

# Modeling Human Hypotheses-Testing Behaviors Using Simulated Evolutionary Processes

Toshihiko Matsuka and Jeffery V. Nickerson

**Abstract**—Human category learning has been modeled using exemplar, prototype, and rule-based theories. Rule-based models are the least discussed. This paper presents a rule-based model based on evolutionary computation techniques. Such techniques allow for the combination of concepts, an important aspect of human cognition that has been largely overlooked in previous cognitive modeling research. We also include other human-like characteristic in the model, namely a simplicity bias and instance-based learning. The results suggest that such an algorithm can replicate well-known results in human category learning. We discuss the broader issue of which of the three models of categorization make sense in particular situations.

## I. INTRODUCTION

Human concept formation, or more generally learning, probably consists of multiple components. In some situations, humans learn by incrementally modifying their current knowledge while in other situations, humans learn by conceptual combination [1] [2]. Although the former process has been widely applied and integrated in computational models of high-order human cognition and learning, the latter has been largely overlooked by cognitive modeling research. Specifically, the gradient descent learning method, which can be considered as a process of incremental modification of a concept, has been widely employed in many models of concept formation and resulted in notable successes in replicating many psychological phenomena (e.g. [3]). Perhaps because of this successful explanation of empirical data by learning-by-modification algorithms, other equally theoretically plausible ways of learning such as learning-by-combination have been neglected by the categorization research community. Alternatively, there may be an implicit belief that humans possess and utilize a universal set of concepts, so that combinations would just yield existing members of the set.

However, there is evidence that learning by concept combination happens (e.g. [1] [2] [4]). In order to better understand the nature of human learning, we decided to develop and evaluate a formal model of learning based on concept combination.

We proceed in the following way. We first discuss related work. Then we describe the model in detail. We provide the results of an illustrative simulation, and discuss the implications for future categorization research.

Toshihiko Matsuka is with the Center for Decision Technologies, Howe School of Technology Management, Stevens Institute of Technology, Hoboken, NJ 07030, USA (phone: +1 201-216-8547; fax: +1 201-216-5385; email: tmatsuka@stevens.edu).

Jeffery V. Nickerson is with the Center for Decision Technologies, Howe School of Technology Management, Stevens Institute of Technology, Hoboken, NJ 07030, USA (email: jnickerson@stevens.edu).

## II. MODEL OVERVIEW & BACKGROUND

### A. Category Learning

The ability to categorize plays a central role in high-order human cognition. Categorization allows us to process, understand, and communicate complex thoughts and ideas by efficiently utilizing salient information while ignoring other information: categorization is a form of data compression. In other words, *categories* are the building units of human knowledge or concepts [5]. Therefore in the field of Cognitive Science the terms 'concept formation' and 'category learning' are used interchangeably, and we follow this trend throughout this paper.

### B. Background

Our new model is called HySEP, for Hypotheses-testing learning with Simulated Evolutionary Process. HySEP is based on character comparisons, and its learning algorithm is based on a multi-objective genetic algorithm, which can be characterized as a simulated evolutionary process: concepts compete with each other and better concepts survive.

Evolutionary algorithms have been discussed in relation to categorization in the past. Holland [6] discussed their use of genetic algorithm in classification. Other researchers have compared their effectiveness to human categorization [7] [8]. Their studies suggested that evolutionary algorithms can replicate, at least generally, the results of a classical study on human categorization, namely the Shepard, Hovland, and Jenkins study [9] and its replication [10]. Although the some degree of effectiveness of evolutionary algorithms in cognitive modeling was demonstrated by both Hartley [7] and Sen [8], their studies placed a lesser emphasis on a qualitative interpretation of the algorithm and had neglect some important aspects of human cognition. For example, they focused on batch learning, which is less plausible in human cognition than instance-based learning. In contrast, we are interested here in the issue of plausibility and the descriptive validity of our model.

We model category learning on the basis of rule representation, assuming concepts are organized by sufficient and necessary rules. We choose a rule-like representation, partly because there is increasing interest in rule-based modeling (e.g. [11] [12]), and partly because it would result in a simpler implementation than some of the previous work in the field, including our own (e.g. [13] [14]).

Throughout this paper, we assume that all feature dimensions are binary and there are only two alternatives in all category structures. Although this is a restriction, we

believe the model’s general principles will still hold for more complex category structures, and extending the model will be straightforward.

### C. Relationships between Genetic Algorithm and Human Cognition

A Genetic Algorithm (GA) is a relatively simple yet robust optimization method based on simulated evolutionary processes [15]. In a GA, there are chromosomes that consist of genes (i.e. coefficients) to be optimized. In HySEP, a gene is a particular rule, and a chromosome is a particular concept or a set of rules (referred to as a concept vector). In a GA, there are multiple chromosomes in a hypothetical environment or population and they compete with each other to pass their genes to their descendant. There are some important evolutionary processes in this hypothetical environment, allowing fitter genes (i.e., a set of rules) to survive. This in turn results in having strong genes in the population, which translate to concepts and knowledge optimization in HySEP.

In a typical GA setup, there are three important processes in each evolution phase: Selection, Crossover (i.e., recombination), and Mutation (i.e., stochastic modification). Here we interpret these processes in terms of human cognition.

1) *Mutation*: In a Mutation process, each gene is randomly altered with some probability  $P$ . The Mutation can be considered as a modification of the concept by randomly creating and testing new hypothesis. Virtually all previous models of human category learning incorporate a Mutation process as the sole mechanism of learning.

2) *Crossover*: In a Crossover process, the selected chromosomes form a pair and exchange gene information to create a new pair of chromosomes. In human cognition, the crossover process is one of conceptual combination, creating new sets of concepts by merging two strong concepts chosen by the (parent) selection process.

3) *Selection*: In a (parent) Selection process, usually about a half of the chromosomes are selected on the basis of their fitness in relation to the environment. Those selected create offspring (i.e., new concepts), while non-selected chromosomes or concepts become obsolete and extinct.

The characteristics of the Selection and Crossover processes are distinct from most models of human category learning, because previous models possess and modify a single concept (i.e., a single set of coefficients), whereas HySEP maintains, modifies, and combines multiple concepts. We consider this to be a contribution of HySEP to cognitive modeling.

The idea of having a population of concepts (vs. having a single concept) is important not only because it allows the Selection and Crossover processes in learning, but also because it allows the creation of diverse concepts that have similar accuracies and/or utility, making knowledge more robust. To our knowledge, this capability has not been discussed in the category learning modeling community, and may warrant future research. The utility of having homogeneous versus heterogeneous concepts probably depends on situational factors, and these factors could be varied.

TABLE I

ALL POSSIBLE RULES FOR A TWO-DIMENSIONAL BINARY STIMULUS SET AND AN EXAMPLE CONCEPT VECTOR FOR CATEGORY STRUCTURE:

CATEGORY A IF DIM2 = ■ AND CATEGORY B IF DIM1=○

	R1	R2	R3	R4	R5	R6	R7	R8
Dim1	*	■	*	○	■	○	■	○
Dim2	■	*	○	*	○	■	■	○
Example Concept Subvectors								
$\rho_j^A$	1	0	0	0	0	0	0	0
$\rho_j^B$	0	0	0	1	0	0	0	0

Another important feature of a GA in cognitive modeling is that it allows the hypothetical error surface to be discontinuous. Some empirical studies have suggested that a human’s concept space might have a non-smooth or discontinuous property [16]. This characteristic has not been successfully incorporated in cognitive model using gradient descent optimization methods. HySEP, because of its stochastic optimization technique, can incorporate multi-objective functions in learning that are consistent with the complexity of human learning and the possibly discontinuous nature of the knowledge utility hypersurface.

### III. REPRESENTATION & ENCODING

There are several ways in which concepts or categorization rules can be represented in a chromosome. For example, a complex encoding approach would incorporate both variables (rules) and operations in a concept vector (a chromosome). A minimal approach would just use a vector indicating the presence or absence of rules. We selected the latter approach for its simplicity. In our approach, category knowledge (a concept) is represented by a vector of *Rules* (i.e., a rule is a gene). For an  $N$ -dimensional binary stimulus set, there are  $(3^N - 1)$  possible rules for each category (Note that those rules are not necessarily mutually exclusive). That is, each dimension can be either one of three possible values – one (e.g. ■) or the other (e.g. ○) or not important (represented by ‘\*’) – but no NULL rule (i.e., a rule with all \*) is allowed (see Table 1). In HySEP, a concept vector  $j$  or  $\rho_j$  consists of two subvectors – one subvector for Category A (i.e.,  $\rho_j^A$ ) and another for Category B (i.e.,  $\rho_j^B$ ). Thus,  $\rho_j = [\rho_j^A : \rho_j^B]$ .

A concept vector consists of two elements either “1” (i.e., the corresponding rule is applicable), or “0” (i.e., the corresponding rule is NOT applicable). Each rule can be represented with  $N$  elements. Table 1 shows all possible rules for a two-dimensional binary stimulus set and an example concept vector. For category structure [Category A if Dim2 = ■ and Category B if Dim1=○], after successful learning, HySEP would acquire the following concept vector: [1000000:0001000], where the first element corresponds to Rule 1 (R1 in Table 1) for Category A, the second to R2, and so forth. The acquired concept vector indicates that R1 (i.e., Dim2 = ■) is applicable for classifying Category A and R4 (i.e., Dim1=○) for Category B. Note that in our modeling framework, it is assumed that more complex rules (i.e., rules defined by more features or less wildcards ‘\*’) always appear

later within subvectors.

As shown above, in order to have multiple rules and/or to acquire Rule-plus-exception type knowledge, we need to incorporate a sufficiently long vector (i.e.,  $3^N - 1$ ) to represent knowledge. Although it might appear that this representation approach suggests that human knowledge representation is cumbersome, we do not think this is true. In fact, we assume the opposite. As will be discussed in detail in the following sections, we incorporate the simplicity principle, making simulated humans prefer simple-sufficient-accurate knowledge over complex but marginally-more-accurate knowledge. In particular, we assume that humans' initial mental states are more likely to be described by many zeros in the rule vector(s), and our learning algorithm is more likely to generate and accept simpler rules over complex ones. Note that we assumed that the processing of 0s in rule vectors requires very little, if any, effort. For example, for the abovementioned category structure, HySEP would acquire a rule vector similar to [1000000:0001000], where only two *active* rules need to be processed. Thus, this apparently more-complex-than-necessary representation approach is to accommodate human cognitive capacity, which, with sufficient training (and motivation), could acquire very complex concepts. In future research we may be able to simplify the representation by incorporating a rule operator in the concept vector.

#### IV. CATEGORIZATION & DECODING

In HySEP, it is assumed that humans would first apply the most complex rules (i.e., a rule defined by more features), followed by simpler rules. That is, HySEP starts comparing an input stimulus with an active rule from the end of the rule vector. For example, for a three dimensional stimulus set, HySEP would apply any active 3D-rules (i.e., "exceptions" or exemplars) first followed by 2D-rules, then 1D-rules. This type of rule (or exception) applying behavior has been empirically suggested [17]. If the input stimulus matches an active rule, then HySEP searches and determines if there is another rule with the same complexity level that is applicable for the current input. If all applicable active rules with the same dimensionality suggest a consistent categorization, HySEP deterministically categorizes according to the rules. In the case of inconsistent applicable active rules, HySEP categorizes probabilistically.

The following equation describes output for category A ( $O_A$ ) for input stimulus  $x$  and  $j$ -th concept vector ( $\rho_j$ ):

$$O_A(x, \rho_j) = \frac{\sum_{\forall i \in H} I(\rho_{ji}^A = 1)}{\sum_{\forall i \in H} I(\rho_{ji}^A = 1) + \sum_{\forall i \in H} I(\rho_{ji}^B = 1)} \quad (1)$$

where indicator function  $I(expression)$  returns 1 if *expression* is satisfied, or 0, otherwise.  $H$  indicates the active rule(s) with the highest dimensionality that is applicable to input  $x$ . The output for category B is obtained in the same manner.

##### A. Processing a rule vector

In HySEP, there are always multiple rule vectors in its conceptual space because of the nature of its evolutionary

learning algorithm. These concept vectors are processed individually when their concept utilities are calculated for the Selection process in learning. However, in order to determine a categorization response, we may need to restrict HySEP to a single response behavior, because people in general can perform only one set of actions at a given instant. That is, when humans are asked to categorize an object, they answer with a single category in a particular level of the category hierarchy. In other words, concepts-to-action is most likely a many-to-one mapping scheme. This in turn raises a theoretical question: how and what concept vectors would humans choose to use in order to categorize stimuli? The following is a list of possible options:

**Option A:** Select the best fit concept vector

**Option B:** Select a concept vector randomly

**Option C:** Create and use the average concept vector

In HySEP, we assume that people would create and utilize the (dichotomized) average concepts for categorizing objects (i.e., Option C) because this approach is less sensitive to mutation. Our preliminary modeling studies showed it was less likely to exhibit "all correct" responses when we incorporated Options A or B because of continuing concept mutation. This, however, does not imply that we are trying to reduce the effect of the Mutation process in learning (it is still very important in learning). Rather, by incorporating the Option C method, we reduce its effect in categorization behaviors. That is, HySEP's asymptotic performance in categorization is most similar to human performance when we incorporate Option C in the preliminary studies. Therefore, we apply the following method to create the average concept vector:

$$\bar{\rho}_i^A = \begin{cases} 1, & \text{if } \frac{1}{N} \sum_j I(\rho_{ji}^A = 1) > 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where  $\bar{\rho}^A$  is the average concept subvector for Category A, and  $N$  is the number of concept vectors in a mind set. The average concept subvector for Category B is calculated in the same manner, and  $\bar{\rho} = [\bar{\rho}^A : \bar{\rho}^B]$

#### V. LEARNING ALGORITHM: MULTIOBJECTIVE GA

HySEP assumes that learning is driven by an optimization of the subjectively and contextually defined utility of knowledge being acquired, rather than by a simple classification error minimization routine [18]. Thus, the learning objective can be represented by the following general utility function  $U$ :

$$U(\rho_j) = E(\rho_j) + \sum_m^M \lambda_m Q_m(\rho_j) \quad (3)$$

where  $E$  defines the accuracy of concept or knowledge, each  $Q$  in the second term characterizes contextual factors, and  $\lambda$ s are scalars weighting different factors. There are many functions or set of functions appropriately defined for describing a variety of contextual factors including motivation. This utility function is used as the basis for defining fitness for the concept vectors.

Recent multiobjective GA applications often search for a Pareto-optimal set rather than a solution based on arbitrary

weights [19]. However, because of the heuristically operating nature of human high-order cognitive processes [20], we doubt that ordinarily humans in ordinary situations would search for a Pareto-optimal set of concepts. Rather, we believe that humans would utilize a heuristic, subjective, and context-dependent weighting scheme to find a set of concepts. Thus, we incorporate the latter simpler and "arbitrary" approach because of its resemblance to human cognition [20].

#### A. Estimating Accuracy of Concept

We believe learning occurs in an instance-by-instance basis in humans (e.g. [21] [22] [23]). Thus for each training instance, HySEP needs to generate and test a set of concepts. However, if the accuracies of concepts are estimated on the basis of the current training instance alone, then HySEP may over-generalize, even though genetic algorithm can be considered to have a memory effect (ideas persist in the gene pools). There are several ways we can make HySEP acquire correct concepts with instance-based learning. In the present paper, we apply a modified version of Anderson and Schooler's [24] (also see also [25]) memory retention model proposed by Matsuka & Chouchourelou [22] [23]. By incorporating their model, the concept accuracy function  $E$  can be formulated as:

$$E(\rho_j) = \sum_{g=1}^G \left[ \Xi \left( x^{(g)} \right) \left[ d_c^{(g)} - O_c \left( x^{(g)}, \rho_j \right) \right]^2 \right] \quad (4)$$

where  $c$  indicates the *correct* category,  $g$  indicates particular training exemplars,  $G$  is the number of unique exemplars in the training set,  $d$  is the desired output, and  $\Xi$  is (training) exemplar retention function, defining the strength of retaining training exemplar  $x^{(g)}$ . The exemplar retention function is given as:

$$\Xi \left( x^{(g)} \right) = \frac{\sum_{\forall i | x^{(i)} = x^{(g)}} (\tau^{(i)} + 1)^{-D}}{\sum_g \sum_{\forall i | x^{(i)} = x^{(g)}} (\tau^{(i)} + 1)^{-D}} \quad (5)$$

where  $D$  is a memory decay parameter controlling for the speed of memory decay, and  $\tau$  indicates how many instances were presented since  $x^{(g)}$  appeared, with the current training being represented with '0'. Thus,  $\tau = 1$  indicates  $x^{(g)}$  appeared one instance before the current trial. The denominator in the exemplar retaining function normalizes the retention strengths, and thus it controls the relative effect of the training exemplar,  $x^{(g)}$ , in evaluating the accuracy of the concept.

As in the original Anderson and Schooler [24] [25] memory retention model, this modified retention model simultaneously accounts for Power Law of Forgetting [26] and the Power Law of Learning [27]. Given the Power Law of Forgetting and the Power Law of Learning,  $E(\theta)$  is strongly influenced by training exemplars shown more recently in early training trials, but it more evenly accounts for various exemplars in later training trials.

## VI. INCORPORATING HUMAN-LIKE CHARACTERISTICS

### A. Simplicity Bias

Recently, a number of cognitive scientists suggested the importance of incorporating a principle of simplicity in high-order human cognition from both theoretical and empirical points of views (e.g. [14] [28] [29]). Therefore, we incorporate the simplicity principle in HySEP. However, for our framework, there are at least three ways to do so: by manipulating initial conceptual complexity, by manipulating the probabilities of mutation for rules with different complexity, and by incorporating a utility function accounting for the simplicity bias.

1) *By Initialization:* Intuitively it is difficult to believe that humans initially possess many higher dimensional rules for unknown categories. Furthermore, Johansen and Palmeri [30] recently observed that subjects applied simple rules for categorization tasks in early learning trials and then their strategies changed and they applied more complex rules (e.g. exemplars). HySEP, therefore, incorporates the initial bias toward simpler rules by manipulating the probabilities of particular rules to be initially activated depending on the dimensionality of rules. In particular, we only allow HySEP to have a one-dimensional rule before learning starts.

2) *By Mutation:* The aforementioned empirical study [30] showed the more complex rules (e.g. exemplars) emerge at later stages of learning. This might have been caused by initial bias. Alternatively it might have caused by differential emergence rates (i.e., probabilities of mutations) for rules with different dimensionality. That is, in early stages of learning, humans would more extensively develop and test hypotheses based on simpler or lower dimensional rules than on complex rules. Although this interpretation is plausible, there is some empirical evidence against it. For example, Sakamoto and Love [17] observed that "exceptions", which are objects that do not share a categorization rule with many other objects within the same category, are memorized with higher accuracy than other objects. (Note that an "exception" is often a rule of its own defined by the highest dimensional rule.) That is, instead of creating and testing incrementally more complex rules, humans can create complex rules in a short amount of time, depending on the structure of categories or concepts. Because of our uncertainty about the cognitive mechanisms associated with a differential mutation process, we have not included such a process in our model.

3) *By Utility Function:* We assume that as the dimensionality of a rule increases then the complexity of the rule increases geometrically. Therefore the complexity of a concept vector ( $\Pi$ ) can be formulated as:

$$\Pi(\rho_j) = \sum_i I(\rho_{ji} \neq 0) \cdot \gamma^{\delta_i} \quad (6)$$

where  $\delta_i$  is the *dimensionality* of rule  $i$  in rule vector  $j$ , and  $\gamma$  is a constant that controls the speed of complexity increment. Thus, if the utility of a concept is defined by its accuracy and simplicity, then

$$U(\rho_j) = E(\rho_j) + \lambda \Pi(\rho_j) \quad (7)$$

Note that HySEP is framed as minimization problem, thus a smaller value in Eq. 7 indicates better utility.

## VII. SIMULATIONS

In order to test the descriptive validity of HySEP, simulation studies were conducted. In particular, we simulated a classical study of categorization [9] which is often used as a benchmarking stimulus set [10]. The stimulus structures are shown in Table I. There were a total of 8 training instances defined by 3 binary feature dimensions (i.e., shape, color, and size). Human subjects were trained to learn to classify those instances into the correct categories with corrective feedback. Shepard et al. [9] created 6 category structures of varying complexity of rules for correct categorization. The results of previous empirical studies showed [9] [10] that Type 1 (T1) was the easiest to learn to classify, followed by T2, T3, T4, T5, and T6 being the most difficult, where the differences in difficulty for T3, T4, and T5 were not statistically significant.

Table III shows all 26 possible rules for the stimulus set used in the present simulation study, along with example rules that were acquired by HySEP.

T1 was easiest to learn, probably because it only requires a single one-dimension rule for each category. T2 can be considered as XOR-logic in Dimensions 1 and 2. T3 – T5 are one-dimensional rules with two exceptions, one for each category. T6 was the most complex as it requires memorization of many if not all exemplars.

### A. Methods

Nosofsky et al. [10] collected data on learning curves for those six category structures, and we use their data for the present study.

The basic training procedures follow that of the original study [10]. HySEP was run in a simulated training procedure with 16 trial blocks, where each block consisted of a random presentation of the eight unique training exemplars shown in Table II exactly twice, in order to learn the correct classification responses for the stimulus set. There are a total of 500 simulated subjects for each category structures. HySEP configurations and parameters were fixed the same for all six conditions. The following describes HySEP configuration and parameters.

1) *Population size*: The size of concept population was fixed at 10 throughout the learning process.

2) *Selection*: We use a tournament method with tournament size = 2 (without replacement). Eq. 7 defines the utility (fitness) of each concept vector, and was used for determining winners of the tournaments.

3) *Crossover*: The uniform crossover was employed with the crossover probability for each rule was fixed at 0.50.

4) *Mutation*: The mutation probability for all rules was fixed at 0.01.

5) *Other parameters*: The memory decay parameter,  $D$  was set at 0.3, the complexity parameter  $\gamma$  was set at 1.05, and  $\lambda$ , which controls the relative importance of concept complexity, was set at 0.5. All model parameters were selected arbitrarily.

TABLE II  
SCHEMATIC REPRESENTATIONS OF THE STIMULUS STRUCTURES USED  
IN THE SIMULATION STUDY

Stim.			Categories					
D1	D2	D3	T1	T2	T3	T4	T5	T6
■	■	■	A	B	A	A	A	A
■	■	○	A	B	A	A	A	B
■	○	■	A	A	A	A	A	B
■	○	○	A	A	B	B	B	A
○	■	■	B	A	B	A	B	B
○	■	○	B	A	A	B	B	A
○	○	■	B	B	B	B	B	A
○	○	○	B	B	B	B	A	B

TABLE III  
ALL RULES FOR THE THREE-DIMENSIONAL BINARY STIMULUS SET  
AND EXAMPLE RULES FOR EACH STIMULUS TYPE

	D1	D2	D3	T1	T2	T3	T4	T5	T6
R1	■	*	*	A	0	A	A	A	A
R2	*	■	*	0	0	0	0	0	0
R3	*	*	■	0	0	0	0	0	0
R4	○	*	*	B	0	B	B	B	0
R5	*	○	*	0	0	0	0	0	0
R6	*	*	○	0	0	0	0	0	B
R7	■	■	*	0	A	0	0	0	0
R8	■	*	■	0	0	0	0	0	0
R9	*	■	■	0	0	0	0	0	0
R10	○	○	*	0	A	0	0	0	0
R11	○	*	○	0	0	0	0	0	0
R12	*	○	○	0	0	0	0	0	0
R13	■	○	*	0	B	0	0	0	0
R14	○	■	*	0	B	0	0	0	0
R15	■	*	○	0	0	0	0	0	0
R16	○	*	■	0	0	0	0	0	0
R17	*	■	○	0	0	0	0	0	0
R18	*	○	■	0	0	0	0	0	0
R19	■	■	■	0	0	0	0	0	0
R20	■	■	○	0	0	0	0	0	B
R21	■	○	■	0	0	0	0	0	B
R22	■	○	○	0	0	B	B	B	0
R23	○	■	■	0	0	0	A	0	0
R24	○	■	○	0	0	A	0	0	A
R25	○	○	■	0	0	0	0	0	A
R26	○	○	○	0	0	0	0	A	0

### B. Results and Discussion

Figure 1 shows the results of the present simulation study. HySEP replicated the observed order of difficulties successfully. It slightly under-predicted the difficulty of T4 category structure at the end. However the figure shows that HySEP could correctly predict the order of difficulty at the 4th training block. Thus its prediction may improve if the model parameters were more carefully selected.

The results of the original study by Shepard et al. [9] and its replication study by Nosofsky et al. [10] have long been interpreted as showing that category learners can selectively allocate attention to stimulus features on a dimension-by-dimension basis, and that they can learn to allocate attention in an optimal or near-“optimal” manner across stimulus dimensions. However, the results of the present simulation study cast doubt on this interpretation, because HySEP without any explicit selective attention mechanisms could replicate the observed phenomena. Our interpretation of the

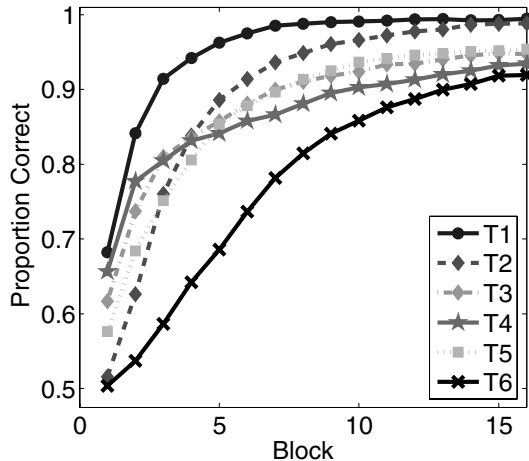


Fig. 1. Results of the Simulation Study.

present simulation study is that a preference for simple-yet-sufficiently-accurate concepts was the key mechanism for the phenomena, resulting in HySEP acquiring simpler concepts (e.g., T1 & T2) faster than more complex ones (e.g., T3 – T6). However, we acknowledge that interdependencies between simplicity bias and selective attention exist. Thus, it might have been that the selective attention process operating via the simplicity bias led HySEP to successfully replicate the phenomena.

### VIII. DISCUSSION

Previous research has show that exemplar theory provides a plausible theory of human categorization. While prototype theory is currently out of favor, we have recently demonstrated that, augmented with a local attention mechanism, prototype theory can replicate human subject experiments [13] [14]. We have shown here that a rule-based approach will also replicate human subject experiments.

Assuming all three theories are plausible, what does this say about human categorization? We can think of several alternatives. First, one of the theories may correctly characterize the way humans categorize, and the others don't. Or it could be that all researchers, including us, are overfitting theories to data. If so, it is even possible that none of the current theories match human mechanisms.

However, it is also possible that all the theories are in some way equivalent. Perhaps we can transform one to the other in simple but not currently obvious ways. If this is the case, then nature may have just picked one. Or nature may use all three: our minds may use alternative strategies depending on the situation.

This idea is worthy of consideration, for in the field of human decision making, there is a strong body of evidence that humans use alternative decision strategies [31]. Sometimes people seem to optimize their self interest. We might see this as computational strategy akin to exemplar or prototype theory: humans calculate the distance to exemplar

or prototypical outcomes, and then decide what to do. But other times, people make decisions based on rules.

If our lower level categorization processes are at all analogous to our higher level decision processes, then we might use alternate categorization techniques. The question becomes: do we use these in parallel? Or do pick our categorization technique to fit the context?

Our hypothesis is that there might be a hierarchy of (subjective) optimization problems. One level, for example, is for choosing an "optimal" representation mode (e.g. exemplar, prototype, or rule representation) while another level, given the selected representation method, optimizes knowledge or coefficients (for example, finding a good set of rules or adjusting association and attention weights), while still other levels optimize additional important resources (e.g. memory).

Alternatively, there may be a single level of optimization processes where various types of operations and concepts are simultaneously optimized.

#### A. Extension

There are several ways in which HySEP can be extended. As stated above, it could be built with multiple levels of optimization processes in order to acquire robust and situationally suitable concepts, possessing the ability to adapt its representation mode, memory usage, and association weighting strategy.

Another potentially significant extension is to integrate self-adaptive learning strategies into HySEP. Several parameters in HySEP's learning method are static, but in order to make it a more realistic cognitive model, a greater degree of self-adaptability may be needed. For example, the population size and the rates of mutation and crossover can be dynamically self-adjusted depending on the success of the learning process, so that more dedicated or motivated learning behaviors are exhibited only when there is a real need for knowledge.

### IX. CONCLUSIONS

The present paper introduced a new model of human category learning based on simulated evolutionary processes. A notable contribution of our model, called HySEP, includes (a) a capability of modeling human learning with both the modification and combination of concepts, (b) a capability of optimizing concept utility on the instance-by-instance basis (vs. batch learning), (c) being sensitive to subjective and contextual factors in multiobjective learning, and (d) a capability of incorporating a discontinuous knowledge utility hypersurface. A simulation study was conducted and showed that HySEP, without any explicit selective attention mechanisms, reliably replicated a classic empirically study on human category learning whose results have long been interpreted as evidence for selective attention processes in human cognition [9] [10].

#### ACKNOWLEDGMENT

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

## REFERENCES

- [1] J. A. Hampton, "Inheritance of attributes in natural concept conjunctions," *Memory & Cognition*, vol. 15, pp. 55-71, 1987.
- [2] J. A. Hampton, "Conceptual combination," in *Knowledge Concepts, and Categories* (Chapter 4), Edited by K. Lamberts, and D. Shanks, East Sussex: Psychology Press, 1997.
- [3] J. E. Kruschke, "ALCOVE: An Exemplar-Based Connectionist Model of Category Learning," *Psychological Review* vol. 99, pp. 22-44, 1992.
- [4] G. Fauconnier, and M. Turner, "Conceptual Integration Networks," *Cognitive Science*, vol. 22, pp. 133-187, 1998.
- [5] K. O. Solomon, D. L. Medin, and E. Lynch, "Concepts Do More Than Categorize", *Trends in Cognitive Sciences*, vol. 3, pp. 99-105, 1999.
- [6] J. H. Holland, "A mathematical framework for studying learning in classifier systems", *Physica D*, vol. 2, pp. 307-317, 1986.
- [7] A. Hartley, "Accuracy-based fitness allows similar performance to humans in static and dynamic classification environments,". *Proc. the Genetic and Evolutionary Computation Conference*, pp. 266-273, 1999.
- [8] S. Sen, "Modeling human categorization by a simple classifier system,". *Online Workshop on Soft Computing*, 1996, Aug. 19-30.
- [9] R. N. Shepard, C. L. Hovland, and H. M. Jenkins, "Learning and memorization of classifications," *Psychological Monographs* vol. 75, 1961
- [10] R. M. Nosofsky, M. A. Gluck, T. J. Palmeri, S. C. McKinley, and P. Glauthier, "Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961)". *Memory and Cognition*, vol. 22 pp. 352-369, 1994.
- [11] R. M. Nosofsky, T. J. Palmeri, and S. C. McKinley "Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961)". *Memory and Cognition*, vol. 22 pp. 352-369, 1994.
- [12] D. J. Navarro, "Analyzing the RULEX model of category learning". *Journal of Mathematical Psychology*, vol. 49 pp. 259-275, 2005.
- [13] T. Matsuka, "A Model of Category Learning with Attention Augmented Simplistic Prototype Representation. *Advances in Neural Networks, LNCS*, vol. 3971, pp. 34-40, Berlin: Springer-Verlag, 2006.
- [14] T. Matsuka, J. V. Nickerson, and J. Jian, "A prototype model that learns and generalize Medin, Alton, Edelson, & Frecko (1982) XOR category structure as humans do," In *Proc. Annual Meeting of Cognitive Science 2006*. Forthcoming.
- [15] D. E. Goldberg, *Genetic Algorithm in Search, Optimization, and Machine Learning*, Reading, MA: Addison-Wesley, 1989.
- [16] G. H. Bower and T. R. Trabasso, "Reversals Prior to Solution in Concept Identification," *Journal of Experimental Psychology* vol. 66, pp. 409-418, 1963.
- [17] Y. Sakamoto, and B. C. Love, "Schematic Influences on Category Learning and Recognition Memory,". *Journal of Experimental Psychology: General*, vol. 133, pp. 534-553, 2004.
- [18] T. Matsuka, T. "Modeling human learning as context dependent knowledge utility optimization," in *Advances in Natural Computation, LNCS* vol. 3601 (pp. 933-946) Edited by L. Wang, K. Chen, and Y. S. Ong, Berlin: Springer-Verlag. 2005. .
- [19] K. Deb *Multi-objective optimization using evolutionary algorithms*, Chichester, England: Wiley, 2001.
- [20] G. Gigerenzer, P. Todd, and the ABC Research Group, *Simple heuristics that make us smart*, New York: Oxford, 1999.
- [21] C. Gonzalez, J. F. Lerch, and C. Lebiere, "Instance-based learning in dynamic decision making," *Cognitive Science*, vol. 27, pp. 591-635, 2003.
- [22] T. Matsuka, and A. Chouchourelou, A. "Model of category learning with dynamic multi-objective multi-hypotheses testing," presented at the 35th Annual Meeting of the Society for Computer in Psychology. Toronto, Canada, November 10, 2005.
- [23] T. Matsuka and A. Chouchourelou, "A Model of Human Category Learning with Dynamic Multi-Objective Hypotheses Testing with Retrospective Verifications". In *Proc of IJCNN 2006*, Forthcoming.
- [24] J. R. Anderson, and L. J. Schooler. "Reflections of the environment in memory," *Psychological Science*, vol. 2, pp. 396-408, 1991.
- [25] J. R. Anderson, D. Bothell, C. Lebiere, and M. Matessa, M, "An integrated theory of list memory," *Journal of Memory and Language*, vol. 38, pp. 341-380, 1998.
- [26] D. C. Rubin, and A. E. Wenzel, "One hundred years of forgetting: a quantitative description of retention," *Psychological Review*, vol. 103, pp. 734-760, 1996.
- [27] A. Newell, and P. S. Rosenbloom, "Mechanism of skill acquisition and the law of practice," in *Cognitive skill and the their acquisition*, Edited by J. R. Anderson (Ed.), Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.
- [28] J. Feldman, "The simplicity principle in human concept learning," *Current Directions in Psychological Science*, vol. 12, pp. 227-232, 2003.
- [29] T. Matsuka, "Simple, Individually Unique, and Context-Dependent Learning Method for Models of Human Category Learning," *Behavior Research Methods* vol. 37, pp. 240-255, 2005.
- [30] M. K. Johansen, and T. J. Palmeri, "Are there representational shifts during category learning?" *Cognitive Psychology*, vol. 45, pp. 482-553, 2002.
- [31] J. March, *A Primer on Decision Making*, New York: Free Press, 1994.