phi$, where $\vdash$ is the usual syntactic entailment relation of first-order logic. An epistemic state can be intended as the syntactic finite representation of a (number of) model(s) of a set of sentences. In cognitive terms, it allows for coping with the limitations of an agent by representing the information that is available to the agent in a certain circumstance.

We shall model concepts by means of sets of weighted formulas, which were introduced to compactly represent utility functions over (finite) combinations of goods in [10,11]. Here, we extend the approach of [10,11] to predicative formulas.

A *weighted formula* is a pair (ϕ, w) where ϕ is a formula of $\mathscr{L}$ and w is a weight, usually a real number $w \in \mathbb{R}$.

A *concept base* is a finite set of weighted formulas $\mathtt{C} = \{(\phi_1, w_1), \dots, (\phi_m, w_m)\}$ where the ϕ_js are either closed formulas or open formulas with a designated single free variable x; we also assume that a concept base has to include at least one such open formula.

The *degree of categorisation* of an object a under the concept (represented by the concept base) $\mathtt{C}$ in the epistemic state $\mathscr{E}$ is defined in two steps: firstly, we replace the occurrences of the free variable x in $\mathtt{C}$ with the individual constant a and we collect the weights of the formulas in $\mathtt{C}$ (after the substitution of x with a) that hold in $\mathscr{E}$; secondly, we aggregate those weights by means of an *aggregation function* $F : \mathscr{P}(\mathbb{R}) \to \mathbb{R}$. Given an aggregator F, the degree of categorisation under $\mathtt{C}$ is represented by the function $v_{\mathtt{C}} : \mathscr{P}(\mathscr{F}) \times \mathscr{C} \to \mathbb{R}$, where $\mathscr{C} \subseteq \mathbf{C}$ is a *finite* set of individual constants, defined as follows:

$$v_{\mathtt{C}}(\mathscr{E}, a) = F\{w \mid (\phi, w) \in \mathtt{C} \text{ and } \mathscr{E} \vdash \phi[x/a]\}. \tag{1}$$

In the remainder of this paper, we assume that F is the sum of the weights, but other choices may be discussed.

To enable degrees of classification in $[0,1]$, we can assume any order-preserving bijective function $f : \mathbb{R} \to [0,1]$ and compose $v_{\mathtt{C}}$ with f, i.e., $\bar{v}_{\mathtt{C}} = f(v_{\mathtt{C}}(\mathscr{E}, a))$. Vice versa to obtain a binary categorisation, i.e. degrees in $\{0,1\}$, we can just set a threshold t, i.e., $\bar{v}_{\mathtt{C}} = 1$ iff $v_{\mathtt{C}}(\mathscr{E}, a) \geq t$.

Example 1. Consider the concept base $\mathtt{C} = \{(Red(x) \vee Green(x), w_1), (Round(x), w_2), (\exists y(Eat(y,x) \wedge White(y)), w_3)\}$. This means that the relevant information to establish the degree of categorisation under $\mathtt{C}$ of an object a is whether (i) it is red or green, (ii) it is round, and (iii) it is eaten by a white entity. Consider the epistemic state $\mathscr{E} = \{Red(a), White(b), Eat(b,a)\}$, i.e., the agent knows that a is red and is eaten by b that is white. In this case, $v_{\mathtt{C}}(\mathscr{E}, a)$ is computed as follows. Firstly, we replace x with a in the open formulas of $\mathtt{C}$ and we obtain $Red(a) \vee Green(a)$, $Round(a)$, and $\exists y(Eat(y,a) \wedge White(y))$. Then we sum the weights of the formulas that are entailed by $\mathscr{E}$—namely $Red(a) \vee Green(a)$ and $\exists y(Eat(y,a) \wedge White(y))$—obtaining $v_{\mathtt{C}}(\mathscr{E}, a) = w_1 + w_3$. □

The concept bases approach is sufficiently expressive to generate all the functions from epistemic states to real numbers, at least on finite epistemic states. This entails that, by means of a concept base, we can represent every possible way of weighting the information conveyed by an epistemic state, when defining a concept. We establish this fact in the following proposition.

Given a set of formulas X, we write $\bigwedge X$ to denote the conjunction of all the formulas in X. In case X is a singleton, $\bigwedge X = X$. We say that two epistemic states $\mathscr{E}$ and $\mathscr{E}'$ are logically equivalent when the conjunction of the formulas in $\mathscr{E}$ is logically equivalent to the conjunction of the formulas in $\mathscr{E}'$: $\bigwedge \mathscr{E} \equiv \bigwedge \mathscr{E}'$. Notice that a concept base cannot distinguish between two logically equivalent epistemic states, therefore we discuss functions from the power set of $\mathscr{F}$ *modulo* logical equivalence $\equiv$: we denote by $\mathscr{P}(\mathscr{F})/\equiv$ the set of equivalence classes of sets of formulas of $\mathscr{F}$ modulo logical equivalence $\equiv$.

Proposition 1. Every function $v : \mathscr{P}(\mathscr{F})/\equiv \to \mathbb{R}$ can be generated by means of a concept base $\mathtt{C}$ and by the sum aggregator. That is, for every $\mathscr{E}$, $v(\mathscr{E}) = v_{\mathtt{C}}(\mathscr{E}, a)$ for some concept base $\mathtt{C}$ and finite set of constants $\mathscr{C}$, where $a \in \mathscr{C}$.

Proof. (Sketch) We can establish this result by means of Theorem (3.2) in [10] that shows that concept bases defined on any propositional formulas can represent every function from finite propositional models (assignments) to real numbers.

Let v be a function from $\mathscr{P}(\mathscr{F})/\equiv$ to $\mathbb{R}$ and (X, w) a pair $X \in \mathscr{P}(\mathscr{F})/\equiv$ and $w \in \mathbb{R}$. To simplify the following argument, we assume that the set $\mathscr{C} = \{a\}$, for a designated constant a that occurs in some of the formulas of $\mathscr{F}$.

For each $(X, w) \in v$ (i.e. in the graph of v), such that $|X| = m$, we include in the concept base $\mathtt{C}$ the weighted formula $(\phi_1' \wedge \cdots \wedge \phi_m', w_X)$ where for $\phi_j \in X$ that does not contain the constant a, $\phi_j' = \phi_j$ and for $\phi_j \in X$ that contains a, $\phi_j' = \phi_j[a/x]$. Moreover, w_X is a weight which is inductively computed as follows, for any non-empty set of formulas $X, Y \in \mathscr{P}(\mathscr{F})/\equiv$ and $Y \neq X$:[4]

$$w_X = v(X) - \sum_{X \vdash \bigwedge Y} w_Y. \tag{2}$$

The concept base $\mathtt{C}$ generates the function v. For every X, we show that $v(X) = v_{\mathtt{C}}(X, a)$. Suppose $(X, w) \in v$. Then $v_{\mathtt{C}}(X, a)$ is the sum of all the formulas in $\mathtt{C}$ that are entailed by X. By construction of $\mathtt{C}$, X entails the following formulas of $\mathtt{C}$: $X \vdash \phi_1' \wedge \cdots \wedge \phi_m'$, and $X \vdash \bigwedge Y$, for every $(\bigwedge Y, w_Y) \in \mathtt{C}$ such that $X \vdash \bigwedge Y$ (where we possibly

[4] Note that the value of v on the empty set has to be set equivalent to the weight associated to a tautology.

replace the occurrences of a with x in the formulas of Y). Thus $v_{\mathtt{C}}(X) = w_X + \sum_{X \vdash Y} w_Y$, which by Definition (2) equals to: $v(X) - \sum_{X \vdash Y} w_Y + \sum_{X \vdash Y} w_Y = v(X)$.[5] $\square$

Note that the function v in the previous proof is represented by a concept base where formulas are all conjunctions of formulas of $\mathscr{F}$. That is, restricting to conjunctions is sufficient to generate all the functions from the epistemic states to the reals.

By restricting the types of formulas in the definition of the concept base, we can single out specific classes of functions. For instance, consider the concept bases with the form $\mathtt{C} = \{(Q_1(x), w_1), \dots, (Q_m(x), w_m)\}$, that is, where formulas are restricted to atomic predicates. The categorisation function generated by $\mathtt{C}$ is additive on the formulas included in the epistemic states, i.e., $v_{\mathtt{C}}(\mathscr{E}, a) = \sum_{e \in \mathscr{E}} v_{\mathtt{C}}(\{e\}, a)$.

Concept bases allow us also to express *super-additive* and *sub-addictive* functions. A function $v_{\mathtt{C}}$ is super-addictive on epistemic states when $v_{\mathtt{C}}(\mathscr{E} \cup \mathscr{E}', a) \geq v_{\mathtt{C}}(\mathscr{E}, a) + v_{\mathtt{C}}(\mathscr{E}', a)$. E.g., consider $\mathtt{C} = \{(Q_1(x), w_1), (Q_2(x), w_2), (Q_1(x) \wedge Q_2(x), w_3)\}$ where $w_3 > w_1 + w_2$. In this case, for the categorisation under $\mathtt{C}$, the relevance of the conjunction of Q_1 and Q_2 outweighs the one of each quality considered separately.

The case of sub-additive functions, i.e., when $v_{\mathtt{C}}(\mathscr{E} \cup \mathscr{E}', a) \leq v_{\mathtt{C}}(\mathscr{E}, a) + v_{\mathtt{C}}(\mathscr{E}', a)$, is related to the issue of monotonicity. A consequence of this modelling is that the categorisation under a concept is *non-monotonic* in the following sense. Suppose that $\mathscr{E}$ and $\mathscr{E}'$ are epistemic states such that $\mathscr{E} \subseteq \mathscr{E}'$, i.e., $\mathscr{E}'$ contains more information than $\mathscr{E}$. There are concept bases $\mathtt{C}$ for which $v_{\mathtt{C}}(\mathscr{E}, a) \geq v_{\mathtt{C}}(\mathscr{E}', a)$. For instance, consider $\mathtt{C} = \{(Q_1(x), 2), (Q_2(x), -1)\}$, $\mathscr{E} = \{Q_1(a)\}$, and $\mathscr{E}' = \{Q_1(a), Q_2(a)\}$. Then we obtain that $v_{\mathtt{C}}(\mathscr{E}, a) = 2$ while $v_{\mathtt{C}}(\mathscr{E}', a) = 1$. Non-monotonicity is a consequence of our view of categorisation as depending on the epistemic state of a cognitive agent, on the actual information that are available to the agent when performing the categorisation.[6]

4. Applications to cognitive views of concepts

We restrict here to a *finite* space $S^n = \Delta_1 \times \dots \times \Delta_n$ of vectors of values within the dimensions Δ_i. S^n serves as a general framework to model concepts. In particular, we consider the function $f_{\mathtt{C}} : S^n \to \mathbb{R}$ introduced in Section 2 to represent the degree of categorisation under the concepts $\mathtt{C}$ of an object, which is associated to a point in S^n. This space is basically shared by many cognitive approaches, although some of them use richer structures. E.g., Gärdenfors assumes a continuous metric structure where the Δ_i are domains possibly composed by several dimensions. However, for the purpose of analysing the $f_{\mathtt{C}}$-functions, this model suffices.

We capture (the structure of) the space S^n in our predicative language as follows. We represent the quality values q_i in each Δ_i by means of atomic unary predicates Q_i of our language $\mathscr{L}$, which are partitioned according to the dimensions Δ_i. Our set of predicates includes then the following typed unary predicates: $\{Q_1^1, Q_1^2, \dots, Q_1^{m_1}, \dots, Q_n^1, Q_n^2, \dots, Q_n^{m_n}\}$, where m_h is the cardinality of Δ_h. We assume a standard first-order model (D, I) to define the semantics of the formulas of our lan-

[5] This results generalises the proof of Theorem (3.2) in [10]. We assume here that the domain is every possible finite set of formulas, and not just sets of atoms. For this reason, the way of computing weights has to take into account the sets Y such that $X \vdash \bigwedge Y$, and not just the subsets of X.

[6] For a characterisation of the types of functions generated by types of concept bases, we may adapt the results in [10].

guage $\mathscr{L}$. In particular, contra Gärdenfors, we do not suppose that the points of S^n are objects, as we interpret our individual constants in the domain D rather than in S^n. Since the values in each dimension are alternative, we assume axioms with form $\forall x(Q^i_h(x) \to \neg Q^j_h(x))$, for every $i \neq j$. That means that an object cannot satisfy two distinct atomic predicates in the same dimension.

In the subsequent sections, we discuss how to represent $f_{\mathtt{C}}$-functions by means of concept bases, then we consider the prototype view and the exemplar view.

4.1. Representing the categorisation functions

Consider now a finite set of individual constants $\mathscr{C} \subseteq \mathbf{C}$ representing a set of objects. A point $p \in S^n$ is represented by the set $\tau(p)$ containing $n \cdot |\mathscr{C}|$ atomic closed formulas of $\mathscr{L}$, one for each instantiation by a constant in $\mathscr{C}$ of the predicate corresponding to the value of the point in the dimension Δ_i. For instance, given $\mathscr{C} = \{a,b\}$, the representation of $p = \langle \mathit{red}, \mathit{1kg}\rangle \in \Delta_{color} \times \Delta_{weight}$ is the set of formulas $\tau(p) = \{\mathit{Red}(a), \mathit{1kg}(a), \mathit{Red}(b), \mathit{1kg}(b)\}$.

Let $\mathscr{F}$ be a set of closed formulas that includes the set of formulas that represent all the points in S^n (given a set $\mathscr{C}$). We can approach the problem of representing the function $f_{\mathtt{C}} : S^n \to \mathbb{R}$ by introducing a concept base $\mathtt{C}$ for which the generated function $v_{\mathtt{C}} : \mathscr{P}(\mathscr{F}) \times \mathscr{C} \to \mathbb{R}$ is such that $f_{\mathtt{C}}(p) = w$ iff $v_{\mathtt{C}}(\tau(p), c) = w$ for all $c \in \mathscr{C}$.

In this specific case, the following construction suffices:

$$f_{\mathtt{C}}(\langle q_1, \ldots, q_n\rangle) = w \text{ iff } (Q_1(x) \wedge \ldots \wedge Q_n(x), w) \in \mathtt{C}. \tag{3}$$

By construction, $\tau(\langle q_1, \ldots, q_n\rangle)$ contains the set of formulas $\{Q_1(c), \ldots, Q_n(c)\}$ for all $c \in \mathscr{C}$ (and no additional information on c). It is easy then to see then that $v_{\mathtt{C}}(\tau(p), c) = w$ for all $c \in \mathscr{C}$. Thus, our framework can represent any $f_{\mathtt{C}}$.[7] In this case, performing the classification under the concept $\mathtt{C}$ requires to know all the qualities (in S^n) of the object. Indeed, this observation is not surprising because it also applies to $f_{\mathtt{C}}$.

Example 2. Consider $S^2 = \mathit{weight} \times \mathit{shape} = \{\mathit{1kg}, \mathit{2kg}, \mathit{3kg}\} \times \{\mathit{red}, \mathit{blue}, \mathit{brown}\}$ and suppose that $f_{\mathtt{C}}(p) \neq 0$ only in the following cases:

$$\{\langle\langle \mathit{1kg}, \mathit{red}\rangle, w\rangle, \langle\langle \mathit{1kg}, \mathit{blue}\rangle, w\rangle, \langle\langle \mathit{1kg}, \mathit{brown}\rangle, w\rangle, \langle\langle \mathit{3kg}, \mathit{red}\rangle, w'\rangle\}.$$

According to the previous construction, we obtain the following concept base:

$$\mathtt{C} = \{(\mathit{1Kg}(x) \wedge \mathit{Red}(x), w), (\mathit{1Kg}(x) \wedge \mathit{Blue}(x), w), (\mathit{1Kg}(x) \wedge \mathit{Brown}(x), w), (\mathit{3Kg}(x) \wedge \mathit{Red}(x), w')\}. \tag{4}$$

[7] This claim can be inferred also as a corollary of Proposition 1 as follows. Let $\tau(S^n) = \bigcup_{p \in S^n} \tau(p)$. Given $f_{\mathtt{C}} : S^n \to \mathbb{R}$, we can (partially) define a class of functions $f'_{\mathtt{C}} : \mathscr{P}(\tau(S^n)) \to \mathbb{R}$ such that $f_{\mathtt{C}}(p) = f'_{\mathtt{C}}(\tau(p))$. Notice that any $f'_{\mathtt{C}}$ is a function from the powerset of a set of formulas to the real numbers. Therefore, by Proposition 1, $f'_{\mathtt{C}}$ can be represented by means of a concept base, hence also $f_{\mathtt{C}}$ can. However, $f'_{\mathtt{C}}$ may contain more information than $f_{\mathtt{C}}$, since $f'_{\mathtt{C}}$ returns a value on any subset of $\tau(S^n)$. In particular, it returns a value for the proper subsets of $\tau(p)$, possibly providing information about the relative weight of the qualities in $\tau(p)$ for the classification degree of $\tau(p)$.

In this case, we can find a more compact concept base $\mathtt{C}'$—where $\mathtt{C}$ and $\mathtt{C}'$ generate the same function—by noticing that, regardless of the shape, weighting 1kg is sufficient to be the categorised with degree w.[8]

$$\mathtt{C}' = \{(\mathit{1Kg}(x), w), (\mathit{3Kg}(x) \wedge \mathit{Red}(x), w')\}. \tag{5}$$

4.2. Prototypes view

We have seen that the prototype view economically represents an $f_{\mathtt{C}}$ in terms of a given prototype and a general categorisation rule, which is valid for all concepts. Usually, the prototype is not reduced to the best example of the concept. Here we consider the approach in [7] where a prototype is represented in terms of a diagnosticity value for each dimension Δ_i and a list of values, weighted by salience, for each Δ_i. Alternative categorisation rules exist. However, to illustrate the general mechanism for capturing the prototype view in our framework, we suppose here a rule that just sums up the products of the diagnosticity and salience weights of all the values that the object to be categorised shares with the prototype. In this way, the categorisation rule can be directly represented by a function F as in Equation (1). It remains to be shown how to build the concept base.

Assume S^n and the prototype $\pi_{\mathtt{C}}$ defined as follows (where q_i^j is a quality value in Δ_j, s_i^j is a salience weight, and $q_i^j \neq q_k^j$ for $i \neq k$):

$$\pi_{\mathtt{C}} = \{(q_1^1, s_1^1), \ldots, (q_r^1, s_r^1), \ldots, (q_1^n, s_1^n), \ldots, (q_m^n, s_m^n)\}.$$

Furthermore, assume that the diagnosticity weight of Δ_i is d_i. The concept base $\mathtt{C}$ that represents the prototype $\pi_{\mathtt{C}}$ can be easily introduced as follows:

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

Firstly, note that in our approach $v_{\mathtt{C}}(\mathscr{E}, a)$ is defined even when $\mathscr{E}$ contains partial information about a, e.g., when only some qualities of a are known. Secondly, it is easy to see that *linear separability* holds in this case, i.e., the degree of categorisation is determined by summing up the weights of independent qualities of the object. Accordingly, the concept base $\mathtt{C}$ generates only additive functions (see [10,11]). However, the concept bases approach allows also to keep track of the *configurations of qualities* that are relevant for the classification under $\mathtt{C}$ (see Example 3). Thirdly, as we noticed, this approach allows also for modelling non-monotonic behaviours that are important, for instance, for representing concepts in the domain of medical diagnosis (see Example 4).

Example 3. Consider the following three alternative concept bases for apples:

$$\mathtt{A1} = \{(\mathit{Red}(x), w_1), (\mathit{Green}(x), w_2), (\mathit{Round}(x), w_3)\},$$

$$\mathtt{A2} = \{(\mathit{Red}(x) \wedge \mathit{Round}(x), w_1 + w_3), (\mathit{Green}(x) \wedge \mathit{Round}(x), w_2 + w_3)\};$$

$$\mathtt{A3} = \{(\mathit{Red}(x), w_1), (\mathit{Green}(x), w_2), (\mathit{Round}(x), w_3), (\mathit{Red}(x) \wedge \mathit{Round}(x), w_4)\},$$

[8] For a precise definition of succinctness of a representation, we refer to [10,11]. Here we notice that the length of a representation can be measured as the sum of the complexities of the formulas included in a concept base.

where $w_1 \neq w_2 \neq w_3 \neq w_4$ and $w_4 > w_1 + w_3$. Take the epistemic state $\mathscr{E} = \{Red(a), Round(a), Green(b)\}$, i.e., we know that a is both red and round, while for b we just know that its colour is green. The functions $v_{\mathtt{A1}}$, $v_{\mathtt{A2}}$, and $v_{\mathtt{A3}}$ are different. For instance, $v_{\mathtt{A1}}(\mathscr{E}, b) = v_{\mathtt{A3}}(\mathscr{E}, b) = w_2$, whereas $v_{\mathtt{A2}}(\mathscr{E}, b) = 0$, and $v_{\mathtt{A1}}(\mathscr{E}, a) = v_{\mathtt{A2}}(\mathscr{E}, a) = w_1 + w_3$, whereas $v_{\mathtt{A3}}(\mathscr{E}, a) = w_4 > w_1 + w_3$. In A1, the qualities are independent, hence by matching one of them, an object has a degree of categorisation greater than 0. This is not the case of A2, where only combinations of colours and shapes are relevant (and $v_{\mathtt{A1}}$ and $v_{\mathtt{A2}}$ coincide on these combinations). Finally $v_{\mathtt{A3}}$ is a super-addictive version of $v_{\mathtt{A1}}$, the combination of the red and round qualities has a weight that is greater than the sum of the weights of the single qualities.

Example 4. Suppose that we want to classify an individual a according to the disease that she may suffer. For instance, the concept of flu may be represented by the following concept base: $\mathtt{FLU} = \{(Fever(x), w_1), (Nausea(x), w_2), (Spots(x), -w_3)\}$, where adding the symptom 'spots' significantly decreases the reliability of the classification under FLU, because it is a strong indication of chickenpox (e.g. assume that $w_3 \geq w_1 + w_2$). Consider the epistemic states $\mathscr{E} = \{(Fever(a), Nausea(a)\}$ and $\mathscr{E}' = \{(Fever(a), Nausea(a), Spots(a)\}$. We obtain that $v_{\mathtt{FLU}}(\mathscr{E}, a) > v_{\mathtt{FLU}}(\mathscr{E}', a)$, although $\mathscr{E} \subset \mathscr{E}'$. In this case we have then a non-monotonic behaviour.

4.3. Exemplar view

We show how a simplified version of the exemplar view can be represented in our approach. Although the proposed simplified version has a purely illustrative purpose, it allows us to highlight some important aspects of the exemplar view and of our proposal.

Suppose to have a set $E = \{\varepsilon_1, \ldots, \varepsilon_h\}$ of exemplars. According to the exemplar view, each exemplar is represented by a point in the space, i.e., $E \subseteq S^n$. To deal with similarity using the same technique adopted in Section 4.2, we represent exemplars by specific kinds of prototypes. More precisely, we associate to each $\varepsilon_i = \langle q_1, \ldots, q_n \rangle$ the prototype $\pi_i = \langle (q_1, 1), \ldots, (q_n, 1) \rangle$ that contains a single quality for each dimension Δ_i and where all the qualities are equally weighted, i.e., there is no information about the salience of dimensions. Following Section 4.2, each π_i is represented by the concept base $\mathtt{E}_i = \{(Q_1(x), 1), \ldots, (Q_n(x), 1)\}$. Furthermore, as the categorisation of each ε_i is known, the set of exemplars E, as well as the set of concept bases $\mathbb{E} = \{\mathtt{E}_1, \ldots, \mathtt{E}_n\}$, is partitioned by the concepts. For each concept C, it is then possible to identify the set $E_{\mathtt{C}} \subseteq E$ containing all the exemplars of C and the corresponding $\mathbb{E}_{\mathtt{C}} \subseteq \mathbb{E}$. Note that exemplars are then subject neither to typicality effects nor to categorisation indeterminacy.

The categorisation function of an object a under the concept C is computed in two steps. Firstly, we compute the similarities of a with respect to all the exemplars in E (in a given epistemic state $\mathscr{E}$). I.e., in our framework, we calculate $v_{\mathtt{E}_i}(\mathscr{E}, a)$ for all $\mathtt{E}_i \in \mathbb{E}$. Then we aggregate these similarities to individuate the degree of categorisation of a under a concept C. Different aggregators can be considered, e.g.:

$$(i)\ \ v_{\mathtt{C}}(\mathscr{E}, a) = \min\{v_{\mathtt{E}_i}(\mathscr{E}, a) \mid \mathtt{E}_i \in \mathbb{E}_{\mathtt{C}}\},$$

$$(ii)\ \ v_{\mathtt{C}}(\mathscr{E}, a) = \sum_{\mathtt{E}_i \in \mathbb{E}_{\mathtt{C}}} \{v_{\mathtt{E}_i}(\mathscr{E}, a) \mid \mathtt{E}_i \in \mathbb{E}_{\mathtt{C}}\} / |\mathbb{E}_{\mathtt{C}}|,$$

$$(iii)\ \ v_{\mathtt{C}}(\mathscr{E},a) = \sum_{\mathtt{E}_i \in \mathbb{E}_{\mathtt{C}}} \{v_{\mathtt{E}_i}(\mathscr{E},a)\} - \sum_{\mathtt{E}_i \in \mathbb{E} \setminus \mathbb{E}_{\mathtt{C}}} \{v_{\mathtt{E}_i}(\mathscr{E},a)\}.$$

The aggregators (i) and (ii) focus only on the exemplars classified under $\mathtt{C}$, whereas (iii) requires the access to all the available exemplars (as usual, in the exemplar view).

Firstly, notice that the exemplar view usually does not consider categorisation indeterminacy. This allows for introducing newly classified objects among the exemplars of a given concept. As suggested in Section 3, we may introduce a threshold to rule out indeterminacy of classification under a given concept. However, nothing prevents objects to exceed the categorisation threshold for several concepts. To avoid multiple categorisations, one could classify an object under the concept with the higher categorisation degree. In this case, categorisation would necessarily require to access all the known exemplars. However, one could assume categorisation indeterminacy also for exemplars, i.e., one could keep track of the degrees of categorisation of an exemplar under all the concepts and take into account these degrees in the $v_{\mathtt{C}}$ functions.

Secondly, the similarity considered in the previous model only counts the qualities that an object has in common with a given exemplar. Assume $S^2 = \{red, orange, blue\} \times \{1kg, 2kg, 3kg\}$ and $\varepsilon = \langle red, 1kg \rangle$ with the corresponding concept base $\mathtt{E} = \{(Red(x),1),(1Kg(x),1)\}$. Take the epistemic state $\mathscr{E} = \{Blue(a), 1Kg(a), Red(b), 2Kg(b), Blue(c), 2Kg(c), Blue(d), 3Kg(d)\}$. In this case, we obtain that $v_{\mathtt{E}}(\mathscr{E},a) = v_{\mathtt{E}}(\mathscr{E},b) = 1$ and $v_{\mathtt{E}}(\mathscr{E},c) = v_{\mathtt{E}}(\mathscr{E},d) = 0$. However, intuitively, c seems closer to the prototype E than d because c and d have the same colour, but the weight in E (1kg) is closer to the weight of c (2kg) than to the weight of d (3kg). The distance defined on dimensions encodes this information about the geometrical structure of qualities. Approaches based on metric spaces, e.g., conceptual spaces and GCM, allow then for a much more powerful treatment of similarity. Following our partitioning of the basic predicates into dimensions, one could assume a distance defined between the predicates of a given dimension and use this information in the concept bases corresponding to the exemplars. For instance, by knowing $d(1Kg, 2Kg)$ and $d(1Kg, 3Kg)$, one could modify the previous $\mathtt{E} = \{(Red(x),1),(1Kg(x),1)\}$ into $\mathtt{E}' = \{(Red(x),1),(1Kg(x),1),(2Kg(x), 1 - d(1Kg,2Kg)),(3Kg(x),1 - d(1Kg,3Kg))\}$ (assuming that the values of the distance function are in $[0,1]$). In this case, although being 2kg heavy and being 3kg heavy are not as important as being 1kg heavy for the categorisation, they still contribute by producing some weight (and the contribution of being 2k heavy is bigger than the one of being 3kg heavy). This information is manifestly onerous, however it actually grounds the high expressive power of conceptual spaces and GCM.

Finally, note that our account is also open to manage partial information about both objects (to be categorised) and exemplars. Partial information about objects can be directly modelled by the epistemic states. Partial information about exemplars can be modeled by abstracting the prototypes. For instance, if the weight of the exemplar ε is not known, as it may not be relevant to categorise ε under its concept, then the previous $\mathtt{E}$ could be modified into $\mathtt{E}'' = \{(Red(x),1)\}$, i.e., the weight does not influence the degree of similarity with ε. Analogously, if one just knows that the weight of ε is 1kg or 2kg, and this is enough to classify the exemplar under its concept, $\mathtt{E}$ can be modified into $\mathtt{E}''' = \{(Red(x),1),(1Kg(x) \vee 2Kg(x),1)\}$.

5. Possible extensions

We conclude by discussing a number of possible extensions of the use of concept bases to represent concepts. A detailed treatment of these points is left to a future dedicated work. Both the definitions of the epistemic state and of the concept base are quite general and they permit to represent a variety of information relative to a categorisation task. In general, we view the epistemic state $\mathscr{E}$ as the information available for the categorisation task and the concept base $\mathtt{C}$ as the information that is needed for the categorisation task. For instance, the epistemic state may include formulas that express correlations between qualities, e.g., $\mathscr{E} = \{Q_1(a), \forall x(Q_1(x) \to Q_2(x))\}$. In this case, $\mathscr{E}$ represents the situation in which an agent explicitly knows that a has Q_1 and she may infer, by knowing the correlation $\forall x(Q_1(x) \to Q_2(x))$, that a has also Q_2. The inferred information can be used to perform a categorisation task. Moreover, the epistemic state can encode various levels of uncertainty about the qualities of an object, e.g., by means of the disjunction $Q_1(a) \vee Q_2(a) \vee Q_3(a)$. Whether the uncertain information is sufficient for a categorisation task depends on the accuracy demanded by the concept base.

The expressiveness of the formulas in the concept base enables also to model *relational concepts*. E.g., consider the concept base $\mathtt{C} = \{(\exists y(\mathit{Eat}(x,y) \wedge \mathit{White}(y)), w_1), (\exists y(\mathit{Drink}(x,y) \wedge \mathit{Red}(y)), w_2)\}$. In this case, the categorisation of an object a under $\mathtt{C}$ relies exclusively on qualities of objects that are related to a by the *Eat* or the *Drink* relation. I.e., the features that are relevant for categorisation concern how a is linked to some objects with given qualities. Admitting universally quantified formulas in the concept base raises an interesting question about the nature of universal quantification. For instance, assume to include in a concept base $\mathtt{C}$ the weighted formula $(\forall y(\mathit{Eat}(x,y)), w)$ or $(\forall y(\mathit{White}(y) \to \mathit{Eat}(x,y)), w)$. One possibility is to view universal quantification in a substitutional way, that is, an universal quantification holds with respect to the epistemic state $\mathscr{E}$ because it applies to all the instances appearing in $\mathscr{E}$. Accordingly, this entails that the categorisation under $\mathtt{C}$ depends on the available information about the objects explicitly mentioned in the epistemic state.

An important issue, that we leave open here, is whether a concept $\mathtt{C}$ may be included in the concept base of another concept $\mathtt{C}'$ and whether categorisation (under a concept) statements may be part of an epistemic state. Apparently, traditional definitions of concepts permit that, e.g., a human is a rational animal. In principle, we may include concepts as weighted formulas in a concept base, e.g. $\mathtt{HUMAN} = \{(\mathtt{RATIONAL}(x), w), (\mathtt{ANIMAL}(x), w')\}$. A first difficulty concerns the fact that, in general, objects are classified under concepts with given degrees. One then needs to combine the weights of the formulas and the degrees of categorisation under the concepts, that may affect the weights of the formulas. For instance, although the weight w of $\mathtt{RATIONAL}(x)$ is high, the degree of categorisation under RATIONAL of the object a could be very low. A second problem is to understand how a concept base that contains another concept deals with the information about that concept. For instance, one may assume that the concept bases of RATIONAL and of ANIMAL are imported in the concept base of HUMAN. If the weights of the formulas in RATIONAL and ANIMAL are not modified in HUMAN, then the concept base of HUMAN is simply the union of the two concept bases. This amounts to requiring two classification tasks, one under animals and one under rational, that are independent of each other. By contrast, the classification tasks under RATIONAL or ANIMAL may be modified when interpreted as part of the concept HUMAN; for instance, the weights of the qualities that are required for categorising animals may need to be adjusted when categorising humans. The problem of combining concepts in a cognitively significant way is a difficult problem, related to compositionality, that we leave open here. We only stress that, by means of the representation in terms of concept bases, a number of possible combination strategies may be defined. Moreover, note that some concepts seem to require both a concept combination and a relational property. Consider the case of the concept *carnivorous*. Its concept base shall include a weighted proposition $(\exists y(Eat(x,y) \wedge \mathtt{ANIMAL}(y)), w)$, which mention the concept ANIMAL and the *Eat* relation.

To conclude, we developed a framework to represent cognitive views of concepts by means of weighted logical formulas. We faithfully represented the prototype view and the exemplar view, by capturing their categorisation functions. A final point for future work is dedicated to extending this approach to larger classes of categorisation functions.

Acknowledgements. We thank the anonymous reviewers for their helpful comments.

References

[1] E. Margolis and S. Laurence, editors. *Concepts: Core Readings*. MIT Press, 1999.

[2] G. L. Murphy. *The Big Book of Concepts*. MIT press, 2002.

[3] P. Gärdenfors. *Conceptual spaces - The geometry of thought*. MIT Press, 2000.

[4] M. Minsky. A framework for representing knowledge. In P. H. Winston, editor, *The Psychology of Computer Vision*. McGraw-Hill, 1975.

[5] G. Frege. On function and concept. In *Translations from the Philosophical Writings of Gottlob Frege*, pages 1–128. Oxford: Blackwell, third edition, 1980.

[6] 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.

[7] 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.

[8] A. Tversky. Features of similarity. *Psychological Review*, (84):327–352, 1977.

[9] Y. Chevaleyre, P. E. Dunne, U. Endriss, J. Lang, M. Lemaitre, N. Maudet, J. Padget, S. Phelps, J. A. Rodriguez-Aguilar, and P. Sousa. Issues in multiagent resource allocation. *Informatica*, 30(1), 2006.

[10] J. Uckelman, Y. Chevaleyre, U. Endriss, and J. Lang. Representing utility functions via weighted goals. *Mathemathical Logic Quartely*, 55(4):341–361, 2009.

[11] J. Uckelman and U. Endriss. Compactly representing utility functions using weighted goals and the max aggregator. *Artificial Intelligence*, 174(15):1222–1246, 2010.

[12] R. M. Nosofsky. Exemplars, prototypes, and similarity rules. In A. Healy, S. Kosslyn, and R. Shiffrin, editors, *From Learning Theory to Connectionist Theory: Essays in Honor of W. K. Estes*, volume 1, pages 149–168. Lawrence Erlbaum Associates, 1992.

[13] S. S. Jones and L. M. Koehly. Muldimensional scaling. In G. Keren and C. Lewis, editors, *A Handbook for Data Analysis in the Behavioral Sciences*, pages 95–163. Lawrence Erlbaum Associates, 1993.

[14] J. Aisbett and G. Gibbon. A general formulation of conceptual spaces as a meso level representation. *Artificial Intelligence*, 133(1):189–232, 2001.

[15] M. Raubal. Formalizing conceptual spaces. In L. Vieu and A. Varzi, editors, *Formal ontology in information systems: Proceedings of the third international conference (FOIS 2004)*, pages 153–164, 2004.

[16] 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.

[17] 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.

[18] Z. W. Pylyshyn. *Things and Places. How the Mind Connects with the World.* MIT Press, 2007.

[19] D. Kahneman, A. Treisman, and B. J. Gibbs. The reviewing of object files: Object-specific integration of information. *Cognitive psychology*, 24(2):175–219, 1992.