Prototypes and Exemplars

An overview by Eric Auer

13. Januar 2002

In this paper I give a summary of *Rosch, Eleanor: Principles of Categorization (1999)* and *Smith, Edward and Douglas Medin: The Exemplar View (1999)*, both found in *Concepts. Core Readings (M.I.T. Press)* and their approaches on conceptual semantics, the prototype theory and the exemplar view. I have added some examples and comments, but most of the information presented is directly taken from the two papers mentioned above. Comments on my comments can be sent to `<eric@coli.uni-sb.de>` ...

1 Rosch: The Prototype Theory

While the *Classical View* fixes concepts by means of lists of necessary and sufficient conditions in terms of (often binary) features, Rosch sketches a more psychologically oriented approach. For her, the two principles of *Cognitive Economy* and *Perceived World Structure* are the driving forces behind the conceptual system of humans and other animals.

1.1 The two principles of Cognitive Economy and Perceived World Structure

Cognitive Economy is described as the general drive to know or predict as much about the world as needed with as little effort as possible: This means for example predicting many properties from knowing few properties or a category. It also means not distinguishing every possible detail, to keep the set of categories reasonably small. For example an office clerk will not have many different categories for kinds of pigs or even the skill to distinguish individuals in a bunch of pigs. It would be possible, but quite useless for a clerk, so simplifying the conceptual grasp of pigs helps to "keep the world model simple" without sacrifying important information. An example of important information could be "Animals that fly have wings", which saves us from the need to remember the *has-wings* property explicitly for every kind of flying animal that we know.

As we have seen, the personal needs to deal with the world are determining the conceptual system. Most of those needs are similar for a whole culture or species, so the conceptual system is assumed to be influenced by interaction with the outside world including society — it is no innate black box that would be the same at least for all humans, only with some parametrical adjustments. Instead, the two principles are the common (and probably innate) seed, while the conceptual system itself is a result of those principles interacting with a real world.

The Perceived World Structure is the other main principle: It describes the perceived correlation between properties of objects in the real world as we perceive it. So the Perceived World Structure is shaped by the way our senses perceive the outside world and based on the assumption that there is structure in it. The world structure itself is not the basis for our concept forming, but what we perceive of it. So the conceptual system of a dog (that has an efficient nose) will be different from the conceptual system of a human being (that has versatile and quite sensitive hands). The way that our body usually interacts with the world is of course closely related to the Perceived World Structure, so our notion of an object we can sit on depends on the posture and motion pattern we show while sitting and sitting down.

1.2 The structure of the conceptual system

Rosch assumes both horizontal and vertical structure in a conceptual system: Vertical structure would determine the "abstractness coordinate" in a mental space of concepts, the level of inclusion: Concepts like THING or ANIMAL have a high level of inclusion, while concepts like TWEETY or TURTLE have lower levels of inclusion. Horizontal structure is the reason why, for example, CAT and DOG are in two different categories for most of us.

An important measure controlling the structuring of the conceptual system is the *Cue Validity*, which can be seen as the importance of a property. A high Cue Validity of a property for a category means that the members of the category are quite likely to have that property, while members of other categories are clearly less likely to have the property. Thus, a property that is, in colloquial terms, *typical* for members of a category and thus an important aspect of that category. Even though having eyes is more common among birds than having wings, the Cue Validity of wingedness will be higher, because other animals are more likely to have eyes as well than to have wings as well — thus we gain more information about birds if we store the fact that they are characterized by being likely to have wings.

Tversky has developed a similarity measure in the 70s that closely resembles Cue Validity. His research can be seen as an early variant of Prototype Theory in some way. The measure has a simple mathematical idea: Similarity is the weighted sum or another function of the overlap minus the differences of properties (for example measured as the sizes of the property sets AB, $A - B$ and $B - A$).

1.3 The Prototype Theory

While Rosch explicitly disavows that her ideas point to a certain theory of cognition, representation or learning of concepts, it is conceivable to take the two principles of Cognitive Economy and Perceived World Structure as a starting point to a theory of categorization, the *Prototype Theory*.

The Perceived World Structure is related to the assumption that there are some combinations of properties that are far more frequent than others. Combined with the concept of Cue Validity (which is related to Tverskys notion of *Category/Family Resemblance*) we can come up with more predictions about structure and creation of our conceptual system. For example Tversky states that if common features outweigh distinctive features, classes (categories/concepts...) tend to combine. Tverskys notion puts somewhat less weight in contrasting categories than Roschs Cue Validity, but both lead to predictions about the link between natural discontinuities (also called basic cuts) distinguishing objects and a *basic level of categorization*.

Finding out more about that *Basic Level* is the main topic of the experimental evidence presented by Rosch in her paper: She assumes the Basic Level to be characterized by the high Cue Validity of the distinctions along the horizontal structure of the Basic Level. The next more abstract level will involve low Cue Validity because more abstract concepts tend to have only few common features along most members. For example members of (the extension of) CAR will resemble each other much more than members of (the extension of) the more abstract category VEHICLE. On the other hand there will not be a big increase in Cue Validity and the predictive power of knowing a category if we go to a less inclusive level than the assumed basic level: For example *chairs* share already lots of properties and we do not gain much additional information if we know that a chair is a *kitchen chair* and not an *office chair*, while we did gain much additional information about a piece of *furniture* by hearing that it is a *chair* and not a *shelf*.

1.3.1 Fixing the Basic Level

As mentioned above, the experiments presented by Rosch aim to find out more about the existence, position and importance of the assumed Basic Level of concepts. Her research domain are concrete objects, as their classification is quite consistent and suited for empirical analysis. For several natural and man-made taxonomies (like *fruit* and *bird* or *clothing* and *furniture*), three levels of abstraction are checked, trying to decide which is the Basic Level. The experiments are looking at four different aspects of perception: *Attributes, movements, shape* and *identifiability*.

For the attribute aspects, the subjects were asked to give a list of attributes for each of the categories. The categories represented different levels of abstraction for different taxonomies, for example TREE, PLANT, OAK, HAT, OBJECT and so on. The hypothesis was that the middle level of the chosen taxonomies would be the Basic Level. Checking the number of common attributes lead to the conclusion, that this was correct for most taxonomies, but that TREE and not less abstract levels was basic for a taxonomy ranging from PLANT to different kinds of *oaks*. At any rate there was clearly one level recognizeable Basic Level in all cases.

The next experiment was about movements involved in interaction between the human body and objects: The subjects had to describe the movements typically related to certain objects in great detail, and again the overlap was analyzed. As predicted, there were not many movements typical to interaction with furniture, but many typical to interaction with chairs. This lead for example to assuming CHAIR to be on the Basic Level, which is supported by the fact that there are not much other typical movements that would be different depending on *what kind of chair* we interact with.

The last two experiments had to do with shape and identifiability: First, the overlap of two-dimensional shape was measured, and second, the identifiability of two-dimensional shapes created from a set of members of a category was checked. As expected, chairs looked much the same while kitchen chairs did not resemble each other much more, and an image of an "averaged chair" was easier to be identified than one of "average furniture" and so on — the results led to the same categories to be considered basic than the results of the other experiments.

All experiments suggest an important role of the Basic Level: There could be mental images of a *prototypical chair* and other objects at the Basic Level — the more abstract categories cannot be represented that way, while such a representation for less abstract categories would probably violate Cognitive Economy. For perception, the idea is that objects would be first identified as members of a certain basic category. Depending on the needs in the current context, we would then decide what more exact (less abstract) or more abstract category they belong to.

This is also supported by the finding that we tend to use basic level names for objects as a default, shifting to more or less abstract levels depending on the context. For example ASL (the American Sign Language of the Deaf) uses single signs for basic level terms far more often than for terms of the other levels.

From the two principles over the basic level the next step would be looking at the importance of *prototypes* in the conceptual system: While they are less useful for reasoning (especially over abstract categories, but also for the other levels), they should be quite useful in everyday life: Only few objects are unclear cases, so categorization of objects by comparing them to some representation of a prototype would be an efficient method. The prototype in this setup would not necessarily be a certain instance (member of the extension) of a concept — this is sometimes misunderstood. It is also interesting to compare this to the representation assumed in the *Exemplar View by Smith and Medin* below. So unless the goal is fixing the boundary between two concepts or reasoning about them, thinking with and in prototypes can be seen as an efficient and psychologically plausible alternative to the Classical View with sets of necessary and sufficient features that would be compared to sets of perceived and otherwise known features.

Even looking at this "prototypical" point of view, Rosch still warns about taking her notion of a prototype too literally: Some prototype does not need to exist at any place, not even in some mental representation of a concept — only a measure of prototypicality is assumed to exist in some way, which does not mean any prototype itself needs to exist. But if there is some prototype, then whenever something resembles it more closely, it will have a higher prototypicality measure (and maximize Cue Validity and family resemblance in some way). If attributes with a metric are involved, the prototype could be the mean or median. Prototypes can then be seen as the members that most reflect the redundancy structure of the category. This allows comparing a category to a cluster of points in some attribute space with the centre of gravity being the prototype.

The relation of prototypicality and cognition seems to show a clear effect in various ways: Things like reaction time, speed and accurateness in learning artificial categories, the normal order of learning categories for children (they can verify category membership earlier and better for prototypical members), the order and probability of listing an item when asked to list category members, as well as many other effects seem to be related to prototypicality of objects.

1.4 Closing comments

Rosch considers the mobility of the Basic Level due to *context* effects to be a flaw in her view of the conceptual structure, and wonders about the psychological reality of perceiving natural discontinuities in the world structure: People may fail to take some correlational patterns to be important or they may fail to notice some patterns completely. She argues that sometimes attributes are not actually perceived but are echoing back from some reasoning process about concepts after the initial categorization.

In the light of those open issues, the next topic for research is a suggested relation between *events* and event scripts as temporal units of experience and context on one hand and our conceptual system on the other hand. Like with the Basic Level for concepts, Rosch found some typical (temporal) size of events. This time did not get longer for events further away, it was just the number of remembered events that decreased (not the temporal resolution). Events can be described in event scripts, which in turn involve/require certain objects. The idea is that typical or good script elements would be also good prototypical objects, presumably of the Basic Level. Rosch claims that event scripts are better described using Basic Level objects than using objects that are too abstract or too specific. Example: Writing a letter with a pen as opposed to moving a pointed object on a flat object leaving some marks on the abstract side and drawing straight and curved parts of letters with a green wooden pencil on white letter sized light paper.

2 Smith and Medin: The Exemplar View

Rosch is focussing issues related to some conceptual system based on the two principles of *Cognitive Economy* and *Perceived World Structure* and is quite reluctant to think about more concrete implications for representation, learning and cognition. She puts great stress on researching the horizontal and vertical structure of it, especially involving *Cue Valitity* and the *Basic Level* of concepts. Announcing some *Prototype Theory* based on that is what she seems to avoid as far as possible.

Smith and *Medin* on the other hand take a Prototype Theory to be a quite real thing, so they develop their own alternative theory to it — the *Exemplar View*. In their theory, any concept (or category) would be represented by a small set of *exemplars*. An exemplar can be either a subset of the category (thus another category) or a single instance of the category. To be more exact: The representation of a single member of the category, described by a list of properties. As with the Prototype Theory, potential instances are not excluded, but real instances seem to be preferred by Smith and Medin. So for example the concept BIRD could be represented by our pet robin ROBBIE, the sub-concepts EAGLE, DOMESTIC-BIRDS and PENGUIN, and so on.

2.1 Contrasting the Exemplar View to others

Both abstraction and prototypes are combined in the Exemplar View, but they are different to the abstraction of the Classical View and the prototypes of what could be the Prototype Theory. For the Exemplar View, concepts are an explicit disjunction of parts, so they do not require the often problematic notion of (globally) necessary and sufficient conditions for category membership, but this can also cause some problems specific to the Exemplar View, as we shall see. In the Classical View, the conditions (also referred to as *summary information* below) are necessary part of every concept and are always the result of some abstraction process.

The disjunctive approach also differentiates the instance type exemplars in the exemplar view from prototypes as centre of gravity of the whole category at least for categorization (as opposed to reasoning, which is better done using properties in the way the Classical View works). While prototypicality ratings based on single prototypes resemble statistical clustering of items into categories, the Exemplar View is less dedicated to abstraction processes.

Smith and Medin argue for their combined model that exemplars have reasonable psychological plausibility: People seem to think in terms of exemplars when they have to deal with abstract categories for categorizations and taking decisions. For example if one gets asked whether all birds could fly, browsing through a set of exemplars to find counterexamples seems to explain better what we are happens in our mind than the other two theories: Neither any conceivable operation with a single prototype nor any simple processing of lists of necessary and sufficient conditions of birdiness give a good model of such reasoning. Again, the Exemplar View could be said to be more real life and less abstraction, with several pros and cons arising from that.

An extreme case or implementation of the Exemplar View would be the *Proximity Model*: In that model, a category is represented by all category members encountered in the past, thus maximizing the disjunctive nature of the approach and minimizing the abstraction process involved. No summary information at all would be stored, and a new item gets incorporated in the category that includes the item which is most similar to the newly arriving item. Not only do we lose a great deal of abstraction, this model is also implausible for larger numbers of items piling up during the time of ones life.

2.2 The Best Examples Model and the Context Model

As the central model for describing the Exemplar View, I chose the *Best Examples Model*, with the variant (or successor) of the *Context Model* described below. In this model, which is somehow based on the Rosch model, a small number of *typical exemplars* represent each category. They are those with a high family resemblance and typicality, they share at least a critical number of properties. The rationale for having several best examples is again empirical data: In many real life contexts, several exemplars share the highest family resemblance or score the same on typicality ratings. This holds especially for more abstract categories: For example the concept ANIMAL is best described as a disjunction of several less abstract concepts like FISH, MAMMAL, BIRD and INSECT.

For the decision *which* exemplars are used to describe a given concept, Smith and Medin propose that the summary information is used. They complain about this being not elegant: Why abstract the summary information first and then throw it away as soon as the exemplar representation works better? I suggest not to assume that the summary information gets discarded at all — it is just less accessible and only used for purposes where exemplars are not well suited. After all, the Exemplar View does not explicitly deny the use of (abstract) summary information, and non-global "summary information" describing instance type exemplars is even a normal part of the Exemplar View.

Another way to solve the decision problem is proposed by Smith and Medin: Newly encountered exemplars are by default added to the representation, but if they turn out to be atypical (or unimportant, because they are marginal variants of already stored exemplars), they get discarded from it after a while. This seems plausible, as typical exemplars tend to be taught/encountered first. But then we just move the question, because we now have to explain why they are taught first...

For categorizing objects, a probabilistic paradigm is assumed: All exemplars are retrieved, better exemplars tend to be retrieved first. If a sufficient match is found fast enough, the object is assumed to belong to the category. Medin and *Scheffer* refine this in the *Context Model* discussed below: The probability (or average speed) of retrieval depends on the similarity to the new object, and if a critical number of exemplars is retrieved in a certain time and before a counterexample is retrieved, the object is considered to belong to the category of the retrieved exemplars. For measuring similarity, similar methods as discussed by Rosch can be used, as well as feature-/value- or template-based methods.

The Context Model allows reducing the number of stored exemplars, because it is conceivable to store only a subset of the features otherwise used with the particular exemplars, along with some weights denoting the importance of the exemplar for the various attributes or features. In an extreme case, the model will converge to some probabilistic representation similar to some variants of the models used with the Prototype Theory. An important difference is that the Context Model does not *add up* matching and mismatching features or differences along dimensions. Instead, it *multiplies* the differences along dimensions of attributes or features (optionally with weights involved, which are determined by salience and perceived contrast). The similarity or probability to retrieve some item X for a concept A would be the fraction of the summed up similarity of that item to the exemplars of A divided by the summed up similarity of that item to the exemplars of the union of A and the contrasting category (or categories) B. Some experimental evidence shows that the multiplicative rule has some advantages over the additive rule in terms of psychological plausibility.

2.3 Seven problems of the Classical View revisited

Smith and Medin select seven typical problems of the Classical View and check the performance of their models in handling those problems. Unless otherwise stated, the Best Examples Model and the Context Model behave quite similar here. The first problem is the handling of inherently *disjunctive concepts* (like ANIMAL, probably): Due to the disjunctive nature of the exemplar based models, this problem does not show up with them at all.

The problem of *unclear cases* is also handled in an elegant way: Because there is no strict definition of the boundary between categories needed, the probabilistic categorization described above can deal with unclear cases of categorization in an appropriate manner. As a related issue, the Exemplar View is also not bothered by the *failure to specify defining features* — the disjunctive representation of categories circumvents problems like the ones described by *Wittgenstein* (he claims that there is no single defining feature that can describe all instances of GAME apart from *is-a-game* itself). Problems with using *unnecessary features* are irrelevant for similar reasons.

The Exemplar View can explain *typicality effects* (where the Prototype Theory shows its strengths): Typical category members are more likely to match well with one of the stored exemplars, and typical members (both instance and subset type exemplars) are more likely part of the representation and thus more probable to be found soon in production tasks[1].

The last problem of the Classical View discussed has also some plausible explanation in the Exemplar View: Some concepts have a *nested* behaviour, for example CHICKEN is a kind of a BIRD but probably a better example of an ANIMAL than of a BIRD (where maybe SPARROW and EAGLE will be found as instance type exemplars instead). Again, the disjunctive structure of exemplar representations gives a plausible solution to our problem.

2.4 Problems special to the Exemplar View

While several problems of the Classical View could be solved quite good by the exemplar based models, there are some problems that show up especially with the Exemplar View. The first of those problems is: *How do we store all knowledge in concepts?* For example how do we find out about the correlation between *small* and *sings* for birds? In an approach mainly based on summary information, the answer could be: We could have explicitly stored the fact that smaller birds sing better.

As I have mentioned above, I think storing both global summary information and a (disjunctive) exemplar representation at the same time is a possible solution and still not in contradiction to the general Exemplar View — so we do not need yet another view, another variant of the Exemplar View will do. Smith and Medin propose a more

[1] This also relates to *typicality detection* — please refer to page 214 of *Concepts* for details.

conservative solution: The correlation is retrieved by abstracting it from a number of exemplars retrieved for the category. For example EAGLE, ROBIN, PEACOCK and SPARROW would be retrieved early, and we can deduce from the fact that robins and sparrows sing better than eagles and peacocks that smaller birds are better with singing. The downside of this approach is that the variability along a dimension for all instances is usually bigger than the estimate that we get from looking at only a few exemplars (which are typically way inside of the multidimensional cluster of possible instances).

The next problem is the *constraining of possible properties*: The Exemplar View does not force us to constrain the possible properties for category members, and depending on the variant, does not even allow global properties (like "it is important to know if a certain kind/instance of bird has wings"). Again, a possible solution would be to allow global summary information to persist.

To avoid the ramifications of having an unbounded set of properties of and relations between exemplars (see below for the latter), some kind of constraint should apply. This does not need to go as far as reverting to necessary and sufficient membership conditions for subset exemplars or the whole category, but after all, a constraint has to be there to avoid too much disjunctiveness. A proposed solution is the implicit constraint imposed by the Context Model with its similarity parameter: Objects of really extreme size are unlikely to be instances of BIRD, so we do not need to consider other properties of those objects to know that they are no birds.

It would be nice if the model could account for *context effects*: Usually, people consider weight to be a more interesting property of an object if the context suggests it — like in "The man lifted the piano". In many other contexts, other properties of the piano would be more salient. Probably this seems to ask for a solution in terms of similarity effects and the like, but a concrete model for handling context models is not found in the Smith and Medin paper.

As mentioned earlier, the ways in which exemplars of a concept are related to each other should be constrained. Again, adding global constraints or global summary information could do the trick, like stating that "furniture is likely to be found in a house" (which greatly limits the set of categories that can provide subset or instance exemplars for furniture). In general it is possible to add summary information to examplars (it can even be added in a separate deliberate act, for example triggered by teacher input). This additional point for retaining summary information in addition to the normal exemplar representation weakens the reluctance of Smith and Medin on that issue: They assume summary information to be less important and less salient in most everyday situations, but they no longer try to avoid longer lasting storage of summary information per se.

2.5 Concluding remarks

The Exemplar View is in some ways similar to probabilistic approaches and the model related to the Prototype Theory, but the Exemplar View comes up with some new ideas and solutions (like the special aspects of the Context Model). There are ways to incorporate summary information as well as obvious advantages of the disjunctive nature of the exemplar representation. On the other hand, it seems to be possible to reduce those problems by making use of explicit summary information as a backup system, even in the tricky case of handling relations between the disjuncts like in the singing/small bird example mentioned. ◇[2]

[2]Layout: LaTeX Editor: `joe` This document online: `http://www.coli.uni-sb.de/~eric/...`