A Prototype Model that Learns and Generalizes Medin, Altom, Edelson & Freko (1982) XOR Category Structure As Humans Do

Toshihiko Matsuka (Toshihiko.Matsuka@stevens.edu)
Jeffery V. Nickerson (jnickerson@stevens.edu)
Jiun-Yin Jian (Jiun-Yin.Jian@stevens.edu)
Center for Decision Technologies
Stevens Institute of Technology
Hoboken, NJ 07030 USA

Abstract

The computational modeling literature suggests that *Exemplar* models of categorization often replicated psychological phenomena better than *Prototype* models. However, those prototype models may have failed because the models' important information processing mechanisms were misspecified. Here we introduce a new prototype model with complex yet realistic learning and selective attention processes. Its attention processes (a) have a prototype specific attention coverage structure and (b) are sensitive to correlations among feature dimensions. In simulation studies, CASPRE, our new prototype model, replicates the results of two important classical empirical studies.

Introduction

The issue of *Internal Representation* has been one of the central theoretical interests and debates in the human categorization research. While many competing theories on internal representations have been advanced, most studies have been dedicated to evaluating the descriptive validity of exemplar theory and/or prototype theory (e.g. Minda & Smith, 2002, Zaki, Nosofsky, Stanton, & Cohen, 2002). Previous modeling studies suggest that *Exemplar* theory is descriptively more valid than *Prototype* theory. More precisely, computational models built upon exemplar theory produced more successful replications of observed phenomena than prototype models. One example of a categorization problem showing a limitation of traditional prototype models is a simple XOR logic stimulus set (i.e., [00, 11] for Category A and [01, 10] for Category B), whose prototypes for Categories A and B are theoretically identical (i.e., [0.5, 0.5]). By internally representing categorical knowledge with these identical prototypes, previous prototype models with traditional selective attention mechanisms failed to categorize or learn to categorize these stimuli. This is because the models' mathematical formulations yield identical psychological similarity measures for any input stimulus to both prototypes, providing no constructive information for categorization.

However, by employing a general and exploratory modeling method, Matsuka (2004) revealed that a model with prototype-like internal representation can learn XOR-logic if the model integrates complex attention allocation mechanisms, namely the capability of attending correlations among feature dimensions and prototype-specific selective attention allocation processes (i.e. each prototype has a customized attention pattern).

In the present research, we introduce a new prototype model for human category learning based on Matsuka's (2004) model. We first test whether or not the new prototype model is able learn and generalize a stimulus structure containing XOR logic like humans do. Then, we conduct an additional simulation study to test the descriptive validity of our new attention mechanisms.

A New Model: CASPRE

Overview

We call our new model CASPRE (Category learning with Attention augmented Simplistic Prototype Representation). It is a cognitive model based on prototype theory. It assumes that categorical knowledge is organized by small numbers of prototypes and that humans utilize psychological similarities between input stimulus and prototypes for categorization.

CASPRE is comprised of two components. The first component assumes a somewhat complex attention process, namely a local attention process (each prototype has a customized selective attention process) and sensitivity to correlations among features to form attention-augmented prototype representations. Selective attention may be interpreted as processes of mental rotation and psychological scaling of proximities or similarities between input stimuli and prototypes. In other words, in CASPRE each prototype is augmented with a customized selective attention process to form a uniquely shaped and oriented prototype *conceptual field*. Unlike traditional prototype models, characteristics of prototypes in CASPRE cannot be explained by centroids alone, but by a combination of centroids and within-prototype psychological scaling processes.

The second component is the principle of simplicity in high-order human cognition. One plausible theoretical justification of prototype theory is that its compact representation of knowledge allows a limited-processing-capacity human brain to handle rich information (whereas exemplar theory assumes that humans utilize information on many if not all exemplars they have previously encountered to categorize an input stimulus). The incorporation of local attention mechanisms could inflate knowledge complexity. To stem the growth of unnecessary complexity, CASPRE incorporates a multi-objective learning algorithm. It tries to acquire manageably simple yet sufficiently accurate concepts.

Assumptions on Attention Processes

Local Attention Coverage Most quantitative models of categorization (e.g. Nosofsky 1986, Smith & Minda 2002) and category learning (e.g. Kruschke, 1992; Love, Medin, Gureckis, 2004) assume that selective attention processes are uniformly applied to all reference points (e.g. exemplars or prototypes). In other words, the models utilize the same attention at all locations along a dimension in the representational space, indicating that the attention coverage is global. However, some laboratory experiments (e.g. Aha & Goldstone, 1992) suggest attention could be specific to the region along a dimension in the representational space, indicating that the attention coverage is in fact local.

Recent studies (Corter & Matsuka 2004; Matsuka & Corter, 2006) provide more direct evidence of differential attention allocation patterns by using the MouseLab experimental paradigm (Bettman, Johnson, Luce, & Payne, 1993). Computational modeling research also indicates that some phenomena require the local attention coverage system in order to replicate some psychological phenomena (e.g. Kruschke, 2002; Sakamoto, Matsuka & Love, 2004). Although more thorough empirical studies on locality-vs-globality in selective attention processes may be necessary, the results of empirical and simulation studies provide sufficient evidence of the possibility of a local attention coverage system.

Attention to correlations among feature dimensions Another notable selective attention mechanism widely applied to models of categorization, yet not extensively sought for alternative possibilities, is the independent dimension-by-dimension selective attention process. Virtually all models of categorization and category learning assume that humans pay no attention to correlations among feature dimensions nor psychologically rotate feature space during categorization: the humans' perceived psychological space is assumed not only to be logically orthogonal, but also to be identical to the space that researchers define. However, Ashby and Maddox (1992) suggested that the mapping of physical coordinates of stimuli to psychologically perceived ones does not have to be linear (orthogonal). In addition, humans are known to be capable of carrying out mental rotation.

Other empirical studies indicate that humans are indeed sensitive to correlations between feature dimensions in categorization (e.g. Anderson & Fincham, 1996; Chin-Parker & Ross, 2002). Although the sensitivity to correlations may not necessarily directly translate to attention in a strict sense, the selective sensitivity to a particular feature dimension has been traditionally interpreted as selective attention (e.g. Kruschke 1992; Nosofsky 1986). Thus, the sensitivity to a particular combination of feature dimensions can be interpreted as attention to the correlations. This is the basis for CASPRE's assumption that humans are indeed capable and do pay attention to correlations among feature dimensions if needed.

Forward Algorithm (Categorization)

CASPRE's forward algorithm resembles that of ALCOVE (Kruschke, 1992), one of the most successful models of category learning. However, there are two crucial differences. First, in CASPRE, it is assumed that people utilize psychological distances or similarities between input stimuli and prototypes (vs. exemplars in ALCOVE). Second, in CASPRE, psychological similarity or distance (d_j) between an input stimulus (*x*) and prototype *j* (π_j) are defined by Mahalanobis distance (in quadratic form) between them, allowing for sensitivity to correlations among features dimensions (vs. Minkowski *r* metric in ALCOVE). Therefore,

$$d_j(x) = \sum_i^I \sum_m^I \alpha_{im}^j (\pi_{ji} - x_i)(\pi_{jm} - x_m) \quad (1)$$

where α^j defines directions and strengths of attention field for π_j, subscripts *i* and *m* indicate feature dimensions, and *I* is the number of feature dimensions. Note that it is assumed that $\alpha_{im} = \alpha_{mi}$, $\alpha_{im}^2 \leq |\alpha_{ii} \cdot \alpha_{mm}|$, and $\alpha_{ii} \geq 0, \forall i$. For off-diagonal entries (i.e., $i \neq m$), an attention weight can be a negative value, where its signum indicates direction of attention field while its magnitude indicates the strength of attention. Psychological distance measures activate prototype units by the following function:

$$h_j = \exp(-c \cdot d_j(x)) \quad (2)$$

where *c* controls overall sensitivity. Activations of prototype units are then fed forwarded to category output nodes, or

$$O_k(x) = \sum_j w_{kj} h_j; \quad (3)$$

where w_{kj} is an association weights between π_j and category node *k*. The output activations will be used to obtain the response probability by the following function:

$$P(k) = \frac{\exp(\phi O_k)}{\sum_l \exp(\phi O_l)} \quad (4)$$

where ϕ scales the decisiveness of response (e.g. Kruschke, 1992). In short, CASPRE assumes that humans utilize psychological similarity between input object (*x*) and prototypes (π), psychologically scaled by correlation sensitive prototype-specific selective attention processes, as evidence for categorizing the input instance into the most probable category.

In this paper we refer to CASPRE's constants that experimenters can manipulate (e.g. λs) as free parameters, and its learnable variables (e.g., $\boldsymbol{w}$ & $\boldsymbol{\alpha}$) as coefficients to avoid confusion.

Backward Algorithm (Learning)

In CASPRE, human learning is not considered an error minimization process, but an optimization of a subjectively and contextually defined utility of knowledge or concepts being acquired. There are many functions or sets of functions appropriately defined for describing a variety of contextual factors, including motivation. However, the rudimentary set of objective functions for CASPRE consists of two elements: concept accuracy and concept simplicity. That is, CASPRE assumes that in ordinary situations humans would prefer and try to acquire manageably simple yet sufficiently accurate knowledge. In order to integrate this multi-objective learning, CASPRE incorporates the gradient descent version of Stochastic Context-Dependent Learning framework (SCODEL: Matsuka, 2005a, 2005b) (Note: SCODEL problems are framed as minimization problems, thus higher values indicate poorer concept utility.).

The minimal set of objective or utility functions for a particular set of coefficients (i.e., concepts or θ), thus, can be formulated as:

$$U(\theta) = \sum_k \frac{1}{2}e_k^2 + \sum_j \sum_i^{I-1} \sum_{m=i+1}^{I} \frac{(\alpha_{im}^j)^2/Z_j}{1+(\alpha_{im}^j)^2/Z_j} \quad (5)$$

where e_k is the difference between the target and predicted outputs for category node *k*, and

$$Z_j = \sum_i^{I-1} \sum_{m=i+1}^{I} (\alpha_{im}^j)^2 \quad (6)$$

The first term in Eq. 5 is a function defining categorization accuracy. The second term is a simplicity bias or an attention elimination function (e.g. Matsuka, 2005b), reducing the number of correlations among dimensions attended on the basis of the *relative* attention strengths. Matsuka (2005b, 2006; Matsuka & Chouchourelou, 2006) discusses a more general utility function, including models for multiple prototypes and various contextual factors.

Learnable coefficients $\boldsymbol{w}$ and $\boldsymbol{\alpha}$ are updated by the following functions,

$$\Delta w_{kj} = -\lambda_w \frac{\partial U}{\partial w_{kj}} + \nu_{kj} = \lambda_w e_k h_j + \nu_{kj} \quad (7)$$

$$\Delta\alpha_{im}^j = -\boldsymbol{\lambda}_\alpha \frac{\partial U}{\partial \alpha_{im}^j} + \nu_{im}^j = -\lambda_\alpha^E \sum_k e_k w_{kj} h_j \delta_{ji} \delta_{jm} c - \lambda_\alpha^A \frac{2\alpha_{im}^j \cdot \sum_{l \notin im} (\alpha_l^j)^2}{\left(2(\alpha_{im}^j)^2 + \sum_{l \notin im}(\alpha_l^j)^2\right)^2} + \nu_{im}^j \quad (8)$$

where $\delta_{ji} = \pi_{ji} - x_i$, λs are learning rates, and $\boldsymbol{\nu}$ are independent Gaussian noise in learning with means equal to zero and some time-decreasing standard deviations (e.g. median $\Delta w^{(T-1)}$ & $\Delta\alpha^{(T-1)}$, where *T* indicates time). Noise is introduced in CASPRE, because recent cognitive modeling studies indicate the importance of stochasticity in human learning for quantitative fits and qualitative interpretations, including probabilistically successful learning, asymmetric utilization of redundant information, and exhibiting arbitrary decisions in learning. (e.g. Matsuka, 2005a, 2005b)

The centroids of category prototypes are updated with a simple competitive learning algorithm (e.g. Kohonen, 2001; Love et al., 2004), where the centroid for only the current category prototype will be updated. Thus,

$$\boldsymbol{\pi}_j^{(T+1)} = \begin{cases} \boldsymbol{\pi}_j^{(T)} + 1/\sqrt{T}(\mathbf{x} - \boldsymbol{\pi}_j^{(T)}) & \text{if } C_\mathbf{x} = C_{\boldsymbol{\pi}_j} \\ \boldsymbol{\pi}_j^{(T)} & \text{otherwise.} \end{cases} \quad (9)$$

where *T* indicates time and $C_\mathbf{x}$ indicates a correct category for stimuli *x*.

There are several reasons for incorporating the simplicity bias in CASPRE. Intuitively, a preference for simpler yet sufficiently accurate concepts appears a plausible phenomenon in high-order human cognition. More importantly, some empirical studies suggested its possibility (Corter & Matsuka, 2004; Matsuka & Corter, 2006). In addition, this bias toward simpler yet sufficiently accurate concepts (vs. complex but marginally more accurate concepts) might have resulted in the emergence of other psychological phenomena such as Basic Categories in human cognition (Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976).

Table 1: Schematic Representation of Stimulus Set Used in Simulation 1 (Medin et al., 1982).

Training					Transfer				
	D1	D2	D3	D4		D1	D2	D3	D4
A1	1	1	1	1	T1	0	0	0	0
A2	1	1	0	0	T2	0	0	1	1
A3	0	1	1	1	T3	0	1	0	0
A4	1	0	0	0	T4	1	0	1	1
B5	0	0	1	0	T5	1	1	1	0
B6	0	0	0	1	T6	1	1	0	1
B7	1	0	1	0	T7	0	1	1	0
B8	0	1	0	1	T8	1	0	0	1

Initialization & Number of Free Parameter The association weights are initialized with small Uniform random numbers around zero. The centroids of prototypes are also initialized with small Uniform random number, but around midpoints between minimum and maximum values. For example, if a feature dimension consists of '0' and '1' (e.g. Table 1), then values for each centroid in that dimension are small Uniform random numbers around 0.5. Initial selective attention weight matrices are diagonal matrices with $\alpha_{ii} = I^{-1}, \forall i$.

There are a total of five free parameters in CASPRE, two for its forward process (i.e., c and ϕ) and three for learning process (i.e., λ_w, λ_α^E, λ_α^A). Thus, CASPRE has only one more parameter than ALCOVE.

Simulations

Simulation 1: Replication of Medin et al. (1982)

In Simulation 1, we simulated a classical study in human category learning (Medin, Altom, Edelson & Freko, 1982). Table 1 shows the schematic representation of the stimulus set used in the present simulation study. Note that in order to perfectly categorize the stimulus set, subjects need to memorize all exemplars, acquire XOR logic in Dimensions 3 and 4, or some combination of both.

In the empirical study, subjects were asked to learn to classify eight unique training exemplars (A1 - B4) to either Category A or B with corrective feedback. The training session was followed by a transfer session in which subjects were asked to categorize the eight training exemplars and eight novel exemplars. The observed profile (see Fig.1) indicates subjects tended to exhibit weak XOR-like classification profiles. Because of this XOR-like conceptualization, to our knowledge, no previous prototype models had successfully replicated the observed classification profile.

Methods CASPRE was run in a simulated training procedure with 50 trial blocks, where each block consisted of a random presentation of the eight unique training exemplars exactly once, in order to learn the correct classification responses for the stimulus set. Note that in the original experiment (Medin et al., 1982), subjects were allowed to study all eight training stimuli simultaneously for 10 or 15 minutes, depending on learning speed. After the training session, CASPRE was run in a simulated transfer procedure with one transfer block, where all 16 exemplars were presented exactly

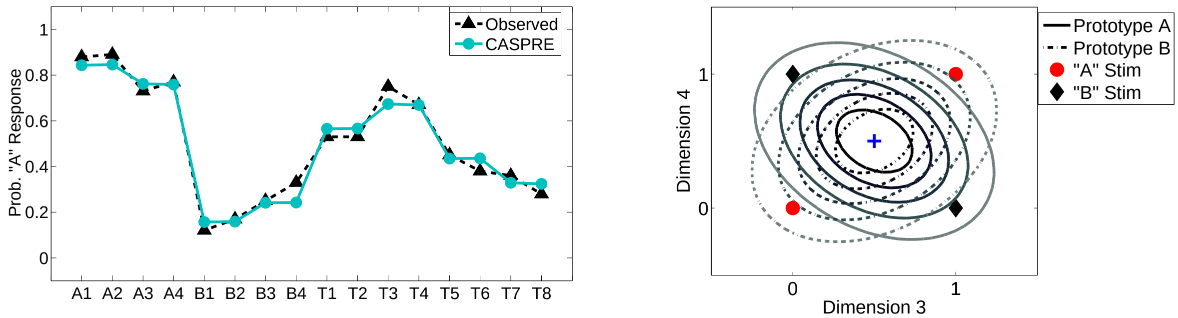

Figure 1: Result of Simulation 1. *Left*: Predicted classification profile by CASPRE and its criterion profiles reported in Medin et al., (1982). *Right*: Predicted prototype conceptual fields by CASPRE.

once. No corrective feedback was given in the transfer session, and thus no learning occurred within the session.

The model parameters were optimized so that the sum of squared errors (SSE) between observed and predicted classification profiles were minimized. The selected parameters were as follows: $c = 3.85$, $\phi = 3.94$, $\lambda_w = 0.38$, $\lambda_\alpha^E = 0.02$, $\lambda_\alpha^A = 0.02$. There were a total of 100 simulated subjects.

Results & Interpretations

Figure 1 shows CASPRE's predicted classification profiles along with the observed profile of Medin et al. (1982). The figure indicates CASPRE successfully replicated the observed response profile (SSE=0.028). The average *relative* attention allocations are shown in Table 2. The attention coverage structures for Prototypes A and B for Dimensions 3 and 4 (i.e., the XOR dimensions) are shown as a contour plot in Fig.1. For these dimensions, Prototype A and B conceptual fields (attention coverage areas) are tilted approximately 45 and -45 degrees, respectively, making an X-shape attention or conceptual coverage area (this reflects the alpha coefficient weighting s of Eq. 1). One plausible interpretation of such attention coverage areas is that CASPRE paid attention to the directions of deviations from the prototypes with the identical centroid; and perhaps mentally rotated its psychological space to make stimulus space more interpretable. That is, CASPRE mentally rotated the dimensions for both prototypes in almost an identical manner, making each prototype's attention coverage area sensitive to a single rotated dimension.

CASPRE, then, compares an input stimulus and the prototypes to see and use the directions of deviations of the stimulus from prototypes as a criterion for deciding which category the stimulus belongs to. Both prototypes have identical characteristics on Dimension 3 and 4, yet because of their local correlation sensitive attention coverage structure, the perceived similarities between the input stimulus to Prototype A and Prototype B were different depending on the directions of deviation from the Prototypes.

We conducted an additional simulation study without the (relative) attention elimination process by fixing λ_α^A at 0 to test the importance of the process. The modified model's SSE was approximately 8 times worse than the original CASPRE, indicating the attention elimination process was a key process in successfully replicating the observed phenomena.

Table 2: Average Relative Attention Distributions for Medin et al. (1982) XOR Stimulus Set Predicted by CASPRE

Prototype A	D1	D2	D3	D4
		Prototype A		
D1	0.136	0.016	0.007	-0.006
D2	-	0.127	-0.022	0.006
D3	-	-	0.255	-0.081
D4	-	-	-	0.255
Prototype B	D1	D2	D3	D4
D1	0.134	0.009	-0.013	-0.006
D2	-	0.136	0.011	0.007
D3	-	-	0.266	0.081
D4	-	-	-	0.266

Simulation 2 - Filtration Advantage

Previous empirical studies (e.g. Gottwald & Garner, 1975; Kruschke, 1993) showed that the Filtration task (e.g. categorization task with one diagnostic dimension) is easier than the Condensation task (e.g. categorization task with two correlated diagnostics dimensions). Kruschke (1993) presented this phenomenon as evidence for an orthogonal dimensional attention process (i.e., attention is allocated dimension-by-dimension independently, but not dependently). That is, if humans are capable of attending correlations among feature dimensions (or mentally rotating feature space), then the Condensation task becomes equivalent to the Filtration task, and there is no Filtration advantage. If this argument is true, this argument might weaken the case for CASPRE.

However, we expect that CASPRE would correctly replicate the Filtration advantage because of CASPRE's bias toward simpler concepts or coefficient configurations: the bias causes CASPRE to learn to categorize a stimulus structure with one diagnostic dimension (i.e., the Filtration task) much more efficiently than a stimulus structure with two conjunctively diagnostic dimensions (i.e., the Condensation task). The present simulation tests this claim.

Methods Figure 2 shows the stimulus sets used in the present study. CASPRE was run in a simulated training procedure. The procedure consisted of eight trial blocks, where each block consisted of a random presentation of the eight

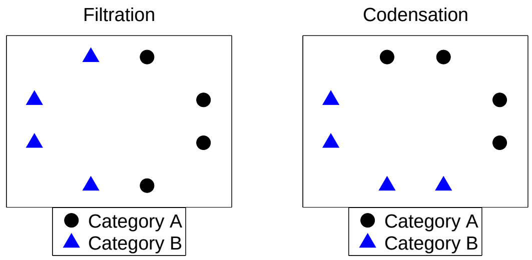

Figure 2: Stimulus Structures used in Simulation 2. *Left* Filtration Stimuli. *Right* Condensation Stimuli

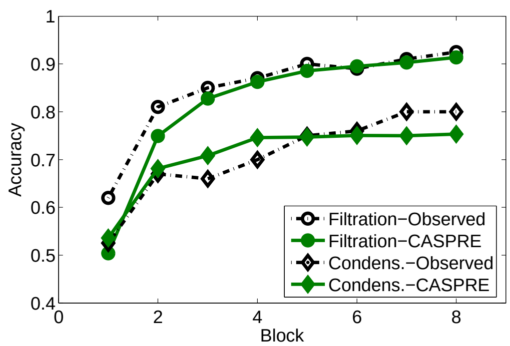

Figure 3: Predicted classification accuracies for the Filtration and Condensation tasks by CASPRE. CASPRE successfully replicated the Filtration advantage.

unique training exemplars exactly once, as to ascertain the correct classification responses for the stimulus set. The model parameters were optimized so that the sum of squared errors (SSE) between observed (Kruschke, 1993) and predicted classification profiles were minimized. The selected parameters were as follows: $c = 1.07$, $\phi = 8.14$, $\lambda_w = 0.96$, $\lambda_\alpha^E = 0.82$, $\lambda_\alpha^A = 0.30$. There were a total of 100 simulated subjects in the present study.

Results & Discussion

Figure 3 shows observed and predicted learning curves for both Filtration and Condensation tasks. CASPRE was able to replicate the Filtration advantage (SSE = 0.027). Although the independent attention allocation principle holds for exemplar models (Kruschke, 1993), the results of Simulation 2 showed that the principle does not necessarily hold for CASPRE.

General Discussion

Model Extension The selective attention process is one of a few cognitive processes that are widely accepted by a majority of cognitive scientists involved in human categorization and category learning. The previous successful models of human category learning incorporate this process in some form. So does CASPRE. However, we have not integrated this process in the process of learning prototypes (i.e., Eq. 9). This was done so in order to make CASPRE to acquire prototypes that meet a traditional definition of a prototype – an average member of a category (e.g. Rosch & Mervis, 1975). Given the impact of selective attention, however, an internally-represented prototype may be only an average of its important and diagnostic feature dimensions, but not of its useless dimensions. This is because there are so many feature dimensions possible for any category; selection may always be taking place. If this is the case, CASPRE's prototype learning or identifying algorithm needs to be modified to incorporate selective attention processes (e.g. using psychologically perceived distances in Eq. 9).

Recently, Matsuka and Nickerson (2006) introduced learning algorithms based on simulated evolutionary processes that offer unique qualitative interpretations of learning process, including (a) combining several ideas into one parsimonious idea, (b) creating "radical" hypothesis, and (c) competition among hypotheses. This learning algorithm may be applied to CASPRE to enhance its qualitative characteristics.

Medin & Schaffer (1972)

One of the most frequently tested stimulus sets in evaluating descriptive validities of models of categorization and category learning is that of Experiment 2 of Medin and Schaffer (1972). In a modeling study, Matsuka (2006) showed that CASPRE indeed performed better than an exemplar model if the criterion classification response profile was estimated based on 30 studies summarized in Smith and Minda (2002). It performed slightly worse when the criterion was the original Medin and Smith's observation. In addition, CASPRE was able to replicate the *A2 advantage* (i.e. a phenomenon in which the less "prototypical" stimulus A2 is more accurately classified than more "prototypical" stimulus A1). This phenomenon has been presented as evidence for exemplar theory and against prototype theory. The argument merits reconsideration because it has been shown that a prototype model, CASPRE, is capable of replicating the phenomenon.

Conclusion

There is long running and heated debate on the descriptive validity of *Exemplar* theory and *Prototype* theory of internal representation of categories. One of the main means of testing the descriptive validity has been computational cognitive models. Because of the less frequently successful replication of some important psychological phenomena by prototype models, many studies have suggested or concluded that Exemplar theory is descriptive more valid than Prototype theory.

We, however, hypothesized that the inadequacy of previous prototype models might be caused by mismatches between three elements: the system of internal representation, the selective attention mechanism, and the routine of learning. In order to test this hypothesis, we introduced two novel mechanisms to a prototype model of human category learning: correlation sensitive local attention process and multi-objective learning (preferring manageably simple yet sufficiently accurate concepts over complex but marginally more accurate concepts). The former mechanism is evident in recent empirical studies (e.g. Chin-Parker & Ross 2002; Corter & Matsuka, 2004). The latter mechanism is based on work suggesting that human learning is not merely characterized by classification error minimization, but by the optimization of subjectively and contextually defined utility of the knowledge being acquired (e.g. Matsuka 2005a; Matsuka & Chouchourelou, 2006).

When we integrated these key mechanisms into a prototype model of category learning, we were able to replicate classical empirical studies. We replicated two important empirical studies. We anticipate conducting additional simulation studies to more thoroughly compare the validity of Exemplar and Prototype theories. This study takes an initial step by facilitating fairer comparisons between prototype and exemplar theories in order to better understand the nature of human categorization processes.

Acknowledgments

This research was supported in part by the Office of Naval Research, Grant # N00014-05-1-00632.

References

Aha, D. W., & Goldstone, R. L. (1992). Concept learning and flexible weighting. In *Proceedings of the 14h Annual Meeting of the Cognitive Science Society.* (pp. 534-539). Hillsdale, NJ. Erlbaum.

Anderson, J. R., & Fincham, J. M. (1996). Categorization and sensitivity to correlation. *Journal of Experimental Psychology: Learning, Memory & Cognition, 22*, 119-128.

Ashby, F. G. & Maddox, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 50-71.

Bettman, J. R., Johnson, E. J., Luce, M. F., & Payne, J. W. (1993). Correlation, conflict, and Choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 931-951

Chin-Parker, S., Ross, B. H. (2002).The effect of category learning on sensitivity to within-category correlations. *Memory & Cognition, 30*, 353-362.

Corter, J. E., & Matsuka, T. (2004). Empirical measures of attention allocation in classification learning: A replication of Medin & Schaffer (1978). Paper presented for the 45th Annual Meeting of the Psychonomic Society. Minneapolis, MN.

Gottwald, R., & Garner, W. R. (1975). Filtering and condensation tasks with integral and separable dimensions. *Perception and Psychophysics, 2*, 50-55.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22-44.

Kruschke, J. E. (1993). Three principles for models of category learning. In G. V. Nakamura, R. Taraban, & D. L. Medin (Eds.), *Categorization by human and machines: The psychology of learning and motivation* (Vol. 29, pp. 57-90). San Diego, CA: Academic Press.

Kohonen, T. (2001). *Self-Organizing Maps.* (3rd ed.) Berlin: Springer.

Love, B. C., Medin, D.L., & Gureckis, T. M. (2004). SUSTAIN: A network model of human category learning. *Psychological Review, 111*, 309-332.

Matsuka, T. (2004). Generalized exploratory model of human category learning. *International Journal of Computational Intelligence, 1*, 7 - 15.

Matsuka, T. (2005a). Simple, individually unique, and context-depending learning methods for models of human category learning. *Behavior Research Methods, 37*, 240 - 255.

Matsuka, T. (2005b) Modeling human learning as context dependent knowledge utility optimization. *Advances in Natural Computation*, LNCS, Vol.3610. (pp. 933-946). Berlin: Springer-Verlag.

Matsuka, T. (2006). A model of category learning with attention augmented simplistic prototype representation. *Advances in Neural Networks*, LNCS, Vol. 3971, (pp 34 - 40). Berlin: Springer-Verlag.

Matsuka, T., & Chouchourelou, A (2006). On the learning algorithms of descriptive models of high-order human cognition. *Advances in Neural Networks*, LNCS, Vol. 3971, (pp 41 - 49). Berlin: Springer-Verlag.

Matsuka, T. & Corter, J. E. (2006). Process tracing of attention allocation in category learning. Under review.

Matsuka, T. & Nickerson, J. V. (2006). Modeling human hypotheses-testing behaviors using simulated evolutionary processes. In *Proceedings of IEEE Congress on Evolutionary Computation*. Forthcoming.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 8*, 37-50.

Medin, D. L., & Schaffer, M. M.(1978) Context theory of classification learning. *Psychological Review, 85*, 207-238

Minda, J. P., & Smith, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 275-292.

Nosofsky, R. M. (1986). Attention, similarity and the identification categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

Rosch, E. Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Sakamoto, Y., Matsuka, T. & Love, B.C. (2004). Dimension-wide vs. exemplar-specific attention in category learning and recognition. In *Proceedings of the 6th International Meeting of Cognitive Modelling.* (pp. 261-266). Mahwah, NJ. Erlbaum

Smith, J. D. & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 3-27.

Zaki, S. R., Nosofsky, R. M., Stanton, R. D., & Cohen, A. L. (2002). Prototype and exemplar accounts of category learning and attentional allocation: A reassessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1160-1173.