PACKER: An Exemplar Model of Category Generation

Nolan Conaway (nconaway@wisc.edu)
Joseph L. Austerweil (austerweil@wisc.edu)
Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Abstract

FILLER

Keywords: Categorization; etc

Introduction

One of the most intriguing capabilities of human cognition is the ability to creatively generate new ideas and concepts. The creative use of knowledge has, however, only infrequently been the subject of inquiry in the field of categorization. To date, most of what we understand about human concept generation comes from foundational studies in the field of creative cognition (e.g., Marsh, Ward, & Landau, 1999; Smith, Ward, & Schumacher, 1993; Ward, 1994, 1995; Ward, Patterson, Sifonis, Dodds, & Saunders, 2002), though some work has recently been conducted using traditional classification learning approaches (Jern & Kemp, 2013).

Existing work on the creative use of conceptual knowledge predominantly explores the role of prior knowledge in generating novel concepts. The core experimental phenomenon is that people generate categories possessing the same distributional properties as known categories. For example, in a canonical study Ward (1994) asked participants to draw and describe novel species of plants and animals that might exist on other planets. Generation was strongly constrained by prior knowledge of Earth plants and animals: people generated alien species using the same structural forms found on Earth (e.g., eyes, legs, wings) and possessing the same feature correlations observed on Earth (e.g., feathers co-occur with wings). Likewise, exemplars drawn from the same species were less variable than animals drawn from different species, and possessed features that were adaptive for their environment (as is the case on Earth).

Similar results were obtained more recently in a study conducted within a traditional categorization paradigm. Jern and Kemp (2013) trained participants on a set of experimenter-defined categories composed of exemplars within an artificial three-dimensional domain (e.g., 2D shapes varying in size, hue, and saturation). After a short training phase, participants were asked to generate exemplars from a new category. Participants were provided with a set of scales to adjust the feature values of each generated stimulus, and were given unlimited time to create each example. As in the classic Ward (1994) experiment, Jern and Kemp (2013) found that people generated artificial categories possessing the same feature variance and correlations as the experimenter-defined categories in the domain.

Results such as these have motivated formal accounts of generation that explicitly invoke the idea that people re-purpose existing knowledge to generate something new – either by copying-and-tweaking an observation retrieved from memory (Ward, 1995; Ward et al., 2002), or by abstracting the distributional structure of categories in the domain (Jern & Kemp, 2013).

In this paper, we introduce a novel exemplar-based approach to category generation, the PACKER model (*Producing Alike and Contrasting Knowledge using Exemplar Representations*), which creates new categories by balancing two constraints: (1) new categories should be different from known categories (minimizing between-class similarity), and (2) new categories should be internally coherent (maximizing within-class similarity). As such, PACKER is a significant departure from previous accounts of generation – rather than proposing that people create new categories by re-purposing existing knowledge, PACKER explains generation exclusively in terms of the well-studied mechanics of exemplar representations and therefore possesses a rich connection to the wider body of research on category learning.

In the sections below, we formally describe the PACKER model. We then explore the model's predictions using a behavioral experiment and simulations. Rather than formally evaluate PACKER on the benchmarks reviewed above, we instead focus on testing predictions made by PACKER that distinguish the model from previous approaches.

PACKER: An Exemplar Model

The PACKER model is an extension of the highly influential Generalized Context Model of category learning (GCM; Nosofsky, 1984). The model assumes that each category is represented by a collection of exemplars within in a k-dimensional psychological space, and that generation is constrained by both similarity to members of the target category (the category in which a stimulus is being generated) as well as similarity to members of other categories.

As in the GCM, the similarity between two examples, $s(x_i, x_j)$, is defined as an inverse-exponential function of distance:

$$s(x_i,x_j) = exp(-c\sum_k |x_{ik} - x_{jk}|w_k) \tag{1}$$

where w_k is a vector of attention parameters ($\sum_k w_k = 1$), weighting to the importance of each feature in the similarity computation, and c (< 0) is a specificity parameter controlling the spread of exemplar generalization. The attention weights are of interest in explaining individual differences in generation strategy (discussed below), but for our formal simulations they are set uniformly.

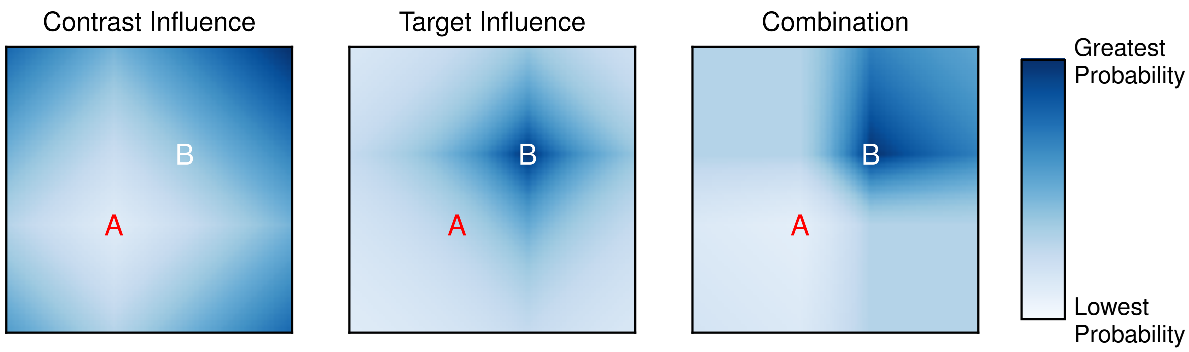

Figure 1: PACKER generation of a category 'B' example, following exposure to one member of category 'A' and one member of category 'B'. *Left*: Predictions given $\{\phi = -1, \gamma = 0\}$ (contrast influence only). *Center*: Predictions given $\{\phi = 0, \gamma = 1\}$ (target influence only). *Right*: Predictions given $\{\phi = -1, \gamma = 1\}$ (both constraints considered).

When prompted to generate a new example, the model considers both the similarity to examples from other categories as well as the similarity to examples in the target category. More formally, the aggregated similarity a between generation candidate y and the model's stored exemplars x can be computed as:

$$a(y,x) = \sum_j \hat{f}_j s(y, x_j) \quad (2)$$

where $\hat{f}_j$ is a function specifying each example's degree of contribution toward generation. Although $\hat{f}_j$ may be set arbitrarily, in PACKER it is set according to class assignment. For known members of contrast categories, $\hat{f}_j = \phi$. For known members of the target category, $\hat{f}_j = \gamma$. ϕ and γ are free parameters ($-\infty \leq \phi, \gamma \leq \infty$) controlling the contribution of target- and contrast-category similarity to generation. Larger absolute values for either parameter produce greater consideration of that type of similarity, with values of 0 producing no effect. Negative values produce a 'repelling' effect (exemplars are less likely to be generated nearby). Conversely, positive values produce a 'pulling' effect (exemplars are more likely to be generated nearby).

Because the model's main proposals are that new categories should be different from existing categories, and exemplars belonging to the same category should be similar to one another, PACKER is commonly simulated under the constraints that $\phi < 0, \gamma > 0$, thus, similarity to contrast categories is effectively subtracted from similarity to the target category. When $-\phi = \gamma$, negative aggregate similarities indicate that y is more similar to members of contrast categories, and positive values indicate y is more similar to the target category.

The probability that a given candidate y will be generated is evaluated using the Luce (1977) choice axiom. Candidates with larger aggregate similarity are more likely to be generated compared to candidates with smaller aggregate similarity:

$$p(y) = \frac{exp(\theta \cdot a(y,x))}{\sum_i exp(\theta \cdot a(y_i, x))} \quad (3)$$

where θ (≥ 0) is a free parameter controlling overall response determinism.

Summary

The PACKER model described above explains exemplar and category generation under a traditional exemplar-similarity approach, proposing that people generate categories by maximizing within-category similarity and minimizing between-category similarity. By consequence, PACKER is unique as a model of generation in several respects. Not only does the model lack mechanisms for copying knowledge of existing categories, but it also specifies precisely how generated classes should *differ* from existing ones. Thus, unlike previous accounts distributional properties of known categories affect generation only insofar as they constrain inter-exemplar similarities. Instead, the *location* of known categories in the domain is crucial because it constrains remaining possible locations for new categories.

Alternative Models

Copy-and-Tweak

The copy-and-tweak model proposes that generation is a two-part process: first, learners retrieve a source exemplar from memory, and then they change it in some way so that the generated item is sufficiently dissimilar. In our model, the probability that a given exemplar z is retrieved from memory is given by:

$$p(z) = exp(\hat{f}_z) / \sum_z exp(\hat{f}_z) \quad (4)$$

$\hat{f}_z$ is a function specifying each item's relative chance of being selected. As in Equation 2, $\hat{f}_j$ may be set arbitrarily. However, in our simulations, $\hat{f}_z = -\gamma$ for members of contrast categories, and $\hat{f}_z = \gamma$ for members of the target category. The resulting free parameter γ (> 0) thus controls the probability that a member of the target category will be retrieved from memory.

After a source example is retrieved, the similarity between generation candidates y and the source is computed as per Equation 1. The model's goal is to generate an item that is similar to z, but exceeds a dissimilarity threshold (such that the new item is similar-but-not-too-similar). Formally, the probability that y will be generated based on a source z, is given by:

$$p(y|z) = \frac{exp(\theta \cdot s(y,z))g(y,z,\tau)}{\sum_i exp(\theta \cdot s(y_i,z))g(y_i,z,\tau)} \quad (5)$$

where θ is a response determinism parameter. g is a function specifying whether y is sufficiently dissimilar to z. The dissimilarity threshold is specified by a free parameter, τ $(0 < \tau \leq 1)$.zZ $g = 1$ when $s(y,z) < \tau$ and $g = 0$ otherwise.

To obtain predictions not depending on a given source example, the model's predictions can be aggregated over all possible sources:

$$p(y) = \sum_z p(z)p(y|z) \quad (6)$$

Thus, the probability that an example y will be selected is a function of its similarity to each of the possible sources, as well as the probability each source will be retrieved.

Hierarchical Sampling

In the Jern and Kemp (2013) model, examples are viewed as samples of an underlying category distribution, which is in turn viewed as a sample of an underlying domain distribution. As in the implementation reported by Jern and Kemp (2013), we assume that exemplars are distributed according to a multivariate normal distribution with parameters (μ, Σ). To obtain a conjugate model, we assume that the domain of categories is inverse-Wishart distributed. Space does not allow for a full derivation of this model, and so we will simply report how generated category parameters (μ_B, Σ_B) are inferred.

Behavioral Experiment

The PACKER model predicts The behavioral experiment described below was designed to test whether the *location* of contrast categories (as opposed to their structure), influences generation. The experiment follows the paradigm developed by Jern and Kemp (2013): first, participants are exposed to members of a known category ('Alpha', or 'A'), and are then asked to generate exemplars belonging to a new category ('Beta', or 'B').

We developed two Alpha categories (see Figure 2), instantiating the two conditions of the experiment. In both conditions, members of the Alpha category are tightly clustered, with equal variance on both features and no correlation between features. Our manipulation is fairly 'weak': the conditions have only a slight difference in the Y-Axis position of the Alpha category. In the 'Middle' condition the Alpha category is placed in the center of the space, in the 'Bottom' condition the Alpha category is placed in the bottom-center of the space.

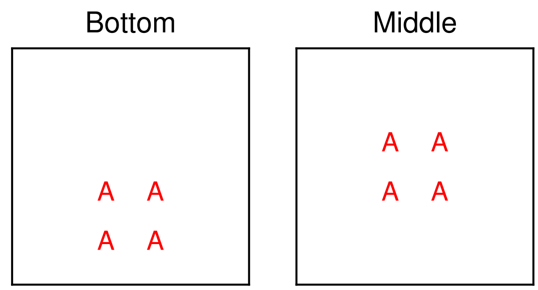

Figure 2: Conditions tested in the behavioral experiment.

Although our manipulation is minimal, the PACKER model is capable of predicting strong between-condition differences. The PACKER model proposes that the nature of the space not occupied by the Alpha category determines where members of the Beta category are likely to be generated. Thus, the lower areas of the stimulus space should be less frequently used for generation in the Bottom condition compared to the Middle (as these areas possess greater similarity to the Bottom Alpha category). Conversely, the upper areas of the stimulus space should be used for generation more frequently in the Bottom condition compared to Middle.

More generally, the PACKER model proposes that the probability a stimulus y will be generated is a function of its similarity to contrast categories *and* to members of the target category. Two more general predictions (not specific to either condition) follow from this proposal: (1) the location of beta examples should be positively related to distance from the Alpha category, (2) Beta examples should be more similar to one another than they are to members of the Alpha category.

Participants & Materials

We recruited 122 participants from Amazon Mechanical Turk. All participants were located in the United States. 61 Participants were assigned to the 'Middle' condition, 61 were assigned to the 'Bottom' condition. Stimuli were squares varying in color (RGB 25–230) and size (3.0–5.8cm). The assignment of perceptual features (color, size) to axes of the domain space (x, y) was counterbalanced across participants.

Procedure

Participants began the experiment with a short training phase (3 blocks of 4 trials), where they iteratively observed exemplars belonging to the 'Alpha' category. Participants were instructed to learn as much as they can about the Alpha category, and that they would answer a series of test questions afterwards. On each trial, a single exemplar from the Alpha category was presented, and participants were given as much time as they desired before moving on. Participants were shown a preview of the range of possible colors and sizes prior to training.

Following the training phase, participants were asked to generate four examples belonging to another category called

Table 1: Behavioral results.

Middle	Used top rows	No top rows
Used bottom rows	31	18
No bottom rows	11	1

Bottom	Used top rows	No top rows
Used bottom rows	22	8
No bottom rows	29	2

'Beta'. Participants were instructed that members of the Beta category could be quite similar or different depending on what they think makes the most sense for the category, but that they were not allowed to make the same example twice.

As in the Jern and Kemp (2013) experiments, generation was completed using a sliding-scale interface. Participants were presented with two scales that could be used to change the color and size of the generated example. An on-screen preview of the example updated whenever one of the features was changed. Participants could generate any example along an evenly-spaced 9x9 grid, and they could change the example as much as they wanted before moving on. Neither the members of the Alpha category nor the previously generated Beta examples were visible during generation. Participants were not informed if they generated a member of the Alpha category, but could not complete the trial if they generated the same Beta two times.

Results

Because the conditions differ only in their location along the vertical axis (see Figure 2), our main interest is in the generation of Beta examples above and below the contrast category. A highly detailed depiction of our results is shown in Figure 3, wherein each participant's generated category is depicted in terms of its use of the vertical axis. As is evident in Figure 3, we observed a great deal of individual differences in generation strategy: whereas some participants generated all four Beta examples within a narrow range of the Y axis, others generated Beta examples along a fairly wide range.

To more specifically evaluate the key predictions made by PACKER, we determined the number of participants in each condition who placed at least one Beta example within the top 'rows' of the space (one of the top two possible Y axis values), as well as the number of participants who placed at least one example in one of the bottom rows. The resulting contingency tables data are shown in Table 1.

Fisher's Exact Tests reveal that a greater number of participants assigned to the Middle condition generated a Beta example in the bottom rows of the space, $p < 0.001$. Participants in the Bottom condition were more marginally more likely to generate a Beta example in the top rows of the space, $p = 0.088$. The conditions did not differ in the number of participants who placed Beta examples in the top rows *and* the bottom rows, $p = 0.14$. Finally, a greater number of Middle participants generated all four Beta examples in the bottom rows of the space, $p = 0.003$, and a greater number of Bottom participants generated all four examples in the top rows of the space, $p = 0.017$.

These results are broadly supportive of PACKER's predictions, and are not predicted by existing accounts (i.e., Jern & Kemp, 2013). However, the results described above are somewhat commonsense: they simply demonstrate that Beta examples tend to be generated in areas not occupied by the Alphas. Beyond these effects, we observed a great deal of differences in generation strategy. Figure 4 depicts the average difference in range between the two features (e.g., horizontal – vertical) across the participant-generated categories, with respect to the location of each category's members. These data reveal highly systematic patterns of generation: whereas many participants generated categories with more vertical range (i.e., 'column'-like categories), others generated categories with horizontal range ('row'-like categories). Furthermore, these two different types of categories appear in distinct locations of the domain space. Whereas vertically aligned categories more often appear to the left or right of the Alpha class, horizontally aligned categories appear above and below the Alpha class. Thus, even beyond generating Betas in locations not occupied by the Alphas, participants appear to modify the internal structure of their categories in order to maximize distance from the Alphas.

Summary

Our behavioral experiment revealed strong effects of the role of category location on generation. Not only are participants more likely to generate Betas in locations

Simulations

We conducted formal simulations to our data using the PACKER model described above, as well as an exemplar-based *copy-and-tweak* model, and conjugate implementation of the hierarchical sampling model proposed by Jern and Kemp (2013). Below, we briefly describe the formal basis of the copy-and-tweak and hierarchical sampling models.

Model-Fitting

In order to obtain an overall sense of each model's ability to explain our results, we fitted each model trial-wise, maximizing the log-likelihood of the its predictions against our results. For each trial, models were initialized with the participant's Alpha category (Middle or Bottom), and the configuration of the participant's Beta category during that trial. Four parameters were fitted for the PACKER model: c, ϕ, γ, and θ. Likewise, four parameters were fitted to the Jern and Kemp (2013) model: κ, γ, ν, and θ. Parameters were not allowed to vary between participants or conditions – the goal was to obtain the best-fit to our entire dataset.

Discussion

The creative use of conceptual knowledge is a highly intriguing yet understudied topic in category learning research. In

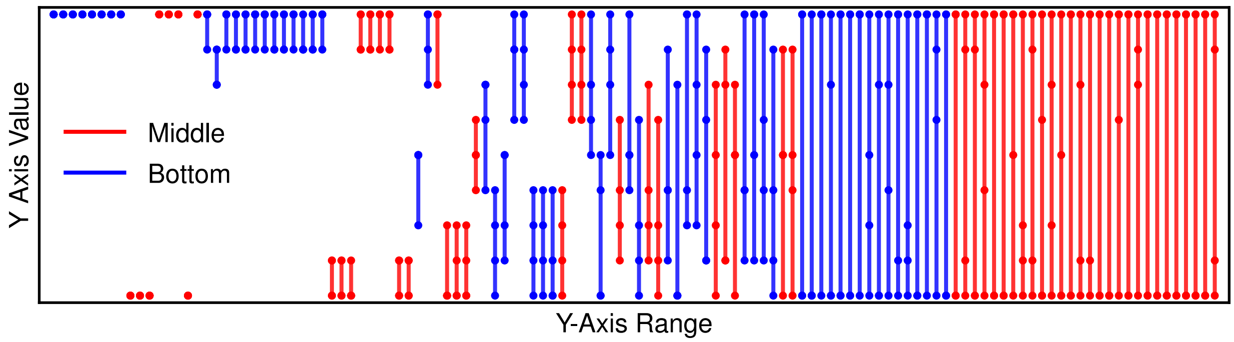

Figure 3: Behavioral results. Each line shows the minimum and maximum value of a generated category along the Y (vertical) axis. Dots along each line represent the positions of individual exemplars in the category, and each participant's category is shown on a separate line. Participants are sorted by overall Y axis range, and then by condition.

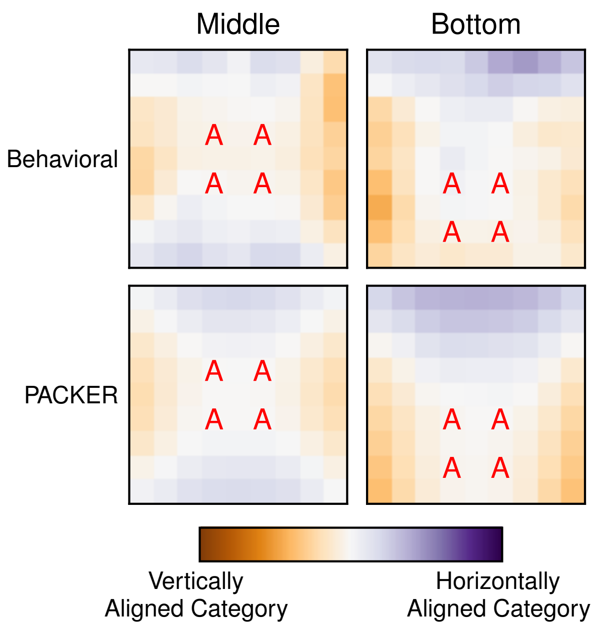

Figure 4: Generated category structure as a function of location in the domain space. Orange areas in each gradient correspond to stimuli that were commonly generated into category possessing greater vertical (Y-Axis) range. Purple areas correspond to categories possessing greater horizontal range. White areas correspond to equal range along both features (or infrequent generation).

this paper, we presented a novel exemplar-based approach to explaining category generation.

Acknowledgments

Support for this research was provided by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin - Madison with funding from the Wisconsin Alumni Research Foundation. We would like to thank Kenneth Kurtz for helpful comments on this work. We also thank Alan Jern & Charles Kemp for providing code and data from their experiments.

References

Jern, A., & Kemp, C. (2013). A probabilistic account of exemplar and category generation. *Cognitive Psychology*, *66*(1), 85–125.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of mathematical psychology*, *15*(3), 215–233.

Marsh, R. L., Ward, T. B., & Landau, J. D. (1999). The inadvertent use of prior knowledge in a generative cognitive task. *Memory & Cognition*, *27*(1), 94–105.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, memory, and cognition*, *10*(1), 104.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, *21*(6), 837–845.

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*(1), 1–40.

Ward, T. B. (1995). Whats old about new ideas. *The creative cognition approach*, 157–178.

Ward, T. B., Patterson, M. J., Sifonis, C. M., Dodds, R. A., & Saunders, K. N. (2002). The role of graded category structure in imaginative thought. *Memory & Cognition*, *30*(2), 199–216.