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, infrequently been the subject of scientific inquiry in the field of categorization. Most research on concepts and categorization has instead conformed to a traditional artificial classification learning paradigm (Kurtz, 2015), wherein the primary use of category knowledge is to discriminate between two or more experimenter-defined classes.

In one of the earliest explorations on the creative use of conceptual knowledge, Ward & colleagues (Marsh, Ward, & Landau, 1999; Smith, Ward, & Schumacher, 1993; Ward, 1994, 1995; Ward, Patterson, Sifonis, Dodds, & Saunders, 2002) provided clear evidence of the role of prior knowledge in concept generation. 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), exemplars tended to obey the same feature correlations observed on Earth (e.g., feathers tended to co-occur with wings), and exemplars drawn from the same species were less variable than animals drawn from different species (as is the case on Earth).

An early account of these these results proposed that people complete the generation task by retrieving memories of known Earth plants and animals, and slightly changing them to create something new (Ward, 1995; Ward et al., 2002). This account, which we will refer to as the *copy-and-tweak* account can be casted within a formal exemplar approach (see Jern & Kemp, 2013). However, due to the nature of the response formats used (e.g., drawing and describing examples), and due the prior knowledge assumed (e.g., knowledge of plants and animals on Earth), these experiments are not suitable for formal modeling.

In more recent work, Jern and Kemp (2013) reported several experiments on exemplar and category generation within an artificial category learning paradigm (and are thus suitable for formal modeling). In their paradigm, participants are first exposed to members from a set of known categories within a reduced feature space (e.g., size, hue, saturation). After observing the assigned examples, participants are asked to generate exemplars from a new category within the domain. Participants provided with a set of scales scales that can be used to adjust the feature values of the generated stimulus (one scale per feature), and are given unlimited time to create each example. As in the classic Ward (1994) experiment, Jern and Kemp (2013) found that people generated categories possessing the same feature correlations as the known categories in the domain.

Jern and Kemp (2013) developed a hierarchical model of generation in order to explain their results. In their model, categories are represented as probability distributions in a multidimensional space (with each stimulus feature corresponding to a dimension of the space). Observed exemplars (represented as points in this space) are viewed as samples from an underlying category distribution (e.g., a multivariate normal distribution). These category distributions are in turn viewed as samples from an underlying domain distribution, which represents the common distributional characteristics among known categories (e.g., feature variance, feature correlations). When asked to generate a novel class, people are thought to sample a category distribution from the domain distribution, and then sample exemplars from the category distribution. Thus, exemplars generated into novel classes tend to obey the same distributional properties of known categories.

The work presented by Ward & Colleagues, as well as that of Jern and Kemp (2013), undoubtedly captures a core principle of the creative use of conceptual knowledge: category generation is informed by prior knowledge of concepts in the target domain. The two models described above (copy-and-tweak, hierarchical sampling) are conceptually similar in that they assume learners *copy* something about existing categories: individuals either directly copy-and-tweak stored observations, or copy the distributional structure of entire categories.

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). PACKER is a significant departure from the accounts of generation reviewed above in that it instantiates assumptions about how exemplars in novel classes should *differ* from existing categories (rather than how they should align to them). Likewise, the PACKER model is
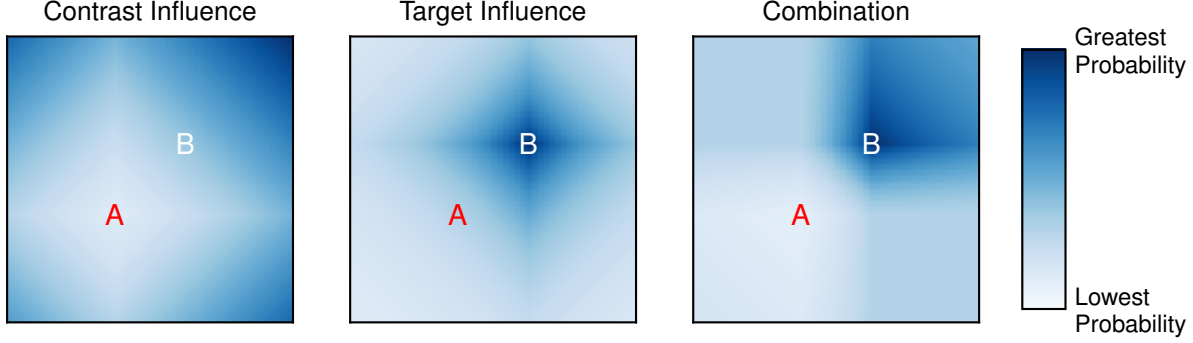
Figure 1: PACKER generation of a category 'B' example, following exposure to one member of category 'A' and one member of category 'B'. *Left*: Predictions given $\{\phi = -1, \gamma = 0\}$ (contrast influence only). *Center*: Predictions given $\{\phi = 0, \gamma = 1\}$ (target influence only). *Right*: Predictions given $\{\phi = -1, \gamma = 1\}$ (both constraints considered).

of interest because it follows directly from the well-studied exemplar-similarity framework that has led the field for over three decades.

In the sections below, we formally describe the PACKER model. We then report a behavioral experiment testing the model's predictions, and report formal simulations comparing PACKER to the copy-and-tweak and hierarchical sampling models.

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:

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

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. The c parameter controls the specificity of examples: larger values of c produce more *specific* representations (exemplars that do not generalize broadly). The attention weights are of interest in explaining individual differences in generation strategy (discussed below), but for our formal simulations they are set uniformly.

When prompted to generate a new example, the model considers both the summed similarity to examples from other categories as well as the summed similarity to examples in the target category. More formally, the summed similarity ss between generation candidate y and the model's stored exem-

plars x can be computed as:

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

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 summed similarities indicate that y is more similar to members of contrast categories, and positive values indicate y is more similar to members of the target category.

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

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

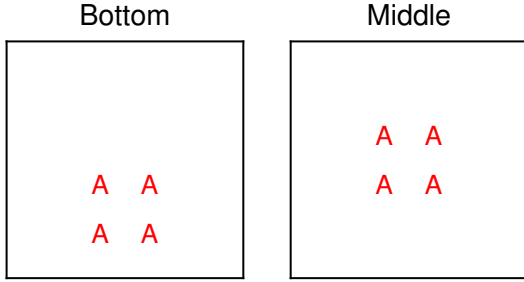where θ (≥ 0) is a free parameter controlling overall response determinism.

Figure 2: Conditions tested in the behavioral experiment.

Figure 3: Sample stimuli (not drawn to scale).

Summary

The PACKER model described above explains exemplar and category generation under a traditional exemplar-similarity approach. The model proposes that people generate categories by maximizing within-category similarity and minimizing between-category similarity.

PACKER is unique as a model of generation because it specifies precisely how generated classes should *differ* from existing ones – the copy-and-tweak and hierarchical sampling accounts explain why generated classes share many properties with known classes, but not how they should be distinct. Likewise, PACKER is unique in that the 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. In the sections below, we report a behavioral experiment testing these predictions. We then report formal simulations with PACKER, a copy-and-tweak model, and the Jern and Kemp (2013) hierarchical sampling model.

Behavioral Experiment

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.

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), see Figure 3 for samples. 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 '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

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

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 4, wherein each participant's generated category is depicted in terms of its use of the vertical axis. As is evident in Figure 4, 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 predic-

tions, 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 5 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.

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: $p(z) = exp(\hat{f}_z)/\sum_z exp(\hat{f}_z)$. $\hat{f}_z$ 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:
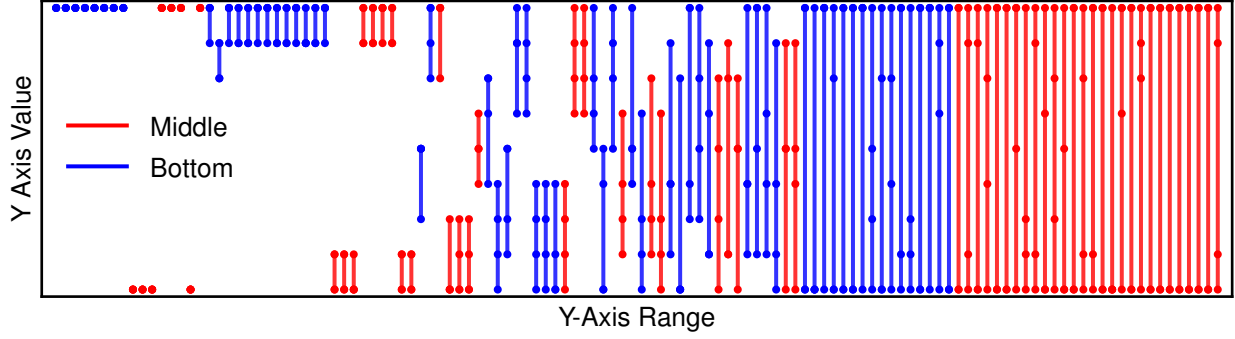
Figure 4: Behavioral results. Each line shows the minimum and maximum value of a generated category along the Y (vertical) axis. Dots along each line represent the positions of individual exemplars in the category, and each participant's category is shown on a separate line. Participants are sorted by overall Y axis range, and then by condition.
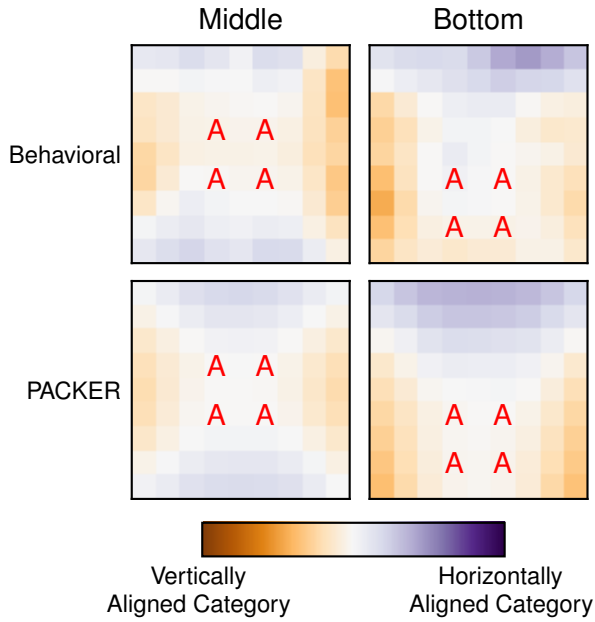

Figure 5: Generated category structure as a function of location in the domain space. Orange areas in each gradient correspond to stimuli that were commonly generated into category possessing greater vertical (Y-Axis) range. Purple areas correspond to categories possessing greater horizontal range. White areas correspond to equal range along both features (or infrequent generation).

$$
\begin{cases}
p(y|z) \propto exp(\theta \cdot s(y,z)), & \text{if } s(y,z) \leq \tau \\
p(y|z) = 0, & \text{otherwise}
\end{cases}
\tag{4}
$$

where θ is a response determinism parameter, and τ is a similarity threshold ($0 < \tau \leq 1$) specifying the amount of similarity tolerated between y and z. Values of $p(y|z)$ are then normalized such that $\sum_i p(y_i|z) = 1$. To obtain predictions

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

$$
p(y) = \sum_z p(z)p(y|z)
\tag{5}
$$

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.

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

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

Acknowledgments

References

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

Kurtz, K. J. (2015). Human category learning: Toward a broader explanatory account. *Psychology of Learning and Motivation*, *63*, 77–114.

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.