

August 2020,  
to appear in Lucas Bechberger, Kai-Uwe Kühnberger and Mingya Liu (eds). *Concepts in Action - Representation, Learning and Application*. Language, Cognition, and Mind. Springer.

## A qualitative similarity framework for the interpretation of natural language similarity expressions

Helmar Gust and Carla Umbach

**Abstract** In this paper, a representational framework is presented featuring a qualitative notion of similarity. It is aimed at issues of natural language semantics, in particular the semantics of expressions of similarity and sameness and their role in comparison and ad-hoc kind formation. The framework makes use of attribute spaces, which are well-established in AI and also in some branches of natural language semantics, e.g., frame-based approaches (Barcalou 1992). What distinguishes attribute spaces and representations as proposed in this paper is the idea of systems of predicates on attribute spaces corresponding to predicates on the domain. On the worldly side, a domain includes a set of relevant predicates talking about individuals. These predicates have counterparts on the representational side talking about points of an attribute space. Counterpart predicates are required to be consistent with their originals; more precisely, they have to agree in truth-value on the set of positive and negative exemplars more precisely, they have to agree in truth-value on the set of positive and negative exemplars thereby approximating the original predicates. Moreover, counterpart predicates will be assumed to have convex and open extensions. This system facilitates a qualitative notion of similarity which is suited to account for the meaning of natural language similarity expressions and, furthermore, their role in comparison and ad-hoc kind formation.

---

Helmar Gust  
Institute of Cognitive Science, University of Osnabrück, [hgust@uos.de](mailto:hgust@uos.de)

Carla Umbach  
Department of German Language & Literature I, University of Cologne,  
[carla.umbach@uni-koeln.de](mailto:carla.umbach@uni-koeln.de)

## 1 Introduction

In this paper, a representational framework is presented featuring a qualitative notion of similarity. It is aimed at issues of natural language semantics, in particular the semantics of expressions of similarity and sameness and their role in comparison and ad-hoc kind formation.<sup>1</sup> Starting point was the interpretation of such expressions in German and English, for example *so/such*, *ähnlich/similar*, and *gleich/same*, which all denote similarity in some sense. It would be unsatisfactory, however, to treat similarity as a primitive predicate because semantic differences between individual similarity expressions would be obscured, for example, the fact that *ähnlich/similar* are gradable while *so/such* and *gleich/same* are not (see Umbach and Gust in print). Furthermore it would be difficult to establish the connection similarity expressed by scalar and non-scalar equative comparison constructions, as shown in (1).

- |     |    |                                  |                      |
|-----|----|----------------------------------|----------------------|
| (1) | a. | Anna is as tall as Berta.        | scalar / adjectival  |
|     | b. | Anna has a car like Berta's.     | non-scalar / nominal |
|     | c. | Anna is dancing just like Berta. | non-scalar / verbal  |

Finally, a primitive similarity predicate would leave no room to account for the observation that certain similarity expressions, in certain contexts, can be used to form ad-hoc kinds. German *so* as well as English *such* combined with nominal expressions may refer to kinds (or concepts) instead of individuals. In (2a,b), for example, *so ein Fahrzeug / such a vehicle* does not refer to a particular vehicle but instead to an ad-hoc created kind of vehicles including the set of vehicles similar to the one the speaker points to. Umbach and Stolterfoht present experimental evidence that features licensing ad-hoc kinds must be principally connected to concepts, excluding factual and statistical properties (Umbach & Gust 2014, König & Umbach 2018, Umbach & Stolterfoht in prep.). Thus, a complex notion of similarity not only provides a detailed semantic interpretation of natural language similarity expressions – it opens a window into mechanisms of concept formation.

- |     |   |
|-----|---|
| (2) | (Speaker points to an oversized car that makes enormous noise:) |
|     | a. So ein Fahrzeug wird in den Innenstädten bald verboten sein. |
|     | b. Such a vehicle will soon be banned in the inner cities.      |

The framework in this paper offers a way to spell out the notion of similarity in some detail without being forced to leave the well-established ground of referential semantics. The core idea is to make use of attribute spaces representing complex features of individuals, and to make use of predicates defined on such features

---

<sup>1</sup> The notion of *kinds* in linguistics is closely connected to the notion of *concepts* in psychology (Carlson 2010). Moreover, *ad-hoc categories* formed by linguistic expressions show core characteristics of concepts (Barcalou 1983). We thus assume that kinds formed ad-hoc by similarity expressions closely correspond to concepts, see Umbach & Stolterfoht (in prep).

determining the granularity of representation. In accordance with referential semantics we assume that natural language expressions refer to entities, or categories of entities, in the real world. However, access is only indirect, mediated by *generalized measure functions* mapping real world entities to points in attribute spaces (this is called a *mediated reference theory* in Färber, Svetashova and Harth, this volume). Similarity is a key concept in our framework because it provides a variable notion of *identity / indistinguishability with respect to a representation*: Individuals count as similar if their features in a particular attribute space, given a particular granularity, cannot be distinguished.

This system provides a powerful and flexible tool in the analysis of natural language semantics facilitating detailed interpretations of similarity expressions (*so, such, similar* etc.). Beyond, and maybe even more relevant, this system offers the possibility to analyze linguistic ad-hoc kind formation constructions, for example, by *so/such* demonstratives and equative comparison as in (1) and (2). It is important to realize, however, that this system is basically a multidimensional generalization of degree semantics (e.g. Kennedy 1999) complemented by a method for varying granularity. From this point of view, our framework is anchored in referential semantics just as much as degree semantics is.

Attribute spaces are well-established methods of representation in AI<sup>2</sup> and also in some branches of natural language semantics, e.g., in frame-based approaches (Barcalou 1992; Minsky 1975). What distinguishes attribute spaces and representations as proposed in this paper from classical frame-based approaches is that we focus on systems of predicates on points in attribute spaces in contrast to the points in these spaces themselves, thereby introducing a qualitative aspect, for instance in modelling comparison. This idea is connected to the idea of micro-theories (see, e.g., in Cyc<sup>3</sup> or other ontology languages) which talk about small parts of the world covered, e.g., by a single concept like *chair, vehicle, elephant, human*, etc., but also about actions and events. We expect that such micro-theories provide some kind of prototypes or exemplars, positive and also negative ones. Maybe we just imagine such exemplars. Here is a typical way how to introduce the concept of a physical object in a beginners lecture in experimental physics by imagination of a positive example:<sup>4</sup> “Think of a red steel ball of ten centimeters diameter in front of you. It need not to be red, it need not to be made from steel, it need not have a diameter of ten centimeters and it need not be a ball.” This shows that even abstract concepts can be characterized by exemplars (real or imagined) together with the specification of relevant dimensions in an attribute space.

This paper is structured in the following way: In section 2 we develop a formal theory of representation making use of predicate systems over attribute spaces. Section 3 gives a brief overview over the interpretation of natural language similarity

---

<sup>2</sup> Starting from Minsky’s frames (Minsky 1975) and feature structures, up to modern approaches based on description logics (for an overview see [https://en.wikipedia.org/wiki/Description\\_logic](https://en.wikipedia.org/wiki/Description_logic)).

<sup>3</sup> For micro-theories in Cyc see, e.g., <https://pdfs.semanticscholar.org/4f28/6fdf9280449588b9d3781c9c897da28e0cff.pdf>

<sup>4</sup> For an overview of the imagery debate see <https://plato.stanford.edu/entries/mental-imagery/>

expressions and the role of similarity in ad-hoc kind formation and equative comparison. Since the focus of this paper is on formal characteristics of the representational framework, we will not go into linguistic details.<sup>5</sup> In section 4 we develop a formal similarity concept based on methods provided in section 2. Section 5 shows how to use granularity and hierarchies of representations in order to model gradability along non-scalar dimensions.

## 2 Representations in multi-dimensional attribute spaces

We start from the idea that natural language expressions refer to entities or categories (or even higher order structures, e.g., relations) of entities in the real world, but in an indirect way. Access to these entities or categories is mediated by a function we call *generalized measure function*, e.g.,  $car_1 \Rightarrow \{\text{horse\_power: } 100\text{ps, weight: } 1680\text{kg, color: green } \dots\}$ . This is related to what is called *observables* in physics:<sup>6</sup> Such a function assigns observable attributes (elements of an attribute space) to entities or classes of entities in the world.<sup>7</sup> The referential power of language predicates like *car* (their meaning in the world) can thus be approximated by classifiers. Such classifiers should be effectively computable characteristic functions of predicates.<sup>8</sup> They operate on attribute spaces (or higher order structures based on attribute

---

<sup>5</sup> Readers primarily interested in formal frameworks might skip section 3. Readers primarily interested in semantics might want to start with section 3 and eventually go back.

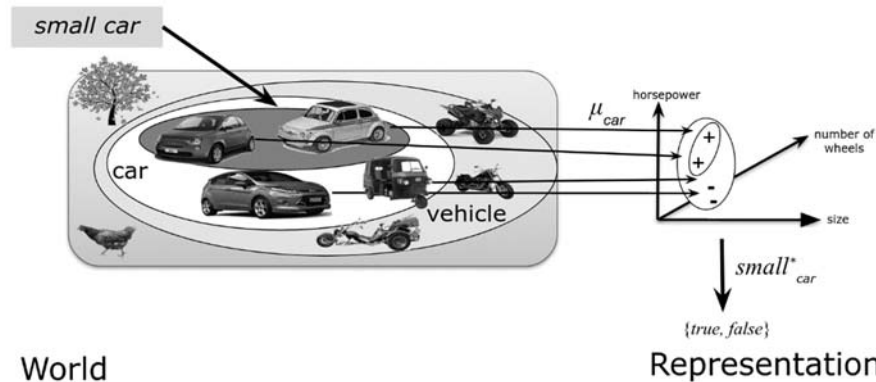
<sup>6</sup> There is a long-standing debate about the dichotomy of observables vs. theoretical terms in philosophy, see <https://plato.stanford.edu/entries/theoretical-terms-science/>. We take a naive view here: observables are functions assigning values to entities in the world which can be determined by ‘simple’ measurements. Examples are *temperature, length, width, height, color, position*, etc., in contrast to values for energy (which in case of heat, for example, depends on temperature, mass and specific heat of the matter).

<sup>7</sup> Our approach is non-constructive since we do not construct representations, but instead have systems of constraints which representations must obey. Bechberger & Kühnberger (this volume) discuss approaches for learning feature space representations by multidimensional scaling. They optimize these representations by using artificial neural networks. From our point of view, they try to learn a feature space  $F$  and a measure function  $\mu$  from similarity and dissimilarity judgments of subjects. In this case,  $\mu$  maps stimuli (elements of a stimuli domain  $D$ ) to points in  $F$ .

Their approach is restricted such that all dimensions of  $F$  have a uniform structure. Essentially  $F$  is an euclidean vector space in their approach. There is no canonical interpretation of the dimensions found, and therefore, no link to natural language expressions. In a second step, the goal is to find classifiers which approximate meaningful subclasses of the stimuli space, which may then lead to interpretations of the dimensions. Bechberger & Kühnberger discuss this as a quality measure suited in determining the number of dimensions of  $F$ . They generalize the approach to handle unseen stimuli.

<sup>8</sup> Classification problems are common in artificial intelligence, where classifiers are trained on huge example sets to be able to classify unseen examples without error. Analogous to our approach, the first step is to find a suitable representation of the real world problems which can be

spaces).<sup>9</sup> Still, we can go back from predicates on points in attribute spaces to predicates on the entities in the world via the inverse image of the generalized measure functions.



**Fig. 1.** A domain of vehicles and a representation featuring positive and negative exemplars of small cars.

On the worldly side, a domain includes a set of relevant predicates  $P$  talking about entities in the world. According to the notion of a representation in this paper, these predicates have counterparts on the representational side marked by a star (\*) in Fig. 1. Counterpart predicates are required to be consistent with their originals; more precisely, they have to agree in truth value on the set of positive and negative exemplars of the original predicate. Moreover, counterpart predicates will be assumed to have convex extensions. As a consequence, they must be true on all points in the convex closure of the images of the positive exemplars (see diagram below). In addition, we stipulate that the extensions of counterpart predicates must be open<sup>10</sup> in some given topology on attribute spaces. This ensures that small changes in the representation (in the sense of the given topology) do not change the truth-values of these predicates.

---

handled by the classification algorithm. Then the example cases have to be translated into this representation in order for the classifier to be able to learn.

<sup>9</sup> We may want to restrict computational complexity of classifiers since there should be efficient algorithms for classification. We will pay with accuracy to get easy to classify areas within the attribute space.

<sup>10</sup> Open sets are sets without a border. Think of a ball in three-dimensional Euclidean space as something like a tomato: It has a crisp border. If we remove the border by peeling, it is unclear where the tomato ends.

## *Domains and representations*

We start the formalization of our approach by introducing domains and representations. For classifiers, given the truth-value  $true \in \Omega$ , we get the extension in the attribute space by its inverse image of  $\{true\}$ , and we get its extension in the real world by applying the inverse image of the measure function. However, given a language predicate like *small* in the context of cars, its reference will in general not be completely determined by a classifier  $small^*_{car}$  and by subsequently applying the inverse image of the measure function. An entity which has all the attributes of a small car may not be a small car, and an entity which is a small car may not have all the attributes we in general assign to cars. In this sense, classifiers approximate the denotation of language predicates. This approximation relation is subject to consistency constraints: If we know that  $x$  is a small car and  $y$  is similar enough to  $x$ , we expect that  $y$  is a small car, too. What should ‘similar enough’ mean? In our approach, we can express this in terms of the attribute space: The attribute values must be similar enough.

If the classifiers cannot discriminate between the representations (points in the attribute space) of two entities  $x$  and  $y$ , they must belong to the same concepts: If one is a small car, then the other must be a small car, too. In particular, this is the case if the representations in the attribute space are equal. Think of a situation where we measure size only with very low precision or specify color only by a few color values. If the above constraint is violated we should probably change our attribute space and/or our measure function, e.g., increase precision of measuring size and/or introduce a more fine-grained color specification.

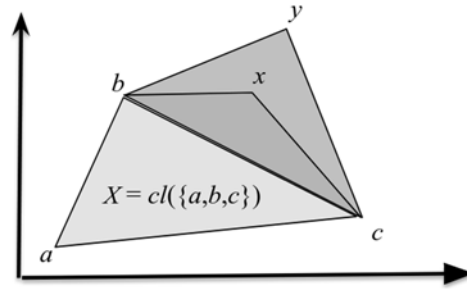
Often we have additional structure on our attribute space, e.g., a (pre)order relation. Assume that  $x$  and  $y$  are small cars, and  $z$  is in the car domain. The number of wheels are  $w_x, w_y, w_z$  respectively;  $x, y$ , and  $z$  differ only in the number of wheels. Then, if  $w_x \leq w_z \leq w_y$ , we expect  $z$  to be a small car, too. If not, we again have an inconsistency in our representation. And again, we probably should change it. The mathematical foundation of this type of inconsistency is the theory of convex closures. The formal definition of a convex closure operator  $cl$  on a set  $X$  is the following (see Korte et al. 1991):

A function  $cl : \wp(X) \rightarrow \wp(X)$  is a convex closure operator iff

- it preserves the empty set  $cl(\{\}) = \{\}$
- it is extensive  $A \subseteq cl(A)$  for all  $A \subseteq X$
- it is monotone  $A \subseteq B \rightarrow cl(A) \subseteq cl(B)$
- it is idempotent  $cl(cl(A)) = cl(A)$
- the anti-exchange property holds  $x, y \notin X, x \neq y,$   
 $x \in cl(X \cup \{y\}) \rightarrow y \notin cl(X \cup \{x\})$

In the two-dimensional Euclidean plane, we can visualize the effect of a convex closure operator. Suppose  $X$  is  $cl(\{a, b, c\})$ . If  $x$  is in  $cl(X \cup \{y\})$ , then  $y$  cannot be in

$cl(X \cup \{x\})$ . The anti-exchange property ensures convexity. In a two-dimensional Euclidean plane, this means that for any two points in  $X$  the connecting line must also be in  $X$ .



**Fig. 2.** Convex closure and anti-exchange property in the Euclidean plane

On a (partially) ordered set  $(M, \leq)$  we can define convex closure operators in a natural way. For  $A \subseteq M$  we define:

$$\begin{array}{ll} \text{left closure:} & cl_{\leftarrow}(A) = \{x \in M \mid \exists y \in A : x \leq y\} \\ \text{right closure:} & cl_{\rightarrow}(A) = \{x \in M \mid \exists y \in A : y \leq x\} \\ \text{convex closure:} & cl(A) = \{x \in M \mid \exists y, z \in A : y \leq x \leq z\} \end{array}$$

To sum up: We approximate the meaning of natural language predicates by classifiers and their inverse images by means of a generalized measure function. Additionally, we request that classifiers respect some consistency constraints: (i) they should classify known examples correctly, (ii) their extension (as a subset of the attribute space) should be convex according to a suitable convex closure operator and (iii) their extensions should be open in a suitable topology. The topology and the closure operator must be compatible: Closures of open sets must be open.

First, we need a notation to refer to the entities we are talking about by a natural language predicate like *small car*: the set of entities (in the world) for which it makes sense to ask if they have car properties, that is, entities for which the attribute dimensions for cars make sense, e.g., *number of wheels*, *horsepower*, *size*, *weight*, *color* etc. We exclude entities for which it does not make sense to ask if they have car properties, e.g., single atoms, trees, hens etc.

Next, we assume that we have clear cases: positive examples such as entities which are definitely cars, and negative examples such as entities for which the attribute dimensions of cars make sense but which are definitely not cars, e.g., motorbikes. Concepts which are related and belong to the same micro-theory are collected as predicates over the same domain. Think of different types of cars, bikes, trikes etc.

We assume that there is a universe  $U$  which includes all the entities in the world. We can start now formalizing our approach by defining a *domain* as a subset of the

universe  $U$  together with a set of predicates and non-overlapping sets of positive and negative examples for each predicate.

**Definition 1: Domain**

A domain  $\mathcal{D}$  is a quadruple  $\langle D, \_+, \_-, P \rangle$  with:

- $D \subseteq U$  a set of individuals / entities (called the carrier of the domain),
- $P = \{p_1, \dots, p_n\}$  a set of identifiers of predicates over  $D$ ,
- $\_D : P \rightarrow \wp(D)$  the extension in  $D$  of a predicate<sup>11</sup> denoted by index  $D$ <sup>12</sup>
- $\_+ : P \rightarrow \wp(D)$  a function which assigns a set of positive examples to each predicate (for  $\_+(p)$  we write  $p^+$ ),
- $\_- : P \rightarrow \wp(U)$  a function which assigns a set of negative examples to each predicate<sup>13</sup> (for  $\_-(p)$  we write  $p^-$ ),
- $\forall p \in P : p_D^+ \cap p_D^- = \emptyset$  (consistency),
- $\exists q \in P \forall p \in P : p_D^+ \subseteq q_D^+ \wedge p_D^- \subseteq q_D^- \wedge q_D^- \cap D = \emptyset$  (universal predicate).

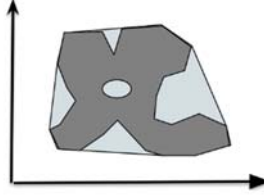


Fig. 3. A non-convex set in the two-dimensional plane and its convex closure.

### Representations and classifier systems

We view the elements of  $D$  as entities to which we have only indirect access via a (generalized) measure function  $\mu$ . The measure function  $\mu$  constructs representations of the entities in  $D$  as points in an attribute space  $F$ , much like observables in

<sup>11</sup> In fact, we will often use characteristic functions in place of predicates. In the structures we are interested in, there is an isomorphism between  $\wp(D)$  and  $\Omega^D$ . We will not restrict ourselves to a special type of logic (e.g. two-valued classical logic). We stipulate a logical system characterized by a set of truth-values  $\Omega$ .  $\Omega = \{true, false\}$  for classical logic,  $\Omega = [0, 1]$  for fuzzy logic.

<sup>12</sup> We will drop the index  $D$  whenever it is clear which domain we are talking about.

<sup>13</sup> Positive examples must be in the domain, negative examples may be anywhere. A small mouse is a negative example for ‘big elephant’, but a small elephant is a more informative example.



physics. Attribute spaces are well-established representational structures.<sup>14</sup> They generalize vector space approaches in allowing heterogeneous dimensions equipped with value sets of different scales (nominal, ordinal, interval, proportional, partially ordered etc.), where value sets may themselves be attribute spaces with multiple dimensions.

An attribute space  $F$  is given by a set of attributes  $A = \{a_1 \dots a_n\}$ , such that for each  $a_i$  in  $A$  there is a set of possible values  $V_{a_i}$  of  $a_i$ . Elements of  $D$  are mapped to points in  $V_{a_1} \times \dots \times V_{a_n}$ , the carrier of the attribute space  $F$ . Think, for example, of *number of wheels* as an attribute with  $\{1,2,3,4,5,6,\dots\}$  as its value set, or *horsepower* as an attribute with the positive real numbers as its value set.<sup>15</sup>

A *representation* includes an *attribute space*  $F$ , a (generalized) *measure function*  $\mu$  mapping elements of a domain into the attribute space, and a set of *classification functions*  $p^*$  applying to points in the attribute space. In the case of the attribute *number of wheels* the measure function  $\mu$  just has to count. In the case of the attribute *horsepower* a complex measurement procedure is required to determine the value of  $\mu$ . The classification functions (short *classifiers*) serve as *approximations*<sup>16</sup> of the predicates in  $P$ .<sup>17</sup> Moreover, the extensions of the classifiers will be assumed to be open and convex. This means that  $F$  comes with a convex closure operator  $cl$  and  $p^*$  must be *true* on  $cl(\mu(p^+))$ .<sup>18</sup> Using the  $n$ -dimensional Euclidean space as an example, the extensions of the classifiers must not have holes, notches or coves in the representation space  $F$ .

**Definition 2: Representation**

A representation  $\mathcal{F} = \langle \langle F, cl \rangle, \mu, \_*, \mathcal{D} \rangle$  of a domain  $\mathcal{D} = \langle D, \_+, \_-, P \rangle$  is given by

- an attribute space  $F$  together with a closure operator  $cl$  and a compatible topology (we write  $F$  for  $\langle F, cl \rangle$  if we are not interested in the closure operator  $cl$ ),
- a measure function<sup>19</sup>  $\mu : D \rightarrow F$ ,

<sup>14</sup> Attribute spaces are related to the classical frame approaches (Minsky 1975). Other related approaches are feature structures which are widely used in linguistic formalisms (Carpenter 1992).

<sup>15</sup> Note that ordinal or metric dimensions as common in degree semantics correspond to one-dimensional attribute spaces in our approach.

<sup>16</sup> More precisely:  $p^* \circ \mu$  approximates  $p$ .

<sup>17</sup> For every  $p \in P$  there is a  $p^* \in P^*$

<sup>18</sup> This includes all points in the convex closure of the images of the positive exemplars. For the concept of convexity in conceptual structures see Gärdenfors (2000). Intuitively, the convex closure of a subset  $X$  of  $F$  is the smallest convex subset of  $F$  containing  $X$ .

<sup>19</sup> In most cases, we do not expect to explicitly compute values of the measure function for entities in  $D$ . Almost no one will be able to compute the horse power of his car. To learn about the horse power of my car I would look-up the value in the data sheet. When you go to the doctor for a general health check-up the chance that she will take a measure stick to measure your height is very small. It might instead be like this: doctor: "How tall are you?", patient: "As tall as you.", doctor: "About 1.75?", patient: "Think so." Nevertheless, it should at least in principle be possible

- a function  $\_ * : P \rightarrow \Omega^F$   
(we write  $p^*$  for  $\_ *(p)$  and call them classifiers).<sup>20</sup>

Representations are subject to three consistency constraints:

- $\forall p \in P$  the extension of  $p^*$  must be open and convex in  $\langle F, cl \rangle$
- $\forall p \in P \forall x \in p^+ : p^*(\mu(x)) = true$
- $\forall p \in P \forall x \in p^- \cap D : p^*(\mu(x)) = false$

From this we get  $\mu(p_i^+) \cap \mu(p_i^- \cap D) = \emptyset$ .

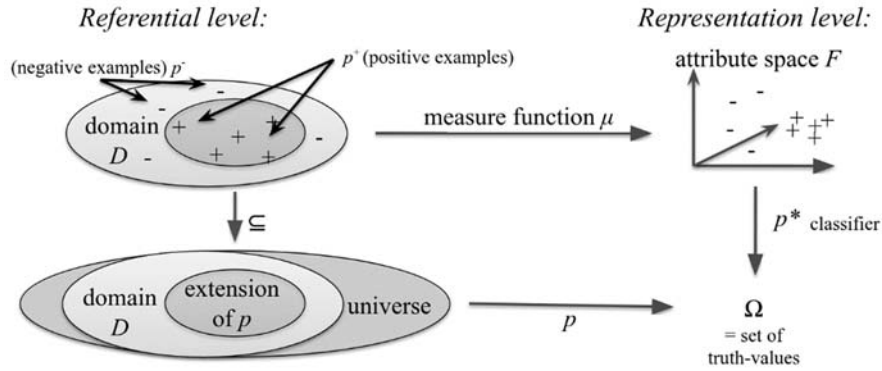


Fig. 4. Domains and representations

As mentioned above, attribute spaces are familiar methods of representation. What distinguishes attribute spaces from the representations proposed in this paper is the idea of classifiers on attribute spaces. On the worldly side, a domain includes a set of relevant predicates  $p \in P$ . On the representational side, these predicates have counterparts, namely classifiers  $p^* \in P^*$ . By  $P^*$  we denote the set of all basic classifiers:  $P^* = \{p^* | p \in P\}$ . These classification functions are required to be consistent with their corresponding predicates over  $D$ ; more precisely, for the set of positive / negative exemplars the truth-values of the classification functions have to agree with the truth-values of the original predicates (see definition 2).

Given a set of basic classifiers,<sup>21</sup> we assume the possibility to construct derived classifiers by logical operations: For the logical conjunction this is unproblematic

to determine the value for a given element in  $D$ . It is even possible to use of machine learning technics to learn suitable dimensions and values by analyzing similarity judgments of subjects (see footnote 7).

<sup>20</sup> where  $\Omega^F$  is the set of characteristic functions  $F \rightarrow \Omega$ . In addition, we expect that classification functions come with algorithmic methods to compute these functions.

<sup>21</sup> There is an interaction between the attribute space  $F$  and the measure function  $\mu$ . While attribute spaces can provide highly structured representations, classifiers can be viewed as attributes with

(convex sets and open sets are closed under intersection). For the logical disjunction we have to apply the convex closure operator  $cl$  to the result. For negation this is not possible. Thus we do not allow to define complex classifiers by applying negation to elementary ones.<sup>22</sup> We name the set of derived classifiers  $\tilde{P}^*$ .

**Definition 3: Classifier systems**

Given a set of basic classifiers  $B$  over an attribute space  $F$ , we define a set of classifiers  $\tilde{B}$  inductively (much like a topology):

- $B \subseteq \tilde{B}$  we expect that elements of  $B$  are convex and open,
- $X \in \tilde{B}, Y \in \tilde{B} \rightarrow X \cap Y \in \tilde{B}$  intersections,
- $X \in \tilde{B}, Y \in \tilde{B} \rightarrow cl(X \cup Y) \in \tilde{B}$  closures of unions.

If  $F$  is (partially) ordered:

- $X \in \tilde{B} \rightarrow cl_{\rightarrow}(X) \in \tilde{B}$   $cl_{\rightarrow}(X) = \{x \in F \mid \exists y \in X : y \leq x\}$  right closures,
- $X \in \tilde{B} \rightarrow cl_{\leftarrow}(X) \in \tilde{B}$   $cl_{\leftarrow}(X) = \{x \in F \mid \exists y \in X : x \leq y\}$  left closures.

It is important to mention that in general  $\tilde{B}$  is not closed under complement. This means that we do not have negation: Complements of convex sets need not to be convex and complements of open sets need not to be open. We start with basic classifiers  $B = \{p_1^*, \dots, p_n^*\}$  and get  $\tilde{P}^*$  as the corresponding system of classifiers.

### 3 Similarity expressions in natural language

In this section, a brief overview will be given of the challenges involved in the interpretation of similarity expressions. This section will not give a full description of the semantic phenomena – references will be given for details – but instead serve as a motivation for the specifics of the similarity framework presented in this paper.

#### *Similarity demonstratives*

The need for a framework that models similarity originated from the problem of how to interpret the German demonstrative *so* ('so'/'such'). It is a genuine

---

values in  $\Omega$ . It is possible to hide all the complex structure of a representation in the measure function by using  $(p_1^* \times \dots \times p_n^*) \circ \mu$  as new measure function and  $\Omega^n$  as attribute space  $F$ . Of course that is not the idea of this approach. We will try to use ‘simple’ measure functions and meaningful attribute dimensions.

<sup>22</sup> In general, complements of concepts are not necessarily themselves concepts – a non-car is not a proper concept.

demonstrative expression, so we expect direct reference in the sense of Kaplan (1989). It does not, however, express identity as does, e.g., *dies/this*, and instead it refers to a set of entities which are in some sense similar to the target of the demonstration gesture (the entity the speaker points to). If the speaker points to a car while uttering "*So ein Auto hat Anna*" ('Anna has a car like this'), Anna's car is said to be, with respect to a particular set of features, indistinguishable from the car the speaker points to. This kind of demonstrative expressions is called *similarity demonstratives* in Umbach & Gust (2014), Gust & Umbach (2016), and *demonstratives of manner, quality and degree* in König & Umbach (2018).

We follow Nunberg's (1993, 2004) adaptation of the Kaplanian analysis, interpreting demonstratives as directly referential expressions, but at the same time dismissing the idea that the target of the demonstration is necessarily identical to the referent of the demonstrative. This allows for a straightforward interpretation of similarity demonstratives such that the target of the demonstration is the individual or event the speaker points to, and the referent of the demonstrative phrase is related to the target by similarity instead of identity. Similarity is then implemented by indistinguishability of points in attribute spaces (see section 4). This implementation of similarity is in fact close to the idea of contextual granularization suggested in Nunberg (2004): When restricting attention to a particular set of features, it may be the case that two entities can no longer be distinguished. It is important to note, however, that this idea requires a framework that distinguishes between a referential and a representational level – you cannot speak about indistinguishability without access to what could have been distinguished.

### ***Ad-hoc kinds***

According to the similarity analysis, demonstratives like German *so* and English *such* create classes of similar items, e.g. similar cars. There is some evidence that in the nominal and verbal case (though not in the adjectival case) these similarity classes constitute ad-hoc kinds. In a nut-shell, *so/such* phrases can be shown to be restricted to particular features of comparison. For example, the feature *number of doors* would be perfect when comparing cars but not when comparing mugs – mugs do not have doors, so the number of doors does not qualify as a feature of comparison for mugs. But mugs as well as cars can be recently purchased and nevertheless *being recently purchased* does not qualify as a feature of comparison for neither cars nor mugs. This suggests that properties qualifying as features of comparison must not be accidental.

There is experimental evidence that features of comparison are restricted to properties which are neither accidental nor evaluative (see König & Umbach 2018 and Umbach & Stolterfoht in prep.). This raises the question of how to characterize these properties, which is a prominent issue in the debate about concept formation in cognitive psychology. Only recently has this debate been connected to the topic of

genericity in linguistics by Greenberg (2003) and Carlson (2010), and by the experimental studies in Prasada & Dillingham (2006) and by Prasada et al. (2013), providing evidence that there are so-called *principled connections* between kinds and properties that an entity has, because it is the kind of thing it is.

There is an alternative analysis claiming that demonstratives like German *so* and English *such* are pro-kind expressions (see Anderson & Morzycki 2015, adapting Carlson's 1980 kind-referring analysis of *such*). The final results of the two accounts are fairly close. However, unlike the pro-kind account, the similarity account not just postulates that *so/such* phrases denote kinds, but in addition shows how these kinds emerge, namely by similarity.

### *Equative comparison*

Another phenomenon where similarity plays a significant role is equative comparison, including non-scalar as well as scalar cases, see (3a-c).<sup>23</sup> In German, scalar as well as non-scalar equatives are uniformly constructed by *so ... wie* where *so* is a correlative pronoun relating to the standard of comparison given in the *wie* clause:

- |     |    |   |                      |
|-----|----|---|----------------------|
| (3) | a. | Anna ist so groß wie Berta.<br>Anna is as tall as Berta         | scalar / adjectival  |
|     | b. | Anna hat so ein Auto wie Berta.<br>Anna has a car like Berta's. | non-scalar / nominal |
|     | c. | Anna tanzt so wie Berta.<br>Anna is dancing just like Berta     | non-scalar / verbal  |

Given that the demonstrative *so* can in general be substituted by *wie dies* ('like this'), it suggests itself to analyze *wie* as expressing similarity as does *so*, though without a deictic component. This allows for a generalized account of equative comparison: The nominal equative in (3b) is interpreted such that Anna's car is similar to Berta's car with respect to a set of contextually given features; the verbal case in (3c) is interpreted such that the event of Anna dancing is similar to the event of Berta

---

<sup>23</sup> It has been argued that (3a) and (3b,c) just differ in being one-dimensional as opposed to multi-dimensional, and that even multi-dimensional comparison is scalar. There are, in fact, multi-dimensional adjectives like *healthy* that allow for comparatives: *A is more healthy than B*. Sassoon (2013) suggests to interpret comparatives of multi-dimensional adjectives by quantification over dimensions in which the compared entities exceed the standard: A is more healthy than B iff the number of dimensions in which A exceeds the standard is greater than that of B exceeding the standard (for alternatives see the subsection on gradability below).

This approach presupposes, however, that the individual dimensions are scalar, which is not generally the case, consider, e.g., color as a dimension in comparing cars or posture as a dimension in comparing dancing habits. Moreover, even though cars and dancing habits can be compared in equatives, forming comparatives is impossible. This is strong evidence that (3b,c) are genuinely non-scalar.

dancing; and the adjectival case in (3a) is interpreted such that Anna is similar to Berta with respect to their height – note that the scalar equative in (3a) does not hinge on contextually given features of comparison but instead 'carries its dimension on its sleeves'.

### ***'Exactly' versus 'at-least' reading***

Scalar equatives like (3a) allow for two readings. On the *exactly* reading, Anna's height is (approximately) the same as Berta's height, while on the *at-least* reading Anna's height is greater than or equal to Berta's height. While both readings are attested in the data, standard degree semantics and the similarity analysis differ with respect to which reading is predicted to be primary. In standard degree semantics equatives are assumed to have an *at-least* interpretation as their meaning while the *exactly* reading is derived by scalar implicature. In the similarity analysis, on the other hand, equatives (scalar as well as non-scalar) are interpreted such that their meaning is symmetric, since similarity is an equivalence relation – *A ist so groß wie B* means that *A* is similar in height to *B* – thereby raising the question of how to account for the *at-least* reading.

The question of which of the *exactly* and the *at-least* reading is basic has been the topic of a continuous debate when addressing numeral expressions. According to the classic analysis by Horn (1972), sentences containing numbers assert lower boundedness and may, depending on the context, implicate upper boundedness – *Anna has three sheep* asserts that she has at least three sheep and implicates, depending on context, that she has at most three sheep. This analysis has been questioned, for example, by Kennedy (2013) who presents, among other things, scope effects that cannot be explained in the classic analysis. Surprisingly, this debate has not been extended to equative constructions, even though according to the classic analysis degree equatives assert *at-least* interpretations, as in the case of Horn's analysis of numerals: *Anna is as tall as Berta* is true if  $\text{height}(\text{Anna}) \geq \text{height}(\text{Berta})$  (see., e.g., Kennedy 1999).

We assume that the semantics of scalar equatives is given by similarity even in contexts requiring an *at-least* reading, and we implement this idea by exploiting the granularity encoded in our framework. Consider the example in (4). In this context, Sophie tells the truth even if she is taller than Larissa. In general, if there is a threshold given in the context, it appears irrelevant by how much it is exceeded.

- (4) Sophie wants to join the police, which requires a certain minimum height. Her cousin Larissa has told their grandma that she has already been accepted by the police. That's why grandma asks Sophie whether she is as tall as Larissa.  
Sophie replies: Ja, ich bin so groß wie Larissa / Yes, I'm as tall as Larissa.

In the case of *at-least* readings, classifiers applying to the standard of comparison, e.g., Larissa's height in (4), are mapped to their right closure.<sup>24</sup> Thereby Sophie counts as similar in height to Larissa even if she is ten centimeters taller. Thus our account is "mildly ambiguous" – in particular contexts, closures involved in determining similarity are adjusted. It has to be noted, though, that this adjustment is licit only if the difference is moderate. But if, for example, Larissa is a six-year-old and Sophie is her mother, it would be absurd to assert that Sophie is as tall as Larissa (which is predicted to be true on the classical analysis of degree equatives).

For negated scalar equatives the prominent reading is asymmetrical: The sentence *Anna ist nicht so groß wie Berta* / *Anna is not as tall as Berta*. is preferably interpreted such that Anna is smaller than Berta. This asymmetry is not influenced by the existence of a contextual threshold and does not appear infelicitous in the case of major differences – *Larissa is not as tall as Sophie* would be acceptable even if Sophie is Larissa's mother. The preference for the asymmetric reading of negated scalar equatives can be explained by the fact that a disjunctive (symmetric) reading according to which Anna is either smaller or taller than Berta would not be convex any longer. Given that convexity plays a primary role in cognitive economy it is hardly surprising to find such effects in natural language semantics (see also Solt & Waldon 2019 on numerals under negation).

## ***Gradability***

Implementing similarity as indistinguishability (see the next section) suggests that it is a nongradable concept. This is plausible considering expressions like German *so/wie* and English *such/like*. On the other hand, the adjectives *ähnlich* and *similar* are gradable – Anna can be more similar to her father than to her mother. This points to the need for a gradable notion of similarity.

Cognitive Science models of similarity usually start out either from a notion of distance in a geometrical space (e.g. Gärdenfors 2000) or from numbers of common and distinctive features (e.g. Tversky 1977). Both approaches facilitate a straightforward definition of the comparative: In geometric models similarity increases if distances decrease, and in feature based models similarity increases if the number of common features increases and that of distinctive features decreases. However, the positive form – the predicate *similar* – would require a threshold from where on two items count as similar, which would be hard to provide in a non ad-hoc fashion.

In our system, the positive form is the primary one – two items are similar if indistinguishable with respect to a given representation (including dimensions of comparison and classifiers, see definition 2 and definition 4 and 5). The comparative will be defined making use of representations of different granularity: Two items *a*

---

<sup>24</sup> See the *quasi exactly* implementation of the *at-least* reading by right closure of classifiers in section 4.

and  $b$  are more similar than two items  $c$  and  $d$  in a representation  $\mathcal{F}$  if and only if there is a less granular representation  $\mathcal{F}'$  such that  $a$  and  $b$  are similar in  $\mathcal{F}'$  while  $c$  and  $d$  are not (see definition 8 in section 5). Suppose, for example, that in representation  $\mathcal{F}$  neither  $a$  and  $b$  nor  $c$  and  $d$  are similar. If there is a less granular representation  $\mathcal{F}'$  such that  $a$  and  $b$  are similar while  $c$  and  $d$  can still be distinguished, then  $a$  and  $b$  must be closer in terms of properties than  $c$  and  $d$ .

Defining a comparative notion *more similar* based on the positive form *similar* is reminiscent of the vague-predicate approach suggested by Klein (1980). In contrast to the standard degree-semantic approach where degrees are compared in interpreting the comparative – *Anna is taller than Berta* is true if her degree of height exceeds that of Berta – in a Kleinian approach the comparative is modelled by varying contexts, that is, varying thresholds for the positive predicate to apply: *Anna is taller than Berta* is true if there is a context such that Anna counts as tall while Berta does not.<sup>25</sup> This way of interpreting the comparative is, first of all, consistent with cross-linguistic findings showing that the majority of languages express the comparative in terms of the positive. Moreover, it does not rely on the existence of a single scale of degrees.

The definition of *more similar* suggested above gives us the means to interpret the comparative form of the adjective *similar*. But beyond that it allows a Kleinian style definition of comparatives for multi-dimensional adjectives like *healthy* and *beautiful*. Comparatives of multi-dimensional adjectives are usually interpreted using degree semantics, either by counting dimensions in which the threshold is exceeded (see Sassoon 2013), or by integrating dimensions such that the result forms an order, where integration may be context-dependent and also judge-dependent (see Solt 2016).

The similarity framework puts us in the comfortable position of not having to treat all adjectives in the same way. Adjectives like *tall* and *old*, which clearly refer to a single ordinal or even metric scale, will be interpreted via a single dimension. In this case, similarity takes the role of specifying the granularity of this scale: *Anna is taller than Berta* is true if all points of the granule of Anna's height are greater than all points of the granule of Berta's height (in the case of overlapping the situation is more complex). Multi-dimensional adjectives like *healthy* and *beautiful*, on the other hand, will be interpreted by similarity to a prototype:<sup>26</sup> *Anna is healthy* is true if Anna's health is similar to the prototype. And *Anna is more healthy than Berta* is true if Anna's health is more similar to the prototype than Berta's health.

---

<sup>25</sup> Contexts have to be consistent with the order of individuals in the domain.

<sup>26</sup> Analogous to thresholds in a single dimension – context-dependent and maybe judge-dependent.



## 4 Indiscernability

In order to realize that two entities in the world are different their representations must differ in some way. This means that they must be recognizably different. In our approach this means that there are classifiers which can discriminate them. The complementary situation is indistinguishability, which means that, on the representational level, we cannot discriminate them. In our approach, given a system of predicates  $P$  there are two reasons why we may not be able to distinguish two elements of  $D$ :

- Two elements may lead to the same value of the function  $\mu$ , i.e., the same point in the attribute space. Then no classifier can discriminate between the two elements.
- The two elements disagree on  $\mu$  (so we see that they are different), but they agree on all classifiers in  $\tilde{P}^*$ .

To account for these types of indistinguishability we borrow the term *indiscernible* from Rough Set Theory (Pawlak 1998):

**Definition 4: Indiscernible**

Given a representation  $\mathcal{F} = \langle F, \mu, \_*, \langle D, \_+, \_-, P \rangle \rangle$  we define:  
For  $x, y \in F$ :  $x \sim_{\mathcal{F}} y \equiv \forall q \in \tilde{P}^* : q(x) \leftrightarrow q(y)$

where  $\tilde{P}^*$  is the set of all derived classifiers.

According to this definition, indiscernibility is relative to the classifiers in  $\tilde{P}^*$  in a representation  $\mathcal{F}$ . The relation of indiscernibility talks about points in  $F$ . However, the similarity relation we are interested in talks about elements of the domain  $D$ . Therefore, we have to apply the measure function before checking indiscernability. This gives us a first simple similarity relation:

**Definition 5: Similar**

$\forall x, y \in D : sim(x, y, \mathcal{F}) \equiv \mu(x) \sim_{\mathcal{F}} \mu(y)$

Obviously, definition 5 defines an equivalence relation on  $D$  and we get a partition of the domain. The indiscernibility relation provides attribute spaces with a level of granularity, facilitating comparison of attribute spaces of distinct granularity which are otherwise identical. Let  $[y]$  denote the equivalence class (similarity class) of  $y$ :  $[y] = \{x \mid x \sim_{\mathcal{F}} y\}$ . In Rough Set theory, such equivalence classes are called granules.

There is a problem with this definition of similarity: The similarity classes in the attribute space may not be convex, as the following example shows. Think of case (3a) *Anna ist so groß wie Berta* ('Anna is as tall as Berta.'). Assume that we have a dimension of height (measured in meter) in the attribute space and classifiers which

specify height with some granularity depending on the measured value: A height of 1.80 is given by some value between 1.78 and 1.82, while a height of 1.81 is given by some value between 1.806 and 1.814, and so on. Therefore, we may not be able to discriminate between 1.80 and 1.815: both belong to the same granule [1.80]. Nevertheless, we can discriminate between 1.80 and 1.81 since we have a classifier [1.81] giving *true* on 1.81 and *false* on 1.80. Therefore, the granule of Berta's height ( $y$ ) in Fig. 5, which is equal to [1.80]) may be not convex because [1.81] forms a hole. This results in the following situation: If Berta's height is 1.80, then Anna's height may be 1.80 or 1.815 but not 1.81 in order for the sentence to be true (as demonstrated in Fig.5). This is counterintuitive.

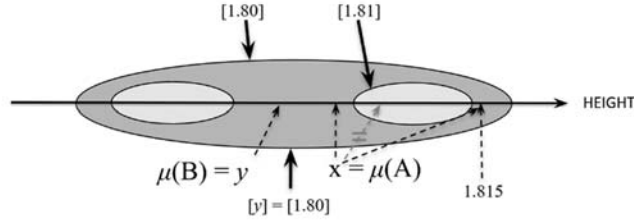


Fig. 5. Granules with holes

We can solve this problem by introducing a new parameter in the definition of the similarity relation: *similarity relative to a point of reference*. This point of reference determines the granules to be selected.

**Definition 6:** *Similarity relative to a point of reference*

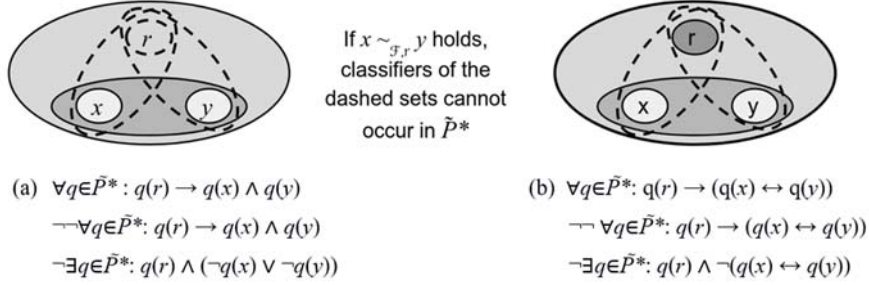
Given a representation  $\mathcal{F} = \langle\langle F, cl \rangle, \mu, \_*, \langle D, \_+, \_-, P \rangle\rangle$ , we can define a similarity relation relative to a point of reference  $r$  in two different ways:

- $$\forall x, y \in F : x \sim_{\mathcal{F}_r} y$$
- (a) iff  $\forall q \in \tilde{P}^* : q(r) \rightarrow q(x) \wedge q(y)$
  - (b) iff  $\forall q \in \tilde{P}^* : q(r) \rightarrow (q(x) \leftrightarrow q(y))$

Definition 6a means that principal filters<sup>27</sup> of  $x$  and  $y$  in  $\tilde{P}^*$  contain the principal filter of  $r$ . In contrast, definition 6b means that elements of the principal filter of  $r$  in  $\tilde{P}^*$  cannot discriminate between  $x$  and  $y$ . It is easy to see that (a)  $\Rightarrow$  (b), but not (b)  $\Rightarrow$  (a).

For an intuitive insight into the functionality of this type of similarity relation, have a look at the following Venn diagrams:

<sup>27</sup> The principal filter of  $x$  is  $\{q \in \tilde{P}^* \mid q(x)\}$ .



**Fig. 6.** If dashed sets occur in  $\tilde{P}^*$ ,  $x$  and  $y$  cannot be similar

Assume that there are four classifiers in  $\tilde{P}^*$ : *small\**, *big\**, *normal\** (concerning size), and *heavy\** (concerning weight). Table 1 shows some possible classifications of  $x$ ,  $y$ , and  $r$ . These possibilities correspond to the dashed sets in Fig. 6. The last two columns show the truth-values of the two similarity relations (a) and (b) in definition 6 for the different cases. All the other cases can be handled by symmetry; only *heavy\** varies. The interesting case is line (2) since the two similarity relations differ: If  $y$  is small but  $r$  and  $x$  are not, and  $x$  is big but  $r$  and  $y$  are not, and  $x$  and  $y$  are normal but  $r$  is not, and  $r$  is heavy but  $x$  and  $y$  are not, then similarity of  $x$  and  $y$  with respect to the reference point  $r$  is *true* according to definition 6b but *false* according to definition 6a. Intuitively, if the properties of the reference point  $r$  differ substantially from the properties of  $x$  and  $y$  then definition 6a gives *false* while 6b gives *true*. We consider definition 6a more plausible than 6b.

|     | <i>small*</i> | <i>big*</i> | <i>normal*</i> | <i>heavy*</i> | $x \sim_{\mathcal{F},r} y$ |       |
|-----|---------------|-------------|----------------|---------------|----------------------------|-------|
|     | $rxy$         | $ryx$       | $rxr$          | $ryr$         | (a)                        | (b)   |
| (1) | --+           | -+-         | -++            | -**           | true                       | true  |
| (2) | --+           | -+-         | -++            | +-            | false                      | true  |
| (3) | --+           | -+-         | -++            | ++            | false                      | false |
| (4) | --+           | -+-         | -++            | ++            | false                      | false |
| (5) | --+           | -+-         | -++            | +++           | true                       | true  |

**Table 1.** Similarity of two points  $x$  and  $y$  in the attribute space with respect to a reference point  $r$  depending on the possible extensions of the predicates *small\**, *big\**, *normal\**, and *heavy\**. The cell [*small\**, (1)], for example, indicates that *small\**( $r$ ) is false, *small\**( $x$ ) is false and *small\**( $y$ ) is true. The cell [*heavy\**, (1)] indicates that *heavy\**( $r$ ) is false, *heavy\**( $x$ ) is either true or false, and *heavy\**( $y$ ) is also either true or false.

For given  $\mathcal{F}$  and  $r$  the relation  $x \sim_{\mathcal{F},r} y$  is a (kind of local) equivalence relation. If we switch the reference  $r$ , the classes will obviously change. If we choose one of the arguments as point of reference, we get an asymmetric similarity relation: In general  $x \sim_{\mathcal{F},y} y$  will be different from  $y \sim_{\mathcal{F},x} x$  because the point of reference changes.

**Definition 7: Similarity classes**

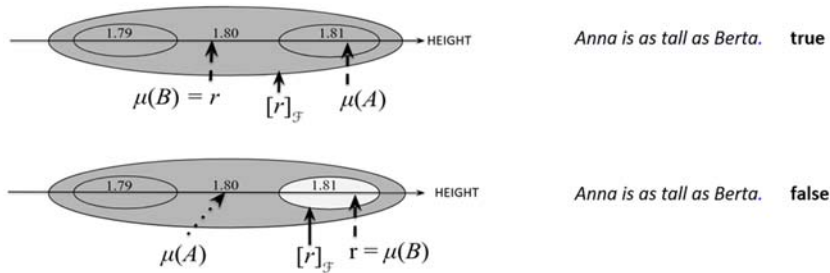
For given  $\mathcal{F}$  and  $r$  we define the similarity class of  $r$  as

$$(a) [r]_{\mathcal{F}} = \{x \mid \forall q \in \tilde{P}^* : q(r) \rightarrow q(x)\}$$

For  $[r]_{\mathcal{F}}$  we borrow the term granule from Rough Set theory. Again we can use the inverse image of the measure function to define similarity relations on the domain. For  $a, b \in D$  we define two different similarity relations. The one in (b) makes use of a point of reference  $r$  that is independent of either  $a$  or  $b$ , whereas in (c) the point of reference is identical to the second argument:

- (b)  $sim_r(a, b, \mathcal{F})$  iff  $\mu(a) \sim_{\mathcal{F},r} \mu(b)$  (+transitive, +symmetric, -reflexive)
- (c)  $sim'(a, b, \mathcal{F})$  iff  $\mu(a) \sim_{\mathcal{F},b} \mu(b)$  (-transitive, -symmetric, +reflexive)<sup>28</sup>

If we again look at our example (3a) *Anna ist so groß wie Berta* ('Anna is as tall as Berta.') we see that the granules depend on the point of reference  $r$  (Fig.7). If we use  $sim'$  from definition (7c), there are two possible situations. In the first situation, we get the information that the height of Berta is 1.80. Since Berta provides the reference point (definition 7c) the relevant granule is  $[1.80]$ . The height of Anna can be an arbitrary value in this granule to make the statement true. It maybe 1.80 or 1.81 – we simply cannot discriminate between both cases because the granule  $[1.80]$  is convex (no holes). In the second situation, we get the information that the height of Berta is 1.81. Now the relevant granule is  $[1.81]$  and not  $[1.80]$  even though 1.81 may be an element of  $[1.80]$ . The height of Anna is restricted to the relevant granule: 1.80 is not a possible value any longer, it falsifies the statement. Although it seems that there is a hole in  $[1.80]$  in the second case, in both cases, the relevant granule is convex.



**Fig. 7.** The effect of holes

<sup>28</sup>  $sim'$  uses the second argument as point of reference.

### **(A)symmetry of similarity**

The notion of *similarity relative to a reference point* is reminiscent of the question of whether the predicate *similar* is symmetrical addressed by Tversky (1977) and also Gleitman et al. (1996).

Tversky's seminal paper on feature-based similarity starts with empirical observations indicating problems of the then predominant geometric notion of similarity and the basic axioms of metric distance:<sup>29</sup> (i) minimality is problematic in view of results concerning the identification probability for identical stimuli, (ii) symmetry is apparently false – the judged similarity of North Korea to Red China exceeds the judged similarity of Red China to North Korea – and (iii) triangle inequality is hardly compelling – Jamaica is similar to Cuba (geographical proximity) and Cuba is similar to Russia (political affinity) but Jamaica and Russia are not similar at all.

However, a closer look reveals that these findings are not generally valid. Before dismissing transitivity of the similarity relation on the basis of the Jamaica/Cuba/Russia example, one should consider the role of switching features within the two comparison steps.<sup>30</sup> And before dismissing symmetry, which is frequently done in the Cognitive Science literature, one should consider the study in Gleitman et al. (1996) and, first of all, Tversky's original study.

In Tversky's study, the linguistic presentation was directional (*North Korea is similar to Red China*), and he himself argues that the asymmetry finding hinges on the directional way of presentation. If the task is to assess the degree to which *A* is similar to *B*, then features of *A* may weigh more heavily than those of *B*.<sup>31, 32</sup> But if the task is to assess the degree to which *A* and *B* are similar to each other, weights are expected to be equal and similarity judgements are symmetric. In Gleitman et al. (1996) the influence of directional vs. nondirectional presentation is experimentally examined for a number of predicates that are intuitively thought to be symmetrical including *similar*, *equal* and *identical*. The authors find that the way of presentation is decisive for the (a)symmetry in the interpretation of these predicates, even if the nouns they are combined with nonsense nouns.

Tversky as well as Gleitman et al. attribute the asymmetry effects triggered by directional presentation to the difference between Figure and Ground. The same idea is found in our second definition of relative similarity (definition 7c), where the second argument takes the role of the Ground in determining the relevant granule.

---

<sup>29</sup> A metric distance function  $\delta$  has to comply with (i) minimality:  $\delta(a,b) \geq \delta(a,a) = 0$ , (ii) symmetry:  $\delta(a,b) = \delta(b,a)$  and (iii) triangle inequality:  $\delta(a,b) + \delta(b,c) \geq \delta(a,c)$ .

<sup>30</sup> *sim'* (definition 7c) is in fact intransitive due to using the second argument as point of reference.

<sup>31</sup> In Tversky's contrast model a function  $S$  takes weighted sums of the feature sets  $A$  and  $B$  of objects  $a$  and  $b$  to an interval scale such that  $sim(a,b) \leq sim(c,d)$  iff  $S(a,b) \leq S(c,d)$ , where  $S(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$ ,  $\alpha, \beta, \theta$  denote weighting functions and  $f$  denotes a nonnegative scale.

<sup>32</sup> There is also the issue of which features are activated in the first place. In a directional presentation the subject will determine which features are relevant in comparison.

### 'Exactly' reading versus 'at-least' reading

As shown in section 3, scalar equatives may have two readings: an *exactly* reading and an *at-least* reading – *Anna is as tall as Berta* may be interpreted such that Anna's height is the same as Berta's height or such that Anna's height exceeds Berta's height. We assume that the semantics of scalar equatives is uniformly given by similarity even in contexts requiring an *at-least* reading, and we implement this idea by exploiting the granularity provided by closures on classifier systems.

The *exactly* reading of equatives is accounted for by the granules defined by the available classifiers and the reference point  $\mu(Berta)$ .  $\mu(Anna)$  must be in the granule of  $\mu(Berta)$ . To account for the *at-least* reading we need a transformation of classifiers such that all degrees above a certain point  $x$  count as similar.<sup>33</sup> Formally, we define a mapping from the classifier set  $\tilde{P}^*$  to a subset  $\tilde{P}_r^*$  such that every  $p^*$  in  $\tilde{P}^*$  that classifies a member of  $cl_{\rightarrow}([x]_{\mathcal{F}})$  as true is mapped to its right closure while the others stay unchanged. Fig. 8 shows such a mapping: All classifiers left to  $[r]$  stay unchanged, while all classifiers to the right of  $[r]$  will be mapped to  $[r]$ . If the classifier extensions overlap, the situation may be quite complex. The right closure of  $[r]$  handles the general case. This procedure makes it possible to derive the *at-least* reading from the *exactly* reading by solely adapting classifiers. We call it a *quasi-exactly* implementation of the *at-least* reading:

**Quasi-exactly implementation of the *at-least* reading** by right closure of classifiers:

$$\tilde{P}_r^* = \{p_r^* \mid \text{for } p^* \in \tilde{P}^* \text{ if } p^* \cap cl_{\rightarrow}([r]_{\mathcal{F}}) \neq \emptyset \text{ then } p_r^* = cl_{\rightarrow}(cl(p^* \cup [r]_{\mathcal{F}})) \\ \text{else } p_r^* = p^*\}$$

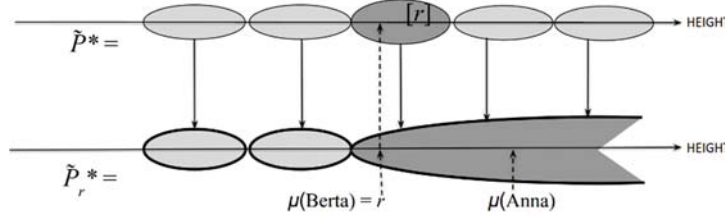


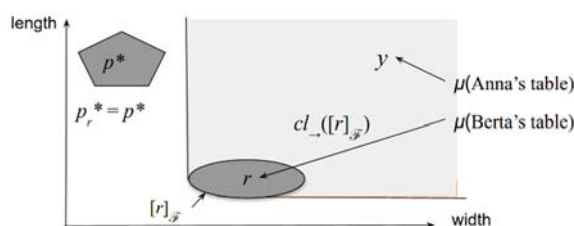
Fig. 8. *Quasi-exactly* implementation, one dimension

Although we get an *at-least* reading, the result still defines an equivalence class:<sup>34</sup> If we select a granule by a point of reference, every element in the granule is

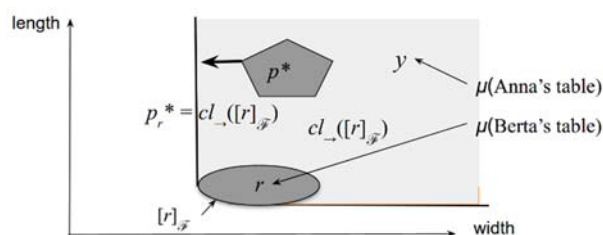
<sup>33</sup> If we have a simple interval scale, we can model the *at-least* reading directly by the order of the attribute values. If we want to model granularity in addition, it becomes more complex since granules may overlap. If the scale is weaker or multiple dimensions are involved, comparison becomes even more complex. Our approach provides a uniform framework for all these cases.

<sup>34</sup> Since we have to select the granule first, it is a kind of 'local' equivalence class.

equivalent to every other element in the granule. This approach can handle multi-dimensional cases, too. Assume that we are talking about the size of tables representing tables by dimensions *length* and *width*, and we use the classical convex closure of the Euclidean two-dimensional space. For non-overlapping classifiers the following two situation may occur (Fig. 9a+b). If the extension of a classifier  $p^*$  is outside  $cl_{\rightarrow}([r])$ , then  $p^*$  stays unchanged. If it is inside, then  $p^*$  will be mapped to  $cl_{\rightarrow}([r])$ , analogous to the one-dimensional case. The general case with overlapping classifiers is again covered by the formula in Fig. 8.



**Fig. 9a.** *Quasi-exactly* interpretation, two dimensions,  $p^* \cap cl_{\rightarrow}([r]_{\mathcal{F}}) = \emptyset$



**Fig. 9b.** *Quasi-exactly* interpretation, two dimensions,  $p^* \cap cl_{\rightarrow}([r]_{\mathcal{F}}) \neq \emptyset$

It is essential in our approach that the *exactly* interpretation is the primary one and is specified by the granularity given by the (contextually determined) classifier system  $\tilde{P}^*$ . The *at-least* interpretation is derived by applying a transformation to the classifier system  $\tilde{P}^*$  depending on the reference element  $r$ .

## 5 Granularity of representations and gradability of similarity

As stated in section 3, granularity of representations provides a notion of *more similar* serving in the interpretation of the comparative form of the adjective *similar*. More importantly, the notion of *more similar* is exploited in the interpretation of multi-dimensional adjectives in general – positive as well as comparative forms. *Anna is healthy* is true if Anna's health is similar to a (contextually determined)

*healthy* prototype. *Anna is more healthy than Berta* is true if Anna's health is more similar to the prototype than Berta's health.

The core of the formalism are sets of *representations* equipped with a preorder structure (transitive, reflexive, but maybe not antisymmetric). This preorder implements a concept of granularity and granularity change. It will be used to construct a predicate *more\_similar* based on a similarity relation defined by indiscernibility. For two representations  $\mathcal{F}$  and  $\mathcal{F}'$  we can ask whether one is more fine-grained than the other, that is, whether there are entities that can be distinguished in one representation but not in the other. Distinguishability is the opposite of indiscernibility and depends on the attribute spaces and the available classifiers. Therefore, these parameters determine the granularity of representations. We will introduce a reflexive and transitive relation on representations (a preorder), which relates granularity levels.

**Definition 8: Granularity of representations**

Given two representations

$$\begin{aligned} \mathcal{F} &= \langle F, \mu, \_*, \mathcal{D} \rangle && \text{with } \mathcal{D} = \langle D, \_+, \_-, P \rangle \\ \mathcal{F}' &= \langle F', \mu', \_*, \mathcal{D}' \rangle && \text{with } \mathcal{D}' = \langle D', \_+, \_-, P' \rangle \end{aligned}$$

we define:

$\mathcal{F}'$  is *at least as coarse* as  $\mathcal{F}$ ,  $\mathcal{F}' \geq \mathcal{F}$  iff there is a function  $f$  such that

(a) the following diagram commutes:

$$\begin{array}{ccc} D' & \xrightarrow{\mu'} & F' \\ \cup \uparrow & & \uparrow f \\ D & \xrightarrow{\mu} & F \end{array}$$

(b)  $\forall x, y \in F : x \sim_{\mathcal{F}} y \rightarrow f(x) \sim_{\mathcal{F}'} f(y)$

This definition states that what is indiscernible in the finer representation cannot be discriminated in the coarser representation. The strict version  $\mathcal{F}'$  is *coarser than*  $\mathcal{F}$ ,  $\mathcal{F}' > \mathcal{F}$ , can be defined by the non-strict one:

$$\mathcal{F}' > \mathcal{F} \text{ iff } \mathcal{F}' \geq \mathcal{F} \text{ and not } \mathcal{F} \geq \mathcal{F}'$$

What we need now is a specification of a relevant set of representations  $\mathcal{H}$ . The *coarser* relation then turns  $\mathcal{H}$  into a preorder. We call such a structure a *hierarchy of representations*. What is missing to get a partial order from a preorder is the anti-symmetry axiom: from  $\mathcal{F} \geq \mathcal{F}'$  and  $\mathcal{F}' \geq \mathcal{F}$  we cannot conclude that  $\mathcal{F} = \mathcal{F}'$ . We may have different possibilities to get the same structure of granules. These hierarchies are related to the concept of context (van Rooij 2011).



**Definition 9: Hierarchy of representations**

A hierarchy  $\mathcal{H}$  is a set of representations such that for any two elements

$$\mathcal{F}_{1/2} = \langle \langle F_{1/2}, cl_{1/2} \rangle, \mu_{1/2}, \_ *_{1/2}, \langle D_{1/2}, \_ +_{1/2}, \_ -_{1/2}, P_{1/2} \rangle \rangle \in \mathcal{H}$$

we postulate the following constraints:<sup>35</sup>

- consistency:  $\forall p \in P_1 \cap P_2 : (p^{+1} \times p^{-1}) \cap (p^{-2} \times p^{+2}) = \emptyset$   
Elements of  $p^+$  and  $p^-$  cannot change roles in different domains.
- discriminative power:  $\forall p \in P_1 \cap P_2 : (p^{+1} \times p^{-1}) \cap (D_2 \times D_2) \neq \emptyset \rightarrow p^{-2} \times p^{+2} \neq \emptyset$   
If a domain contains a discriminating pair of another domain for a shared predicate identifier, it must itself contain a discriminating pair.<sup>36</sup>
- connectedness:  
 $\exists \mathcal{F} = \langle \langle F, cl \rangle, \mu, \_ *, \langle D, \_ +, \_ -, P \rangle \rangle \in \mathcal{H} : D_1 \subseteq D \wedge D_2 \subseteq D \wedge P_1 \subseteq P \wedge P_2 \subseteq P$

and there are continues closure preserving functions  $f_{1/2} : F \rightarrow F_{1/2}$

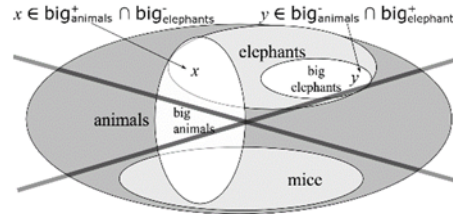
with  $\mu_{1/2} = f_{1/2} \circ \mu$ .

For any two domains there is an enclosing domain.

These constraints can be visualized by the following Venn-diagrams:

- a) the **consistency** constraint rules out cases like this: if  $y$  is a big elephant and  $x$  is a small (not big) one, then  $x$  cannot be a big animal if  $y$  is a small (not big) one.

**Fig. 10.**



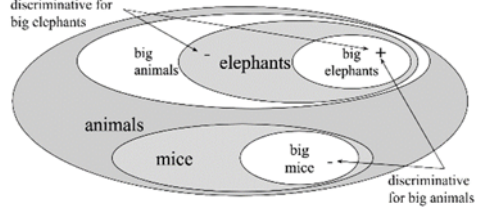
- b) **discriminative power**: If there are big and small elephants there must be big and small animals, too, because animals are different in size: We have big and small elephants which are animals.

<sup>35</sup> (a) and (b) are adaptations of the context constraints in (van Rooij 2011: Definition 1).

<sup>36</sup> If we have big and small elephants and view them as animals, then there should be big and small animals, too. Either there are small animals like mice or, if all animals have the size of elephants, then small elephants must be small animals, too. See the Venn-diagram in Fig 11 and 12.

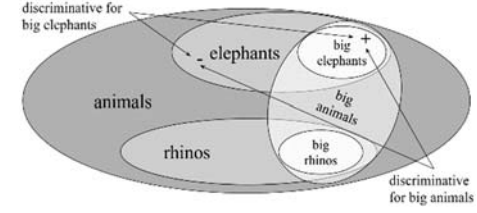
b1. If we collect elephants and mice in one animal domain, then a mouse (big or not) is a negative example for big animals. Thus we have a discriminating pair for big animals.

**Fig. 11.**



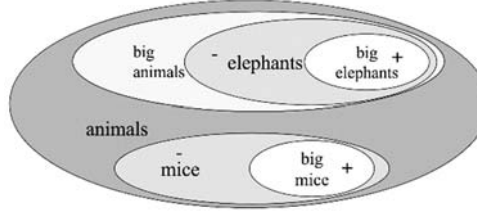
b2. If we collect only big animals in one animal domain, say elephants, hippos, and rhinoceroses, then any discriminating pair for these species is also discriminative for big animals.

**Fig. 12.**



c) **connectedness:** For any two domains there must be a super domain containing both (upward directed).

**Fig. 13.**



In the remainder of this section we assume that there is a contextually given hierarchy of representations  $\mathcal{H}$ . Our approach is non-constructive in the following aspect: We do not construct representations and hierarchies, but instead have systems of constraints which hierarchies must obey. The instantiations must be given by, e.g., the situation of the utterance.

We will now demonstrate how to define a general relation  $more\_sim(a, b, c, d, \mathcal{F})$  based on our similarity relation  $sim$  and the preorder on representations.  $more\_sim(a, b, c, d, \mathcal{F})$  is intended to be true if  $a$  is more similar to  $b$  than  $c$  is to  $d$  with respect to a representation  $\mathcal{F}$ .

**Definition 10: More similar**

Given a hierarchy  $\mathcal{H}$ , a similarity relation<sup>37</sup>  $sim$ , and a representation  $\mathcal{F} \in \mathcal{H}$ , we define

$more\_sim(a, b, c, d, \mathcal{F})$  iff

$$(a) \quad \exists \mathcal{F}' \in \mathcal{H} : \mathcal{F}' \geq \mathcal{F} \wedge sim(a, b, \mathcal{F}') \wedge \neg sim(c, d, \mathcal{F}')$$

$$(b) \quad \forall \mathcal{F}' \in \mathcal{H} : \mathcal{F}' \geq \mathcal{F} \rightarrow (sim(c, d, \mathcal{F}') \rightarrow sim(a, b, \mathcal{F}'))$$

<sup>37</sup> We discussed different similarity relations (see section 4). In this definition, we can use any of these.

The widely used version  $more\_sim(a, b, c, \mathcal{F})$  in the sense that  $a$  is more similar to  $b$  than  $c$  is similar to  $b$  can be defined straightforwardly by:

$$more\_sim(a, b, c, \mathcal{F}) \equiv more\_sim(a, b, c, b, \mathcal{F})$$

If  $a$  is more similar to  $b$  than  $c$  to  $d$  in a given representation  $\mathcal{F}$ , it must be possible to discriminate between  $c$  and  $d$ . Otherwise, because  $c$  and  $d$  are maximal similar,  $a$  and  $b$  cannot be more similar than  $c$  and  $d$ . If we can discriminate between  $c$  and  $d$  in  $\mathcal{F}$ , then we can discriminate between  $c$  and  $d$  in every finer representation but maybe not in every coarser one. If we can find a representation  $\mathcal{F}'$  (maybe coarser than  $\mathcal{F}$ ), such that we can discriminate between  $c$  and  $d$  but not between  $a$  and  $b$  (definition 10a), we are almost done. It remains to exclude contradictions, that is, representations in which we can discriminate between  $a$  and  $b$  but not between  $c$  and  $d$  (this is excluded by definition 10b).

The following diagrams show example hierarchies of representations talking about color and size of objects (each circle stands for a representation):

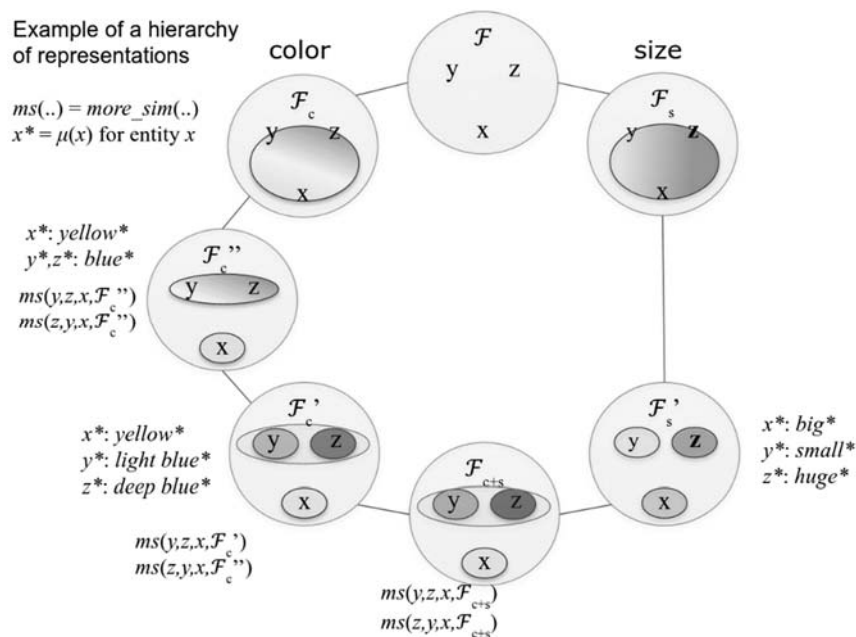


Fig. 14. Hierarchy of representations, example 1

Representations which are higher in the hierarchy are coarser than lower ones. On the left branch we introduce a dimension *color* and a classifier system based on  $\{yellow^*, light-blue^*, blue^*\}$  which can classify colors by convex subsets of a (three-dimensional) color space. On the right branch, we introduce a dimension *size* with a corresponding classifier system  $\{small^*, big^*, huge^*\}$ . The bottom

representation integrates the left branch and the right branch (definition 9c). Again, the *size* dimension need not to be a simple proportional scale. It can itself be a three-dimensional vector space with sub-dimensions *length*, *width*, and *height*.

According to the definition 10a, the *more\_sim* relation will be inherited from top to bottom along the coarser relation. In the circles, we see the extensions of the corresponding  $\tilde{P}^*$  elements. Next to the circles we see the statements about *more\_sim* which are true in these representations. These statements depend not only on the representation they are attached to, but on the whole upper structure (the filter) of the representation. If we look at the circle at the bottom  $\mathcal{F}_{c+s}$ , we see that we inherit two statements, both from the left branch:

$$\text{more\_sim}(y, z, x, \mathcal{F}_{c+s}) \text{ and } \text{more\_sim}(z, y, x, \mathcal{F}_{c+s}).$$

From the right branch, we inherit nothing because the classifier system is too weak. Representations may inherit inconsistent information from different paths which rule out some of the statements (by definition 10b). We can see this when we add more powerful classifiers to the right branch, see Fig. 15:

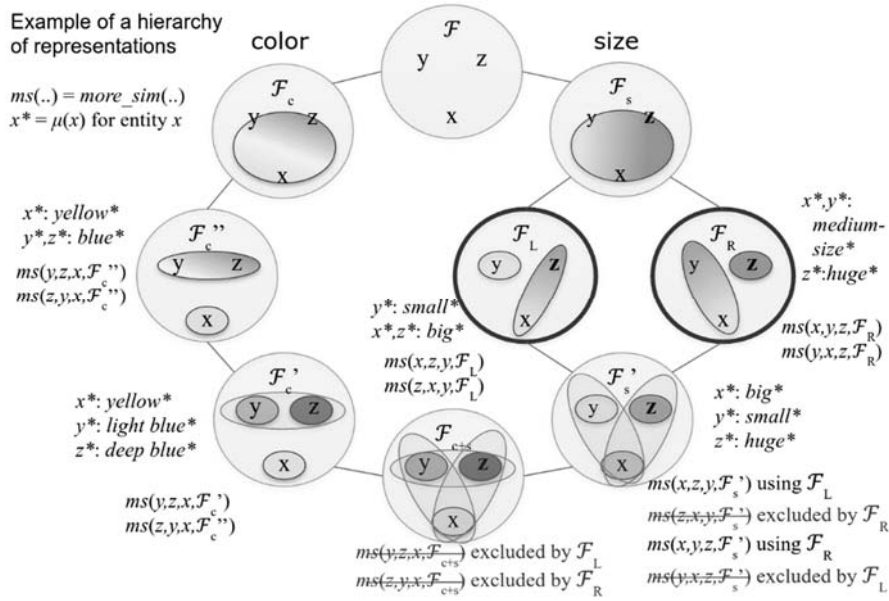


Fig. 15. Hierarchy of representations, example 2

The two heavily bordered circles ( $\mathcal{F}_L$  and  $\mathcal{F}_R$ ) are alternatives which have different effects on the more fine-grained representations (below). The representation  $\mathcal{F}_s$  (circle below  $\mathcal{F}_L$  and  $\mathcal{F}_R$ ) inherits *more\_sim* statements though some are ruled out by the consistency constraint (definition 10b). In the bottom circle  $\mathcal{F}_{c+s}$  all statements are ruled out by the consistency constraints if both  $\mathcal{F}_L$  and  $\mathcal{F}_R$  are present in  $\mathcal{H}$ . In

$\mathcal{F}_{c+s}$ , *more\_sim*( $z, y, x, \mathcal{F}_{c+s}$ ) would be true ( $z$  is more similar to  $y$  than  $x$  is) according to color because of  $\mathcal{F}_c$  and definition 10a. In this case, we cannot discriminate between  $z$  and  $y$ , but we can discriminate between  $x$  and  $y$ . According to the existential quantifier in definition 10a, this is propagated downwards. On the other side, in  $\mathcal{F}_R$  we cannot discriminate between  $x$  and  $y$ . According to the definition 10b and the universal quantification we should not be able to discriminate between  $z$  and  $y$  in this representation, but we are. Therefore, we get a contradiction.

Since in a natural language utterance the hierarchy of representations is not explicitly expressed, we can interpret the meaning of an utterance like *A is more similar to B than C* only as constraint on the relevant hierarchy of representations.

## 6 Conclusion

We presented a framework introducing a non-metric and qualitative concept of similarity suitable for the interpretation of similarity in natural language.

The basic idea is to "measure" properties of individuals with the help of multi-dimensional attribute spaces representing relevant features of comparison (thus generalizing the idea of degree semantics). In our framework, attribute spaces are complemented by classifiers which are predicates on points in attribute spaces approximating domain predicates; this is what we define as a representation. Individuals count as similar with respect to a particular representation if their values are indistinguishable.

In our framework, the granularity of the similarity relation may vary due to different dimensions of comparison and classifier systems. This leads to sets of representations forming hierarchies of different granularity levels, where the order on representations facilitates a Kleinian style notion of *more similar*.

This system provides a powerful and flexible tool to capture the meaning of natural language similarity expressions and account for the role of similarity in ad-hoc kind formation as well as equative comparison. Future work will explore its capacity in, e.g., multi-dimensional comparison of adjectival, nominal and verbal properties. The general idea of our approach is to reconstruct comparison in natural language in a qualitative way, with the help of different levels of granularity imposed by constraints on systems of classifiers.

**Acknowledgments** We presented previous versions of this paper at the *SFB 991 Kolloquium* (Düsseldorf, 2016), the *Semantics Colloquium* of the Institute of Linguistics (Frankfurt/M, 2017), the ZAS workshop *Records, Frames and Attribute Spaces* (Berlin, 2018), and the workshop *Concepts in Action: Representation, Learning, and Application* (Osnabrück 2018). We would like to thank the colleagues in the audience, and in particular Robin Cooper, Peter Gärdenfors, Wiebke Petersen, Stephanie Solt, Henk Zeevat and Ede Zimmermann for their valuable comments. Finally, we would like to express our gratitude to the editors of this volume for their patience and support. The first author acknowledges financial support by the DFG (UM 100/1-3).

## References

- Anderson, C. & M. Morzycki (2015) Degrees as kinds. *NLLT* 33-3.
- Barsalou, L. W. (1983) Ad hoc categories. *Memory & Cognition* 11 (3), 211-22.
- Barsalou, L.W. (1992). Frames, concepts, and conceptual fields. In E. Kittay & A. Lehrer (eds.), *Frames, fields, and contrasts*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carlson, G. (1980) *Reference to kinds in English*. New York and London: Garland.
- Carlson, G. (2010). Generics and Concepts. In F. J. Pelletier (ed.) *Kinds, Things and Stuff*. Oxford, Oxford University Press, 16-36.
- Carpenter, B. (1992) *The Logic of Typed Feature Structures*. Cambridge Tracts in Theoretical Computer Science, Cambridge University Press.
- Gärdenfors, P. (2000) *Conceptual Spaces*. MIT Press.
- Gleitman, L., Gleitman, H., Miller, C., Ostrin, R. (1996) Similar, and similar concepts. *Cognition* 58, 321-376.
- Greenberg, Y. (2003) *Manifestations of Genericity*. New York: Routledge.
- Gust, H. & C. Umbach (2015) Making use of similarity in referential semantics. In H. Christiansen, I. Stojanovic, G. Papadopoulos (eds.) *Proceedings of the 9th Conference on Modeling and Using Context*, Context 2015, LNCS Springer.
- Kaplan, D. (1989) Demonstratives. In J. Almog, J. Perry & H. Wittstein (eds.), *Themes from Kaplan*, Oxford University Press, 481–563.
- Kennedy, C. (1999) *Projecting the adjective*. Garland Press, New York.
- Klein, E. (1980) A semantics for positive and comparative adjectives. *Linguistics & Philosophy* 4.1, 1-45
- König, E. & C. Umbach (2018) Demonstratives of Manner, of Quality and of Degree: A Neglected Subclass. In M. Coniglio, E. Schlachter & T. Veenstra (eds.). *Atypical demonstratives: syntax, semantics and typology*. Berlin, de Gruyter Mouton.
- Korte, B., Lovász, L., Schrader, R. (1991) *Greedoids*. Springer.
- Minsky, M.L. (1975) A framework for representing knowledge. In P.H. Winston (ed.) *The psychology of computer vision*. New York.
- Nunberg, G. (1984) Individuation in Context. *Proceedings of the West Coast Conference on Formal Linguistics*. Stanford University, 203-217.
- Nunberg, G. (2004) Indexical Descriptions and Descriptive Indexicals . In M. Reimer & A. Bezuidenhout (eds) *Descriptions and Beyond*. Oxford University Press.
- Pawlak, Z. (1998) Granularity of knowledge, indiscernibility and rough sets. *Proceedings of the IEEE International Conference on Fuzzy Systems*, 106-110.
- Prasada, Sundeeep & E. M. Dillingham (2006). Principled and statistical connections in common sense conception. *Cognition* 99:73–112.
- Prasada, Sundeeep, Khemlani, S., Leslie, S-J, Glucksberg, S. (2013). Conceptual distinctions amongst generics. *Cognition*, 126, 405-422.
- van Rooij, R. (2011) Implicit versus Explicit Comparatives. In: Égré P., Klinedinst N. (eds) *Vagueness and Language Use* (pp. 51-72). Palgrave Studies in Pragmatics, Language and Cognition. Palgrave Macmillan, London.
- Sassoon, G. (2013) A typology of multidimensional adjectives. *Journal of Semantics*, 30, 335 -380.
- Solt, S. (2016). Multidimensionality, subjectivity and scales: experimental evidence. In E. Castroviejo, L. McNally, and G. Sassoon (eds.) *Gradability, scale structure and vagueness: experimental perspectives*. Springer.
- Solt, S. & B. Waldon (2019) Numerals under negation: Empirical findings. *Glossa*: 4(1): 113. 1–31.
- Tversky, A. (1977) Features of similarity. *Psychological Review* 84: 327-352.

- Umbach, C. & H. Gust (2014) Similarity Demonstratives. *Lingua* 149, 74-93.
- Umbach, C. & H. Gust (in print) Grading similarity. In Löbner, S., Gamerschlag, T., Kalenschner, T., Schrenk, M. & Zeevat, H. (eds.) *Concepts, Frames and Cascades in Semantics, Cognition and Ontology*. Springer.
- Umbach, C. & Stolterfoht, B. (in prep.). Ad-hoc kind formation by similarity.