DOI: 10.17185/duepublico/40717

Prototype-based Knowledge Representation for an Improved Human-robot Interaction ⋆

Alisa Volkert * **Alexandra Kirsch** *

* *Human-Computer Interaction and Artificial Intelligence, Eberhard Karls Universität Tübingen, Tübingen, Germany (e-mail: Alisa.Volkert@uni-tuebingen.de)*

Abstract: We propose a knowledge representation based on prototype theory in order to improve human-robot interaction. Since robots are becoming increasingly important in our everyday life, one day they might be used to do the chores, e.g., in kitchens. In order to tidy objects, however, robots have to be able to find the places where items belong; thus they need to categorise them. We develop a paradigm that mimics human categorisation, in order to provide flexible, human-like solutions. In order to identify a suitable approach to the prototype theory, we augmented and implemented the approach by Hampton and the one described by Minda and Smith, and compared their performance. We found that prototype models represent similarities between objects well. Furthermore, we found that the approach described by Minda and Smith is preferable over Hampton's, although on the whole they do not differ to a great extent. We provide an idea how the approach by Minda and Smith — augmented by us — could be used in order to contribute to a human-like knowledge representation. One final question is to what extent the proposed knowledge representation is able to reflect human categorisations and if the resulting behaviour of a robot is intuitively understandable for users.

1. INTRODUCTION

In order to improve human-robot interaction we propose a knowledge representation based on prototype theory. A successful and well-functioning interaction between users and robots is becoming increasingly important, since robots will be used more frequently. One example is the use in households, especially in an ageing society. It is crucial, however, that the actions taken and the solutions found by a robot are intuitively comprehensible for the user.

One task in a household is to tidy up, for example in the kitchen. Tidying up, however, is not that easy: the place where an object usually belongs, might already be occupied, new items may need to be tidied, or after a relocation all the dishes have to be placed in the kitchen in the new flat.

Such tasks rely on a classification of objects: similar objects are usually stored in the same place. But human classification is not necessarily unique and unambiguous. Most people store the cutlery in one place, but what about the eggbeater? Is that a piece of cutlery or is it rather a cooking aid? The answer to such questions depends on the person asked and the situation, for example whether the drawer for the cutlery offers enough room for an eggbeater. A good classification algorithm for a household robot should be able to model such ambiguities and the existence of different solutions. Rather than mimicking one specific classification of objects, our goal is to implement a flexible classification scheme based on existing theories from psychology.

⋆ This work was supported by the Ministry of Science, Research and Art in Baden-Württemberg and by the Bavarian Acedemy of Sciences and Humanities.

Vernon et al. (2007) distinguish between two basic approaches for artificial intelligence systems: rationalist methods and self-organising systems. We find this distinction helpful to position our method among existing knowledge representations.

Rationalist methods — as the name suggests — try to make an agent act rationally. In a rational view of the world, facts are either true or false (or have a certain probability of being true or false). The basic paradigms of rationalist methods are logic (also in the form of ontologies) and probability theory (e.g. in the form of Markov models). The advantages of rationalist knowledge representations are their expressive power and the strong mathematical foundation. Ontologies, for example, are able to store a wide range of everyday knowledge and they work well with automatic reasoning techniques (Lemaignan, 2012; Tenorth, 2011). But ontology classifications are unambiguous: an object either belongs to a category or not and reasoning gives definite answers. As we have argued above, a human-like and human-understandable classification must allow for grey areas and situation-specific adaptations.

This may be the case, if our household robot wants to put flowers in a vase, but there is none available. So it has to find something that is up to the mark of a vase, except e.g., looking nice. This vagueness can not even be well represented by probability theory. A drinking glass is not a vase with a certain probability, but may serve as one in this special situation.

Self-organising methods, in contrast, try to find classifications based on data. Jäkel et al. (2008) show how exemplar models for classification can be well represented by kernel methods from machine learning. Such methods inherently enable an agent to adapt to a changing environment. They

DOI: 10.17185/duepublico/40717

are also robust against modelling errors (because there is practically no manual modelling involved) or data outliers. A drawback of self-organising methods is their black-box character. The results come from intricate computations that are not explicable to a user. This also prohibits the possibility of learning from human advice.

Our goal is to take the best of both worlds to find a cognitively plausible knowledge representation that can learn from observed data and adapt to changes and user requirements, and at the same time comes with high expressive power and the ability to explain its decisions.

In psychology there are several competing paradigms of representing objects based on similarity: the prototype, the exemplar-based model, and the theory paradigm of concepts (Machery, 2009) In the exemplar-based model items are compared to all stored exemplars that are already classified (Medin and Schaffer, 1978). Thus representing a high number of objects, each constituted by a certain set of features, is infeasible for realistic scenarios.

In the theory paradigm, in contrast, a concept of a category contains knowledge that is able to explain the properties of the category members. A theoretical concept of a bird, for example, stores — amongst others — causal and functional knowledge about birds; e.g., that they have wings to fly (Machery, 2009). It is significantly vaguer described and thus harder to implement.

In the prototype model items are represented by means of prototypes, each constituting something like a *best example* of a category (Rosch, 1975). According to this theory, categories are not defined by clear boundaries but by a prototype, which could be generated by taking the average of all values of the relevant numerical features of already learnt examples. Relevant features could be — amongst others — height, width, and material. These items are represented as vectors $\mathbf{x}$ in a multidimensional space, the feature space (Valentine, 1991), with indices referring to the dimensions of the space.

Our suggested knowledge representation picks up on this theory of category representations, since items are only compared to the prototypes and thus computationally less expensive than exemplar-based models. In order to be able to deal with all kinds of situations occurring in our scenarios, the robot has to be able to broaden categories and calculate similarities between any item and any stored category. In prototype models similarities are representable without difficulties. For instance, if an object is needed that is similar to a vase, a program can easily go through all items and identify the most similar object.

In order to develop a comprehensive knowledge representation we first had to decide which approach to the prototype model to work with. We chose two recent models of prototype theory that are described in enough depth that allows for an almost direct implementation. Thus, we implement, augment and compare the approaches by Hampton (1993) and by Minda and Smith (2011) with respect to their suitability for human-robot interaction.

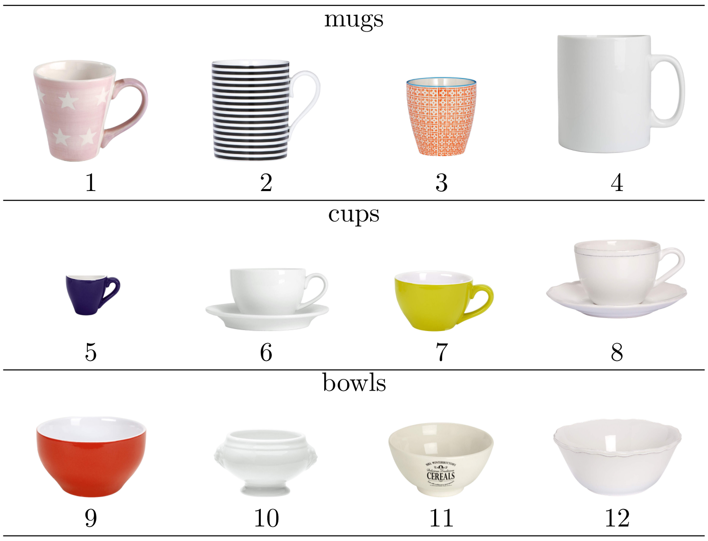


Fig. 1. All tested items. The rows show the *mugs*, *cups*, and *bowls*. See also Appendix A. (All pictures from www.butlers.de)

2. PROTOTYPE MODELS

Both approaches represent an object by a vector, where each component represents one attribute. They state the prototype as a set of averages of all feature values.

We first created a database of twelve items: four mugs, four cups, and four bowls of a European interior dinnerware and decoration department store chain. On this basis, we defined three categories: mugs, cups, and small bowls (Figure 1). We used the attributes *height* in cm, *upper diameter* in cm, *charge* in ml, and *number of handles* (s. Appendix A).

The objects were chosen to have a high resemblance and to allow for ambiguities. For example one of the mugs has no handle and we have cups of different sizes. We think that most people would agree with our categorisation into mugs, cups and bowls, but other classifications could still be acceptable, for example one might use Cup 8 to eat cereals from, or Mug 3 might be accepted as a cup because of its small size. We restricted the data set to these 12 objects to allow for an instance-based validation.

According to our goal of a human-understandable representation, we consider a representation as successful, if it

- correctly classifies all 12 objects into the three assumed categories, but also
- allows for different classifications, possibly by reconfiguring parameters;
- allows for automatic adaptation (learning).

2.1 Approach 1: Hampton

Hampton calculates the similarity S of an item $\mathbf{x}$ to a category C by summing up the attribute-value weights w (Machery, 2009):

$$S(\mathbf{x}, C) = \sum_{i=1}^{n} w(\mathbf{x}, i) \tag{1}$$

DOI: 10.17185/duepublico/40717

where i is the attribute and n the total number of attributes of an item. He does not provide a formula for calculating the weights though. Furthermore, he does not account for different categories at all.

Hampton, however, provides a decision rule for the categorisation (Machery, 2009) that states that items with a similarity S to a category C larger than a threshold t are members of this category:

$$S(\mathbf{x}, C) > t \rightarrow \mathbf{x} \in C. \tag{2}$$

He does not account for the case that an item might have a similarity larger than the threshold to two different categories (Machery, 2009).

Further, this approach does not account for different units of values, considerably diverging values, or discrete values.

2.2 Our augmented version of Hampton's approach

We defined the similarity S of an item $\mathbf{x}$ to a category C accordingly to Hampton's approach:

$$S(\mathbf{x}, C) = \sum_{i=1}^{N_a} \mathbf{sim}_C(\mathbf{x}_i) \tag{3}$$

where i is the attribute and N_a the number of attributes that an item contains. Since in our opinion Hampton's notation of *weights* does not represent any kind of weights at all, we called them attribute similarity *sim*. There is one attribute similarity $\mathbf{sim}_\mathrm{C}$ for each attribute-value of an item. Since Hampton, however, does not account for different classes at all, but we *had* different categories, we had to specify the attribute similarities with respect to them. That is why our attribute similarities depend on a certain category. So we defined $\mathbf{sim}_\mathrm{C}(\mathbf{x}_i)$ [1] as being the attribute similarity of the value — possessed by the instance $\mathbf{x}$ for the ith attribute — to the respective category. As mentioned above we had the following categories:

$$Categories\ C = \{mugs, cups, bowls\}. \tag{4}$$

We calculated the attribute similarities of all items to all predefined categories. Since we had

$$items\ \mathbf{x} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \tag{5}$$

(see Appendix A), we had one 12 items x n attribute similarity matrix for each category and, therefore, three attribute similarity matrices on the whole. Since Hampton does not give any formula for calculating the attribute similarities, we defined the attribute similarities $\mathbf{sim}_\mathrm{C}(\mathbf{x}_i)$ with the help of the Gaussian function. For this we had to calculate the mean over all attribute-values of each attribute i that appeared in a certain category C. These means P_{C_i} constituted the prototype $\mathbf{P}_\mathrm{C}$ for each category C. Formally a prototype $\mathbf{P}_\mathrm{C}$ is a vector of the means P_{C_i}. So for each prototype, there are as many averages as attributes. As attributes we had

$$attribute\ \mathbf{i} = \{height, diameter, charge, handles\}. \tag{6}$$

Our definition of the attribute similarities using the Gaussian function was as follows:

$$\mathbf{sim}_C(\mathbf{x}_i) = exp(-\frac{(\mathbf{x}_i - P_{C_i})^2}{2s_{C_i}^2}), \tag{7}$$

[1] This notation is equivalent to Hampton's $w(x, i)$.

with s_{C_i} being the standard deviation of this sample. In this way the values got normalised, so different units and considerably varying sizes would not affect results. We calculated s_{C_i} for each category C and attribute i. So we had as many standard deviations as averages P_{C_i}. Each category was constituted by four items (see Appendix A):

$$Mugs = \{item\ 1,\ item\ 2,\ item\ 3,\ item\ 4\}, \tag{8}$$

$$Cups = \{item\ 5,\ item\ 6,\ item\ 7,\ item\ 8\}, \tag{9}$$

$$Bowls = \{item\ 9,\ item\ 10,\ item\ 11,\ item\ 12\}, \tag{10}$$

and N_C being the number of instances of the category C:

$$N_{mugs} = |Mugs| \tag{11}$$

$$N_{cups} = |Cups| \tag{12}$$

$$N_{bowls} = |Bowls|. \tag{13}$$

Thus, for the category *mug*, for instance, we calculated the standard deviations as follows:

$$s_{C_i} = \sqrt{\sum_{x=1}^{N_{mugs}} \frac{(\mathbf{x}_i - P_{C_i})^2}{N_{mugs}}}, \tag{14}$$

with $\mathbf{x} \in$ *mugs*. We chose N_{mugs} as denominator, instead of $N_{mugs} - 1$, since we did not want to estimate the standard deviation of a population of unknown *mugs*, but calculating s for our known population of *mugs*. Thus, we did not have to estimate the mean P_{C_i}, since we just could calculate it.

The advantages of using the Gaussian function are that the attribute similarity values are non-dimensional, the minimum attribute similarity is zero, and the maximum value equals one. The latter is advantageous, since the attribute similarities have to be summed up together with all the other ones of an item in a category, so attribute similarities do not diverge to much from each other. The first reason might be even more important, since attributes have different dimensions. So working with values in centimetres and millilitres is not a problem.

Since Hampton's approach, however, does not account for discrete values at all, but we had discrete attribute values, we had to find a way to handle with this situation. Fortunately, if the standard deviation did not become zero, we could treat the discrete values the same way we treated the continuous ones. We had categories though, where all instances had the same value in their discrete attribute *handle*. Thus, the standard deviation of the distribution of this attribute in these two categories became zero. Since dividing by zero is never a good option, we set s to 0.25. In this way we simulated a distribution of attribute similarities with a standard deviation of 0.25. This we could do, since 95 % of this simulated distribution would lie between the prototype for the attribute *handle* $P_{C_h} + 2s$ and $P_{C_h} - 2s$. For a P_{C_h} of 1, almost all values of the simulated distribution would lie between 0.5 and 1.5. If one rounds values between these two, it would always result in 1, which is the prototype P_{C_h} in turn.

Since Hampton does not provide a formula how attribute similarities of the values are calculated exactly, it remains speculative, if our formula represents a way that Hampton would have totally agreed with.

DOI: 10.17185/duepublico/40717

Concerning the decision rule, we only used it in order to compare the results to the ground truth data.

2.3 Approach 2: Minda and Smith

Similarly to Hampton's approach, in the one described by Minda and Smith (2011) the classification decision is done in two steps: 1) items are compared to the prototypes, and 2) it is decided, which category is the most probable one.

In this first phase of comparison, the so-called psychological distance d_{iP} between the item $\mathbf{i}$ and the prototype $\mathbf{P}$ is calculated. For this the distances between the values x_{ik} of the items and the prototype $\mathbf{P}$ are calculated for each attribute k:

$$d_{iP} = \sum_{k=1}^{N} w_k |x_{ik} - \mathbf{P}_k|, \tag{15}$$

where w_k is the weight of the attribute k of N attributes. $\mathbf{P}_k$ is the prototype of a certain attribute [Smith and Minda, 2000; Minda and Smith, 2011].

According to Minda and Smith (2011) the weights can take values between 0 and 1.0, where 0 means no attention has to be paid to this attribute and 1 meaning exclusive attention has to be paid. The weights w_k of a certain category across all the attributes are constrained to sum 1.

The psychological distance d_{iP} is then converted into a measure of similarity η_{iP} of an item $\mathbf{i}$ to a prototype $\mathbf{P}$ (Smith and Minda (2000) followed Shepard (1987)):

$$\eta_{iP} = exp(-c \cdot d_{iP}) \tag{16}$$

where c is a *sensitivity parameter*. It is freely estimated and responsible for the steepness of the decay of similarity. The higher its value, the steeper the decay and the stronger the category attribution for typical instances. A low value leads to a flat descent and the classification probability is closer to chance. According to Minda and Smith (2011) it can take values between 1 and $+\infty$.

In the decision phase (2), it is decided which category is the most probable one of an item. In order to do this, the probability of a Category A-response R_A is calculated for each stimulus S_i:

$$P(R_A|S_i) = \frac{\eta_{iP_A}}{\eta_{iP_A} + \eta_{iP_B}} \tag{17}$$

where η_{iP_A} is the similarity of an item i to the category A, which as the denominator has to be added to the similarity of this item to a second category B (Smith and Minda, 2000).

As Hampton's approach this one does not account for different units of values, considerably diverging values, or discrete values.

2.4 Our augmented version of the approach described by Minda and Smith

In order to stick to Hampton's notation (1993), we had to adjust the one used by Minda and Smith (2011). A list of variable notations is shown in Table (1).

Tab. 1. Variable notations

Minda and Smith	here
item $\mathbf{i}$	instance $\mathbf{x}$
dimension k	attribute i
value x of an item $\mathbf{i}$ in dimension k: x_{i_k}	$\mathbf{x}_i$

Tab. 2. Weights

category	height	diameter	charge	handle
mug	0.4	0.4	0.1	0.1
cup	0.25	0.15	0.1	0.5
bowl	0.15	0.25	0.1	0.5

Our focus, however, does not lie on category decisions and hard boundaries between them; that is why we implemented only the first step. Since we had to account for different categories, we had to specify variables accordingly. Thus, we defined the normalised psychological distance d_{xP_C} between an item $\mathbf{x}$ and a prototype $\mathbf{P}_C$ of a specific category C as follows:

$$d_{xP_C} = \sum_{i=1}^{N_a} w_{C_i} \frac{|\mathbf{x}_i - P_{C_i}|}{s_{C_i}}, \tag{18}$$

where w_{C_i} is the weight of the attribute i of N_a attributes in the category C. P_{C_i} is the prototype of a certain attribute in a specific category [Smith and Minda, 2000; Minda and Smith, 2011]. We supplemented this formula by a normalisation that is done with the help of the standard deviation s_{C_i}. We used the same formula (14) as in the first augmented approach. So it is specific for each category and attribute. Since the second approach, however, does not account for discrete values neither, we treated the standard deviation — in case it became zero — the same way as in our augmented version of Hampton's approach.

Since we did not measure any classificatory significance of dimensions, we set the weights as much reasonable as possible (Tab. 2). We chose exactly these weights for the following reasons: in the category *mug* we set the weights of the attributes *height* and *diameter* to the same value, since we estimated them to be equally important. In contrast, in the category *cup* we gave more importance to the attribute *height* than to the attribute *diameter*, because for us the height of cups seemed quite significant to distinguish them from mugs. In the category of *bowls* it seemed to us to be the other way around and we changed weight values. In all categories we gave only little weight to the feature *charge*. Since all predefined cups had and all bowls did not have a handle, we put the weight for this attribute quite high and low for the category of mugs. In future work we will learn the weights from data automatically.

We had to adjust the formula provided defining the similarity η_{xP_C} to our approach:

$$\eta_{xP_C} = exp(-\alpha \cdot d_{xP_C}) \tag{19}$$

where α is the sensitivity parameter. We set α to one, because otherwise the decay would have been too steep. (For a more detailed explanation see section 4.)

Figure (2) shows the exponential decay of similarity plotted against the distance of all items to the prototype *cup*.

DOI: 10.17185/duepublico/40717

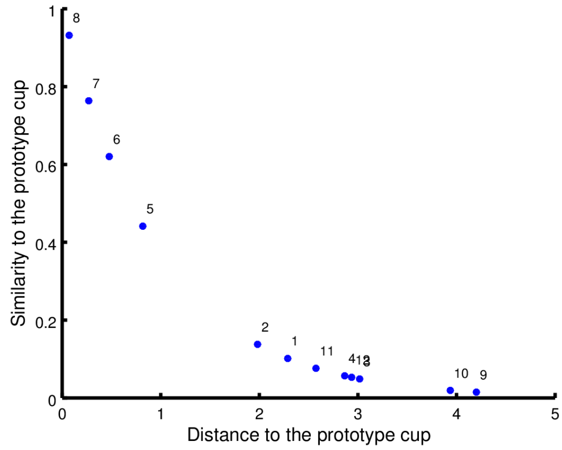


Fig. 2. Exponential decay of similarities of items to the prototype *cup* according to the second approach. Similarity is plotted on the y-axis, while the normalised psychological distance of objects to the prototype is plotted on the x-axis. The points correspond to the items shown in Fig.1.

3. VALIDATION

We evaluated the theoretical approaches with our collection of items.

3.1 Approach 1: Hampton

Figure 3 shows the similarities of all instances compared to the three prototypes. The horizontal line shows the thresholds with which the instances would fall into the predefined categories. It is only shown in order to represent the ground truth data. Since Hampton does not give any further information about the thresholds, we specified them accordingly to the predefined category (Tab. 3).

In the comparison to the prototype *mug* all predefined *mugs* would be categorised as *mugs* with a threshold of 1.6 (Fig. 3(a)). Only item no.8 would then fall into this category as well. Item no.3 is a *mug* without handles. All other items are much less similar.

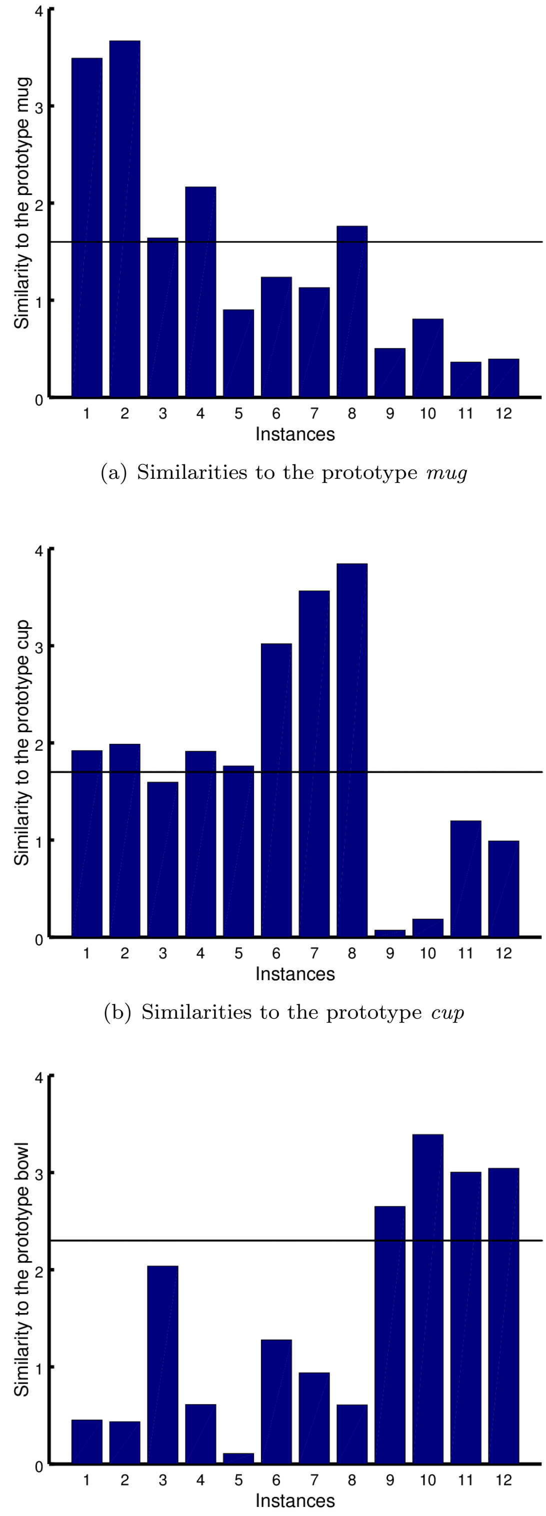
When items were compared to the prototype *cup* three of four predefined *cups* were rated much more similar to the prototype than all other items (Fig. 3(b)). If one wants to set a threshold so that all predefined *cups* are categorised as so, item no.1, 2, and 4 would fall into the class *cups* as well. Also item no.3 is relatively similar to the prototype as well.

In the case of the last category all predefined *bowls* are much more similar to the prototype than all other items (Fig. 3(c)). Thus all and only predefined *bowls* would fall into the category, if a threshold is set. Only item no.3 is quite similar to the prototype as well.

Discussion Three of four predefined *mugs* are more similar to their prototype than other items. Only item no.3

Tab. 3. Category thresholds

mugs	cups	bowls
1.6	1.7	2.3



(a) Similarities to the prototype *mug*

(b) Similarities to the prototype *cup*

(c) Similarities to the prototype *bowl*

Fig. 3. Similarities to each of the three prototypes according to Hampton's approach. All instances were compared to the prototypes *mug*, *cup*, and *bowl*. The horizontal line shows the threshold with which an item would fall into the predefined category. It is only depicted in order to compare the results with the ground truth data. The x-axis shows all instances, while the y-axis depicts the similarity to the respective prototype.

DOI: 10.17185/duepublico/40717

did not reach that much similarity compared to all other instances of the category *mug*. That might be, because this instance does not have a handle, while all other instances of this category do have one. Cup no.8 is quite similar to the prototype *mug* as well. The reason may be that its upper diameter is quite small compared to the other diameters of the normal cups. So no.8 resembles the prototype *mug* most compared to all other cups.

In the comparison of the items to the prototype *cup* three *mugs* are, however, more similar than item no.5 — the espresso cup. This is probably the case, because the latter one is much smaller than all other *cups* and *mugs*, and thus not a usual cup. Results of this comparison are, however, clearer than those of the comparison to the prototype *mug*. At least three of four predefined *cups* are much more similar than all other items, whereat *mugs* are still more similar to the prototype than *bowls*. This is what we had expected, since *mugs* and *cups* resemble each other in appearance more than *bowls* and the rest.

On the whole, similarities of all predefined *mugs* to the prototype *cup* (Fig. 3(b)) are much higher than the similarities of most of the predefined *cups* to the prototype *mug* (Fig. 3(a)). This phenomenon is quite intuitive, since in the German culture area people would classify *mugs* as *cups* anyway, but they would not categorise all *cups* as being *mugs* as well.

In the case of the comparison to the prototype *bowl*, the situation is even clearer: all predefined *bowls* are very similar to their prototype. Additionally, item no.3 — a *mug* — obtains a high similarity score, since it is the only non-*bowl* item that does not have a handle.

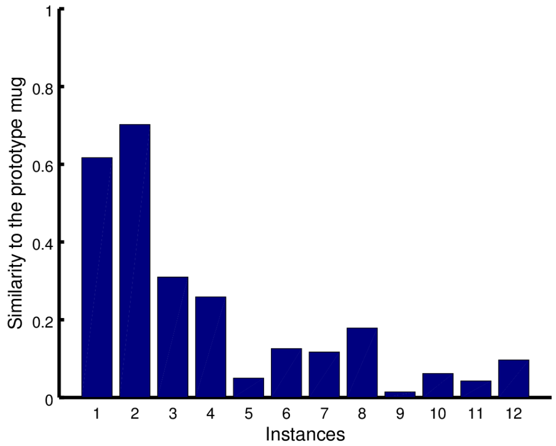
According to our success criteria from Section 2 on page 2, Hampton's model fails to classify all objects according to our assumptions. But the misclassifications can be explained and might be acceptable to a user. It thus models some realistic confusions of objects, but there is no parameter to explicitly adjust the level of flexibility. Options for learning are discussed in Section 4.

3.2 Approach 2: Minda and Smith

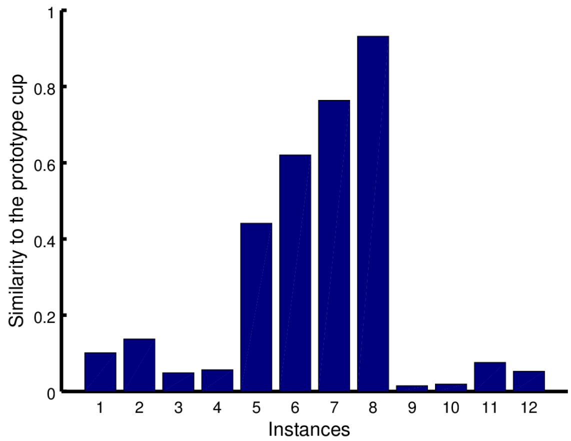
For the approach described by Minda and Smith Figure 4 shows the similarities.

When items were compared to the prototype *mug* the first two items obtain a similarity to the prototype *mug* larger than 0.6, while the other two predefined *mugs* are still above 0.2 (Fig. 4(a)). Three *cups* follow, while all other items only got similarity scores below 0.1. In the comparison to the prototype *cup* all predefined *cups* are much more similar to the prototype than all other items (Fig. 4(b)). In the case of the category *bowl* again the predefined *bowls* are much more similar to the prototype *bowl* than most of all other items; except item no.3 (Fig. 4(c)).

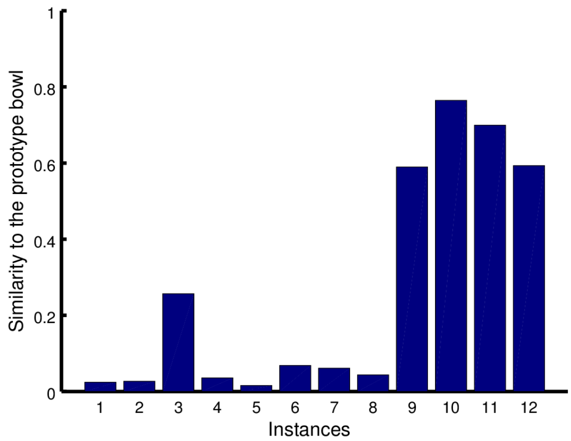
Discussion In the case of the comparison of the items to the prototype *mug* item no.1 and 2 get higher similarity scores, since they are quite prototypical, while item no.3 does not have a handle and item no.4 is relatively big. In the comparison to the prototype *cup* similarities diverge more between predefined categories: predefined *cups* are much more similar to their prototype than all other items. When items are compared to the prototype *bowl* results look similarly clear. Only item no.3 obtained still a quite



(a) Similarities to the prototype *mug*



(b) Similarities to the prototype *cup*



(c) Similarities to the prototype *bowl*

Fig. 4. Similarities to the three prototypes according to the second approach. Here we did not define any thresholds, since this approach does not contain any. For further explanation see Fig. 3.

DOI: 10.17185/duepublico/40717

high score like it was the case in Hampton's approach. The same reason as in his approach might apply here.

The steepness of the exponential function (Fig. 2) depends on the sensitivity parameter α. While Minda and Smith (2011) stated it could only take values between 1 and $+\infty$ we could not find any mathematical reason, why α must not take values smaller than one. We even find it advantageous if it can take values between zero and one as well, since it provides a greater margin with respect to the steepness of the exponential function. While in some cases a greater category endorsement might be favourable, in other situations a more gradual decay is advantageous. Here we set α to one, which still led to a steep decline of the exponential function. If a more gradual decay is favoured, the sensitivity parameter must be set lower. On the other hand, setting the sensitivity parameter α of the second approach to 0.5 does not change the fact that its results represent ground truth data better than those of Hampton's approach. So we still obtain good results, even if we prefer a more gradual decay in the second approach. We are going to give an example in the outlook (Sec. 5) where a more gradual decline is advantageous.

With respect to the success criteria from Section 2 on page 2, the model of Minda and Smith perfectly models our intended classification. In addition, the parameter α can adjust the level of vagueness and ambiguity (Nosofsky, 1987). Options for learning are discussed in the next section.

4. SUMMARY AND GENERAL DISCUSSION

In this section we will compare the two approaches on the whole. We will first focus on similarities of both approaches and then discuss their differences and drawbacks.

With our extension of the formulae by Hampton both approaches now use — in albeit slightly different ways — the distance between an item and the prototype: $|\mathbf{x}_i - P_{C_i}|$. Additionally they calculate the negative exponential function of it (equation (16) for the second approach and (7) for the first one). Thus, larger distances between a prototype and an item result in smaller similarity scores. So in the end both approaches need the distance between an item and the prototype and calculate the negative exponential function of it. Both add up these distances over all attributes of a given instance at some point in time providing the similarity between the item and the prototype compared. Consequently, both approaches perform analogue calculations.

Both approaches, however, do not say anything about discrete values, e.g., the number of handles; and the given formulae do not work with discrete values properly. That is why we had to set the standard deviation to a certain value, if it otherwise would have been zero. So the approaches can now account for discrete values as well. And both approaches ignore the problem of how to handle different units or attributes that vary substantially, since different kinds of objects sometimes have the same features, but very different sizes. That is why we normalised all values with the help of the standard deviation s, so different units and considerably varying sizes do not affect results.

One considerable advantage of both approaches is that items can easily be classified to more than one category, albeit in the approach by Minda and Smith this can be done better if the sensitivity parameter is set smaller than one. This is not disadvantageous, since one could classify mugs as cups as well. Depending on the purpose items can be used in very different ways: think of a French *bol*, a cup for white coffee that looks like a cereal bowl. You can use it to drink white coffee and you can easily use it for eating cereals.

Finally we will focus on the differences between the two approaches and the drawbacks of Hampton's approach.

Although we managed to bring about results that are quite similar to those of the second approach, Hampton's approach leaves some details open. Since he does not provide any formula on how attribute similarities of the attribute-values are calculated, we can only speculate whether our formula represents a way that Hampton would have agreed with. Yet, the resemblance of the results lets us assume that our choice is adequate.

While Hampton does not account for different categories at all, the approach described by Minda and Smith at least gives a formula in the decision phase for the calculation of the probability of a certain category response compared to others. However, it does not provide more detailed information in the formulae about considering different categories neither. They only say that the process of item-to-prototype comparison has to be repeated for all prototypes and thus for all categories.

The first approach (Hampton, 1993), however, can be extended by a formula analogous to the one used in the second approach: Guaus i Termens (2009) proposed to use Luce's Choice Rule (Luce, 1959) in order to decide, which is the most probable category of an item. This is not a big issue, however, since we focus on calculating similarities and not on deciding which category is the most probable one. Yet, someone who is interested in category decision should consider these aspects.

Although Hampton mentions weights for the different attributes, he does not give any formula where these weights are implemented. That is why we did not implement them in our version of his approach. It remains to be seen how these weights could influence results and whether they would become similar to the ones evoked by the second approach. This is at the moment, however, a huge drawback of Hampton's approach, since without weights for different attributes, these cannot be adjusted to specific situations.

In both approaches discrete attributes are especially challenging: in the categories *cup* and *bowl* all instances have the same value in the attribute *handle*. This implied that the standard deviation s, which calculated value is zero, is set to 0.25 and led to quite clear-cut results at least in the second approach. If decreasing the sensitivity parameter α is for some reason out of the question, these clear-cut results might become a problem. For instance, in our category *bowl* all instances do not have a handle. If now an item looking altogether like a *bowl*, but having one or two handles (since it is a soup bowl), the distance to the prototype of the category *bowl* would be quite

DOI: 10.17185/duepublico/40717

large. It could be used as a *bowl* though, but a function programmed in order to find the next similar instance (outside the category *bowl*) would never consider this item as suitable for the same purposes as all instances of the category *bowl*. A possible solution may be a more intricate scoring method that not only averages over all attributes, but can ignore attributes or weigh some attributes higher than others, depending on the situation.

In sum, according to the first two success criteria of Section 2, the method of Minda and Smith is preferable to that of Hampton. So far we have ignored the aspect of learning and adaptation. Basically, both approaches support adaptation by the definition of prototypes. The classification of previously unseen objects by the distance to any existing category is a basic constituent of prototype theory. But an open question is whether a new object should be considered as a member of an existing category (i.e. is close enough to an existing prototype) or whether it constitutes a new prototype.

5. OUTLOOK

Our original motivation was to build a knowledge representation that provides classification results for realistic household situations similar to those that humans would find or at least accept. In this paper we have shown how cognitive models can classify objects realistically, but with some vagueness that can even be parametrised in the model by Minda and Smith. But a classification result as such is not enough for a robot to make decisions about where to put objects in a kitchen. In this last section we sketch some aspects of further work that we hope will lead to a flexible decision-making process based on prototype-based representations.

For the classification as such, we need to find out which object attributes are relevant in specific situations. For this purpose we are currently building a larger database of everyday objects and are preparing a similarity scaling study. We also hope to get insights as to when humans generate a new prototype and how prototypes are grouped themselves into categories on a more abstract level.

The next step will be a flexible decision-making algorithm that uses the distance measures of a prototype model to make decisions in household chores, for example where to put a newly acquired mug if the cupboard of mugs is already full. We will use and extend a heuristic problem solving method (Kirsch, 2011) that takes into account different aspects of the situation such as space in cupboards, the objects that are already in a cupboard and their distances, and reachability of places by specific users.

It may well be possible that humans use different classification schemes when asked about an object class and when using or storing the object. This is why we plan to compare our prototype model and the decision-making algorithm to human decisions, both for abstract classification tasks where an object category has to be named and real-world tasks such as storing objects.

ACKNOWLEDGEMENTS

Special thanks go to Juan Purcalla Arrufí for his profound mathematical advice.

REFERENCES

Enric Guaus i Termens. *Audio content processing for automatic music genre classification: descriptors, databases, and classifiers.* PhD thesis, 2009.

James Hampton. Prototype models of concept representation. In Iven Van Mechelen, James Hampton, Ryszard S. Michalski, and Peter Theuns, editors, *Categories and Concepts - Theoretical Views and Inductive Data Analysis*, pages 67 – 95. Academic Press, London, 1st edition, 1993.

Frank Jäkel, Bernhard Schölkopf, and Felix A. Wichmann. Generalization and similarity in exemplar models of categorization: Insights from machine learning. *Psychonomic Bulletin & Review*, 15(2):256–271, 2008.

Alexandra Kirsch. Humanlike problem solving in the context of the traveling salesperson problem. In *AAAI Fall Symposium on Advances in Cognitive Systems*, 2011.

Séverin Lemaignan. *Grounding the Interaction: Knowledge Management for Interactive Robots.* PhD thesis, CNRS - Laboratoire d'Analyse et d'Architecture des Systèmes, Technische Universität at München - Intelligent Autonomous Systems lab, 2012.

Robert Duncan Luce. *Individual choice behavior: a theoretical analysis.* Wiley, New York, 1st edition, 1959.

Edouard Machery. *Doing without concepts.* Oxford University Press, 2009.

Douglas L. Medin and Marguerite M. Schaffer. Context theory of classification learning. *Psychological Review*, 85(3):207–238, 1978.

John Paul Minda and J. David Smith. Prototype models of categorization: basic formulation, predictions, and limitations. In Emmanuel M. Pothos, editor, *Formal approaches in categorization*, pages 40–64. Cambridge Univ. Press, Cambridge, 1st edition, 2011.

Robert M. Nosofsky. Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1987.

Eleanor Rosch. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104(3):192–233, 1975.

Roger Shepard. Toward a universal law of generalization for psychological science. *Science*, 237(4820):1317–1323, September 1987. ISSN 0036-8075. doi: 10.1126/science.3629243.

David J. Smith and John Paul Minda. Thirty categorization results in search of a model. *Experimental Psychology: Learning, Memory, and Cognition*, 26(1): 3–27, 2000.

Moritz Tenorth. *Knowledge Processing for Autonomous Robots.* PhD thesis, Technische Universität München, 2011.

Tim Valentine. A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 43(2):161–204, May 1991.

D. Vernon, G. Metta, and G. Sandini. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *IEEE Transactions on Evolutionary Computation*, 11(2):151–180, 2007.

DOI: 10.17185/duepublico/40717

Appendix A. THE RAW DATA

item	height [cm]	diameter [cm]	charge [ml]	handles
		mugs		
1	10	8	260	1
2	9.5	8	250	1
3	8	8	200	0
4	11	10.5	270	1
		cups		
5	5.8	6	95	1
6	7	11	200	1
7	6.7	10.8	190	1
8	6.5	9.5	195	1
		bowls		
9	7.9	14	630	0
10	8.5	13	400	0
11	6.5	13	300	0
12	7.2	11	350	0