A Computational Account of Basic Level and Typicality Effects

Douglas H. Fisher
Department of Computer Science
Box 67, Station B
Vanderbilt University
Nashville, TN 37235

Abstract

Cognitive psychology has uncovered two effects that have altered traditional views of human classification. Basic level effects suggest that humans prefer concepts at a particular level of generality, while typicality effects indicate that some instances of a class are more readily recognized as such than others. This paper describes a model of memory that accounts for basic level effects, typicality effects, and interactions between them. More generally, computer experiments lay the groundwork for a formal unification of basic level and typicality phenomena.

1. Introduction

A significant finding in cognitive psychology is that within hierarchical classification schemes there is a *basic* or preferred level of human classification. In a *forced naming task* [Rosc76, Joli84], a subject identifies a pictured item; when shown a picture of a particular collie, subjects will respond that it is a *dog*, not a *collie*, *mammal*, or *animal*. In a *target recognition task* [Rosc76], a subject will more quickly confirm that a pictured collie is a *dog* than confirmation will be given for *collie*, *mammal*, or *animal*. These two tasks indicate that for a hierarchy containing (*collie*, *dog*, *mammal*, *animal*), *dog* is the basic level concept.

A second influential class of phenomena are *typicality* effects. Psychological studies indicate that some members of a class are treated preferentially or as more typical of a class. For example, in a target recognition task a *robin* will be recognized as a *bird* more quickly than will a *chicken*. The evidence for a typicality ranking has accrued from many sources [Merv81, Smit81, Rosc78].

This paper describes a cognitive model of hierarchical classification and memory that accounts for basic level and typicality effects during target recognition and forced naming tasks. Apparently this is the first computational model of any basic level effect. In addition, typicality effects emerge from the same classification procedures. Interactions between basic level and typicality phenomena are also demonstrated in computer experiments. These findings further confirm the model's psychological consistency and suggest unexplored behavioral possibilities.

2. Concept Models and Hierarchies

Many psychological and AI studies have assumed that concepts are logical summaries of features that are common to concept members; instances are classified by insuring a 'perfect' match between a concept and the instance. This *classical* view [Smit81] implicitly treats each instance as 'equal', but typicality effects suggest that humans do not treat instances equally. In response, a number of representations have been proposed that reflect the variable importance of concept features [Rosc75]. In particular, *probabilistic* representations [Smit81] associate a probability or weight with each concept feature.

Recognition using a probabilistic concept involves summing the weights of concept properties that are present in an instance. *Independent cue models* only record probabilities of individual properties (e.g., $P(\text{COLOR} = \text{red})$); the time required for summation to reach a predefined threshold varies with object typicality, thus accounting for target recognition data. However, independent cue models are limited; summing over primitive property weights constrains recognition to linearly separable classes. This has motivated *relational cue models* that record probabilities for property combinations (e.g., $P(\text{COLOR} = \text{red} \wedge \text{SIZE} = \text{large})$) and *exemplar* models [Smit81, Kibl87] that do not represent concepts by abstractions (probabilistic or logical), but by selected instances. Exemplar models are equivalent to relational cue models since instances can be used to compute joint-property distributions as needed.

Computational considerations motivate two models of *individual* concepts beyond the independent cue type: relational cue and exemplar models. However, another view [Fis87a] is that the weaknesses of independent cue models can be overcome by concept *organizations*. This view is illustrated by a conceptual clustering system, COBWEB [Fis87a, Fis87b], which builds *probabilistic concept trees*. Each node of the tree, N_k, contains conditional probabilities for observed attribute values, $A_i = V_{ij}$. For example, Figure 1 shows a tree over U.S. senators, where each senator is described by a legislative voting record – i.e., 14 attribute values (e.g., Contra-aid=yes). Only a few probabilities, $P(A_i = V_{ij}|N_k)$, are shown, but probabilities conditioned on node membership are stored for

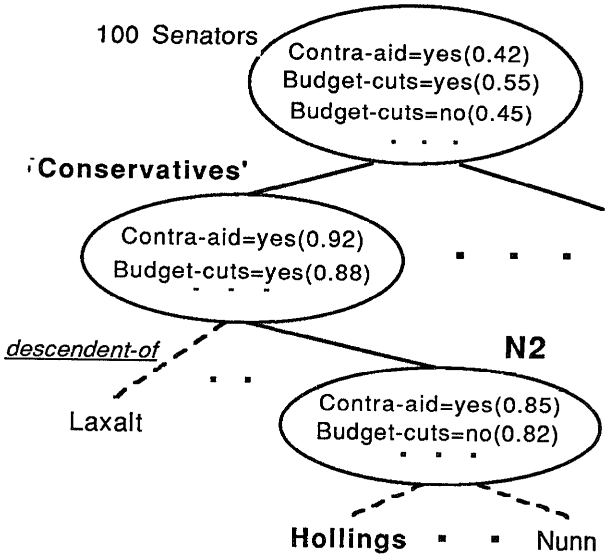

Figure 1: Probabilistic tree of senate voting records

each attribute value. Values along certain attributes (e.g., Budget-cuts) tend to distinguish members of a subclass (N2) from a higher-level node ('conservatives'), although there must be agreement on other attributes (e.g., Contra-aid) for objects to have been grouped together under 'conservatives'. Object classification proceeds along a path of best matching nodes – i.e., those that maximize a summation of individual attribute value probabilities.

Probabilistic concepts can be organized so as to *optimize* prediction of a *single* 'teacher-specified' (perhaps nonlinear) class attribute as with ID3 [Quin86] decision trees. However, COBWEB does not regard any attribute as special – rather, COBWEB trees facilitate *good* prediction along *many* attributes. In theory, probabilistic concept trees capture the the same information found in relational cue or exemplar models. Ideally, the tree should capture joint probabilities that occur most frequently or are most 'important', thus improving classification accuracy and/or efficiency. Probabilistic concept trees are best viewed as efficient implementations of exemplar and relational cue models, rather than as alternatives to them. This view opens the way for a unified account of basic level and typicality effects – behaviors traditionally treated as disparate because of distinctions drawn between representations of individual concepts (i.e., the scope of typicality) and concept hierarchies (i.e., the scope of basic level effects).

3. Hierarchical Classification

Basic level effects suggest that there is a preferred level in human classification hierarchies. Several measures [Rosc78, Jone83] for predicting the basic level have been proposed. Most recently, Gluck and Corter [Gluc85] have formulated *category utility*, which presumes that the basic level maximizes 'predictive ability'. For example, very few correct predictions can be made about an arbitrary *animal*, but those that can be made (e.g., animate) apply to a large number of objects. In contrast, knowing something is a *robin* assures many predictions, but they apply to a small number of objects. The basic level concept (e.g., *bird*) is where a tradeoff between the *expected* number of correct predictions (e.g., has-feathers, beaks, flies) and the proportion of the environment to which the predictions apply,

$$P(N_k)E(\#\text{ correct predictions}|N_k),$$

is maximized. If $P(A_i = V_{ij}|N_k)$ is the probability that an attribute value will be predicted and this prediction is correct with the same probability then this measure can be further formalized as:

$$P(N_k)\sum_i \sum_j P(A_i = V_{ij}|N_k)^2. \qquad 3\text{–}1$$

Gluck and Corter verify that category utility correctly predicts the basic level (as behaviorally identified by human subjects) in two experimental studies [Hoff83, Murp82].[1]

COBWEB uses category utility as a partial matching function to incrementally guide object incorporation at each successive level of a classification tree. Despite its psychologically-motivated underpinnings, COBWEB's strict top-down classification procedure does not jibe with findings that an intermediate or basic 'entry point' is preferred by humans. To account for basic level effects the classification scheme used by COBWEB is modified along two dimensions of classification [Rosc78]. The *horizontal* dimension is concerned with object placement among contrasting categories at the same tree level. The *vertical* dimension is concerned with object placement among categories at various levels of generality.

3.1. The Horizontal Dimension

A number of systems [Lebo82, Kolo83] use attribute-value indices to constrain classification along the horizontal dimension. Indices filter out nodes that bear little similarity to a new object. A similar approach can be developed in the COBWEB framework. In particular, the best host for a new object is the category that maximizes

$$P(N_k)\sum_i P(A_i = V_{ij}|N_k)^2. \qquad 3\text{–}2$$

This measure favors the class whose attribute-value distributions are most reinforced by the new object. This function is *not* guaranteed to identify the same best host as 3–1, but empirical analyses indicate that there is very close agreement [Fis87a]. Using Bayes rule, 3–2 equals:

$$\sum_i P(A_i = V_{ij})P(N_k|A_i = V_{ij})P(A_i = V_{ij}|N_k). \qquad 3\text{–}3$$

[1] Category utility is actually the expected *increase* in correct predictions. However, for our purposes, 3–1 is equivalent.

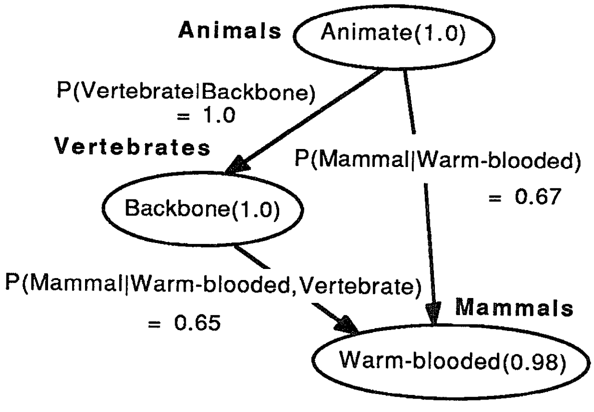

Figure 2: Index placement along the 'vertical' dimension

Intuitively, $P(N_k|A_i = V_{ij})$ is the *predictiveness* of V_{ij} towards class N_k, while $P(A_i = V_{ij}|N_k)$ is the *predictability* of V_{ij} among N_k members. $P(A_i = V_{ij})$ is V_{ij}'s predictability at the root of a classification tree. $P(A_i = V_{ij})$ resides at the root, $P(A_i = V_{ij}|N_k)$ at N_k, and $P(N_k|A_i = V_{ij})$ weights the index between the root and N_k.

3.2. The Vertical Dimension

Classification efficiency can also benefit from indices that 'jump' levels of a hierarchy. In particular, attribute value indices are directed at nodes that maximize the *collocation* [Jone83] of V_{ij}: $P(N_k|A_i = V_{ij})P(A_i = V_{ij}|N_k)$. This is a tradeoff between the predictiveness and predictability of V_{ij} with respect to N_k. Collocation-maximizing nodes tend to be the most specific nodes for which a value is still significantly predictive.

Figure 2 illustrates how indices may skip levels. The collocation for *Backbone* is maximized at *Vertebrates*: $P(\text{Vertebrate}|\text{Backbone}) \times P(\text{Backbone}|\text{Vertebrate}) = 1.0 \times 1.0 = 1.0$. However, the index for *warm-blooded* is directed at the subordinate node *Mammals* since collocation is maximized there: $0.67 \times 0.98 = 0.66$. Moreover, each node (e.g., Vertebrates) is the root of its own subtree; indices whose probabilities are conditioned on subnode membership are also defined [Fis87a]. However, in this paper there is little need to detail the recursive case.

After adding variance on the vertical dimension, recognizing an object $O = \{A_1 = V_{1j_1}, A_2 = V_{2j_2}, ..., A_m = V_{mj_m}\}$ is a matter of finding the node that maximizes

$$\textit{total-predictiveness} = \textstyle\sum_i P(N_k|A_i = V_{ij}). \qquad 3\text{–}4$$

Intuitively, this is the node that is most predicted by the object's attribute values. Because indices are directed at collocation maximizing nodes, $P(A_i = V_{ij}|N_k)$ tends to be high (e.g., close to 1.0). In addition, $P(A_i = V_{ij})$ is constant across all nodes. For these reasons, $P(A_i = V_{ij})$ and $P(A_i = V_{ij}|N_k)$ have little impact on best host selection; in practice 3–4 closely approximates 3–3. Classification of an object, O, is summarized by:

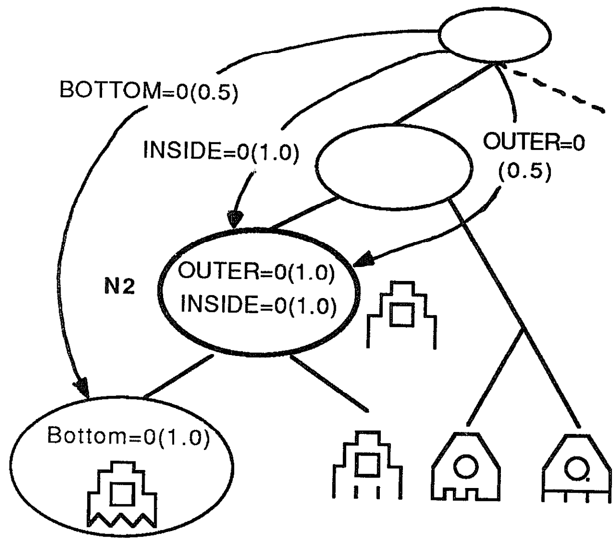

Figure 3: Partially-indexed tree to test basic level effects

```
FUNCTION Classify (O, Root [of tree])
   total-predictiveness(N_k) ← 0 for each N_k
   FOR each V_ij (∈ O)
      FOR each V_ij index from Root (to N_k)
         increment total-predictiveness(N_k)
               by P(N_k|A_i = V_ij,Root)
   Best ← N with max[total-predictiveness(N_k)]
   IF terminate-condition THEN RETURN(Best)
      ELSE Classify (O, Best)
```

Recursion may terminate when a leaf is reached or when a desirable prediction can be made with certainty.

4. An Account of Basic Level Effects

The indexing scheme's consistency has been demonstrated with findings from two psychological studies of basic level effects. In one study [Hoff83] subjects learned a classification tree over 'nonsense' objects like the one shown in Figure 3. Each class (node) had a 'nonsense' name that subjects used to identify class membership in target recognition tasks. Objects were defined in terms of three attributes: the shape of the *inside* subcomponent with values square, triangle, star, or circle (encoded as 0, 1, 2, and 3, respectively); the *outer* shape with (encoded) values of 0 and 1; and the shape of the *bottom* with values 0, 1, 2, and 3. For the tree of Figure 3 subjects consistently 'preferred' level 2 (e.g., N_2); the root is level 0. In addition to this tree, two separate subject groups were trained and tested on trees with different basic levels.

The trees were encoded as probabilistic concept trees as shown for the leftmost portion Figure 3's tree. Value probabilities are only shown at nodes where the value's collocation is maximized.[2] The object {OUTER = 0, BOTTOM = 0, INSIDE = 0} is first recognized with respect N_2 since this node maximizes 3–4: $P(N_2|\text{OUTER}=0) + P(N_2|\text{INSIDE}=0) = 0.5 + 1.0 = 1.5$.

For each of the three variants of the Hoffman and Ziessler study, the model identified objects with respect to the appropriate basic level node. The model is also consistent with experiments by Murphy and Smith [Murp82].

5. An Account of Typicality Effects

Humans also exhibit preferences along the horizontal dimension as evidenced by typicality studies. In this regard, the indexing scheme is consistent with findings by Rosch and Mervis [Rosc75]; they demonstrate that typicality increases with the number of features shared with other objects of the same class and varies inversely with the number of features shared with members of contrasting classes.

5.1. Typicality and Intra-Class Similarity

Rosch and Mervis used the 'nonsense' strings of Table 1a to demonstrate the relation between typicality and intra- (within-) class similarity. Members of category A vary in the extent that they overlap with other members of the same class. For example, on *average* the symbols of 'QBLFS' appear in 2 other strings of class A, while the symbols of 'HMQBL' are shared by 3.2 other members of class A. The *inter-class* overlap between members of A and B is constant (i.e., no overlap). Subjects learned to distinguish categories A and B and then participated in target recognition tasks for members of A. Recognition time decreased as within-class overlap increased, supporting the hypothesis that typical instances shared more properties with other members of the same class.

To model these effects some presumptions must be made about how categories A and B can be hierarchically structured. This did not pose a problem in modeling basic level effects, since classification trees were explicitly taught. In contrast, typicality studies assume that 'flat' categories are taught, but the model assumes that they are stored as a hierarchy of probabilistic concepts. At least two principles might dictate the hierarchical organizations used by humans to encode A and B. Subjects may segregate instances based entirely on the external label (A or B) or they may base an organization, as COBWEB does, on the similarity of objects irrespective of external label. Presumably, these represent the extremes of possible organizational principles. Conveniently, since there is no overlap between categories A and B, COBWEB's approach of grouping similar objects results in the same classes (at the top level) as those based solely on external label.

[2] The level 1 nodes do not maximize collocation for any value. In COBWEB such a node would not be created during concept formation or would be removed once all arcs to it were lost. However, the Hoffman and Ziessler study trained subject's on this classification (i.e., a tutored learning task) – the objective here is to simply test recognition on an existing tree, regardless of how it was learned.

	Letter String	Intra Overlap	Letter String	Inter Overlap	*Typicality*
A	JXPHM	low	4KCTG	high	*low*
	QBLFS	"	GKNTJ	"	"
	XPHMQ	med.	4KC6D	med.	*med.*
	MQBLF	"	HPNSJ	"	"
	PHMQB	high	HPC6B	low	*high*
	HMQBL	"	HPNWD	"	"
B	CTRVG		8SJKT		
	TRVGZ		8SJ3G		
	RVGZK		9UJCG		
	VGZKD		4UZC9		
	GZKDW		4UZRT		
	ZKDWN		MSZR5		
	(1a)		(1b)		

Table 1: Letter strings used to test typicality.

The tree of Figure 4 classifies instances with respect to the node that maximizes 3–4; this is N_1 for each class A member. At N_1 a prediction can be made that a recognized instance is in class A since $P(\text{Class=A}|N_1) = 1.0$. However, symbols that are relatively unique among class A members will cause certain instances to activate arcs to subordinate nodes. In turn, this will detract from the total evidence with which N_1 is predicted. For example, 'HMQBL' has symbols common to most other class A members and it predicts N_1 with a score of 4, while 'QBLFS' has several relatively unique symbols, which reduces prediction of N_1 to 2. A strong assumption of the model is that the time required to reach a node is inversely proportional to the total predictiveness towards that node. Simulated time is computed as

$$\text{time} = \text{distance/rate} = 1.0/\text{total-predictiveness},$$

The distance between any two nodes that are connected by one or more indices is assumed to be 1.0. 'HMQBL' is recognized in $1.0/4 = 0.25$ time units. Figure 5 indicates that in both the human and simulated case, instances with greater intra-class overlap are recognized more quickly.

5.2. Typicality and Inter-Class Similarity

Rosch and Mervis explored the impact of inter- (between-) class similarity using the data of Table 1b; within-class overlap was held constant for class A members, but the extent to which class A members overlapped with B varied from 0 ('HPNWD') to 1.3 ('4KCTG'). Subjects were taught to distinguish categories A and B; category A instances that shared few symbols with strings in category B were recognized more quickly (i.e., were more typical).

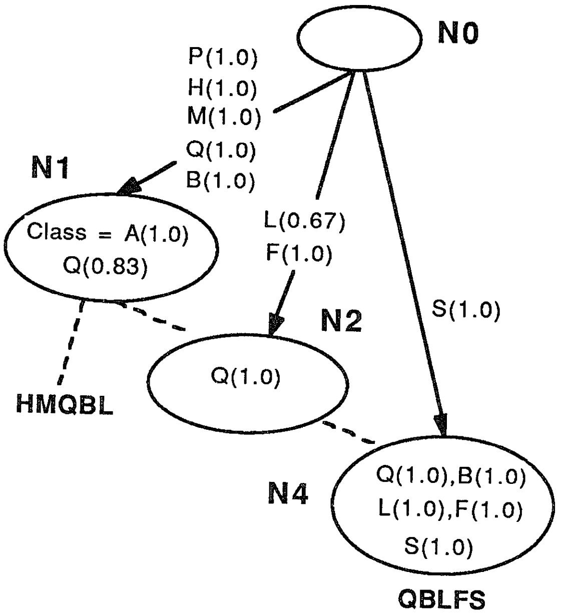

Figure 4: Partial tree that models typicality as function of intra-class similarity

Fisher [Fis87a] used two classification trees to test the indexing scheme: one segregated categories A and B into different subtrees and the second tree was formed by COBWEB, which grouped instances based on similarity. Recognition using both trees agreed with Rosch and Mervis' findings. Figure 6 shows part of the indexed tree produced by COBWEB. Inter-class similarities tend to diffuse evidence across lateral subtrees or a value may not be predictive of any subnode (i.e., collocation is maximized at the root). For example, HPNWD predicts N14 with a total predictiveness of 2.67 at which point a prediction of category A can be made. In contrast, 4KCTG predicts N1 with a score of 1.8; K predicts a subordinate node, and G and T are not predictive of any node. Even when recognition is made with respect to N1, a prediction of class membership (A or B) cannot be made with certainty, causing the classification process to recurse.

5.3. Summary

A strong assumption of the computer model is that the time to transit from one node to a descendent is inversely proportional to the total predictiveness of attribute value indices that are activated during recognition. The model predicts that objects with less intra-category similarity will be recognized slowly (i.e., be less typical) because relatively unique attribute values will diffuse index activation across several levels (i.e., the vertical dimension) of the classification tree. Instances with high inter-category similarity will be be less typical because common inter-class values will diffuse activation across lateral subtrees or will not be predictive at all (i.e., the horizontal dimension).

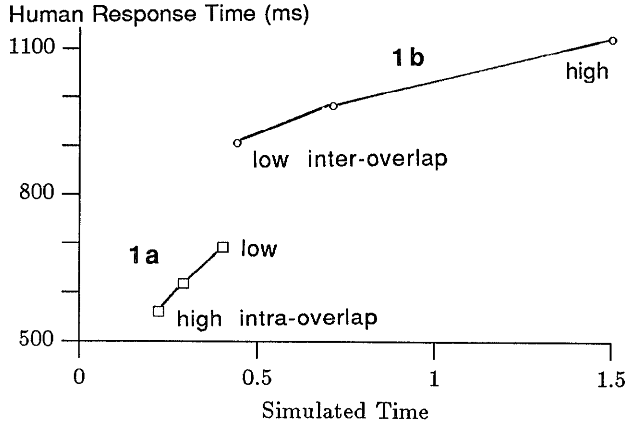

Figure 5: Simulated and human recognition times of letter strings

6. Interactions Between Basic Level and Typicality Effects

Studies by Jolicour, Gluck, and Kosslyn [Joli84] qualify the human preference for the basic level. In particular, an instance (e.g., a particular chicken) may be sufficiently atypical of its basic level class (e.g., *bird*) that it will be first recognized as an instance of a subordinate class (e.g., *chicken*). The model explains the impact of atypicality on basic level preference. Low intra-category overlap results in greater prediction of subordinate nodes. In addition, there is a simultaneous decrease in prediction of the basic level node for atypical objects due to less intra- and more inter- category overlap. These tendencies may coact so that classification is initiated at a subordinate level. Unfortunately, experimental results in easily encoded artificial domains are lacking. However, Fisher [Fis87a] gives a speculative demonstration of the model's consistency in a domain of thyroid patient case histories. Several cases are first classified at a subordinate node; these are the most atypical cases by the intra- and inter- similarity criteria.

A second interaction between basic level and typicality effects is suggested by the model. Traditionally, target recognition tasks that test for typicality have focussed on typicality with respect to a basic level category [Rosc75]. Because classification usually passes through the basic level, an expectation is that recognition with respect to a *subordinate* node will be mediated by the the object's typicality to the basic level, as well as the subordinate node. For example, in the domain of congressional voting records, the tree of Figure 1 shows that N_2 is subordinate to 'conservatives'. N_2 classifies objects that tend to be atypical conservatives like 'Hollings' (a southern-democrat). As expected, 'Hollings' is recognized slowly as a 'conservative' since he is (relatively) atypical of this class. However, 'Hollings' is also slow to be recognized as a member of N_2 even though he is typical of this class (by intra- and inter- similarity criteria)! The model predicts

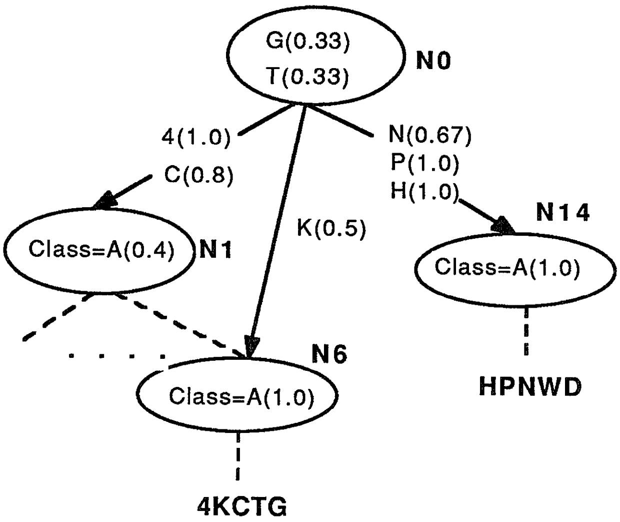

Figure 6: Partial tree that models typicality as function of inter-class overlap.

that atypicality with respect to a basic level concept can offset advantages associated with subordinate node typicality. Apparently, there is no psychological data to support this hypothesis, but it may bolster, weaken, or alter claims for hierarchical representations of category structure should psychological data be forthcoming.

7. Concluding Remarks

This paper presents a memory model that is consistent with data on human basic level and typicality effects. The model draws significantly from previous investigations of the basic level, notably Gluck and Corter [Gluc85] and Jones [Jone83]. However, this work demonstrates how explicit calculation can be 'compiled' into an indexing scheme. Apparently, this is the first computational account of any basic level effect. In addition, the model accounts for typicality data and basic level/typicality interactions. The model also predicts a previously unexplored interaction between basic level and typicality effects. Finally, the COBWEB framework offers a unique opportunity for speculating on the evolution of basic level and typicality effects during learning. Learning with indices is reported in [Fis87a], but it has been downplayed here so that a clear picture of the static memory structure could be described and evaluated. Furthermore, there is little psychological data on which to base claims for basic level and typicality *development*. Thus, there are several areas in which the model can guide experimentation.

Acknowledgements

This work owes much to Dennis Kibler and Pat Langley, as well as early discussions with Mark Gluck and Jim Corter. AAAI reviewers helped clarify exposition.

References

[Fis87a] Fisher, D. *Knowledge acquisition via incremental conceptual clustering*, Technical report 87-22 (doctoral dissertation), Department of Information and Computer Science, University of California, Irvine, (1987).

[Fis87b] Fisher, D. Knowledge acquisition via incremental conceptual clustering. *Machine Learning* 2 (1987), 139-172.

[Gluc85] Gluck, M., and Corter, J. Information, uncertainty, and the utility of categories. *Seventh Annual Conference of the Cognitive Science Society* (Irvine, CA, 1985), Academic Press, Orlando, FL, 283–287.

[Hoff83] Hoffman, J., and Ziessler, C. Objectidentifikation in kunstlichen begriffshierarchien. *Zeitscrift fur Psychologie* 16 (1984), 243-275.

[Joli84] Jolicoeur, P., Gluck, M., and Kosslyn, S. Pictures and names: making the connection. *Cognitive Psychology* 16 (1984), 243-275.

[Jone83] Jones, G. Identifying basic categories. *Psychological Bulletin* 94 (1983), 423-428.

[Kibl87] Kibler, D., and Aha, D. Learning representative exemplars of concepts. *Fourth International Workshop on Machine Learning*, (Irvine, CA, 1986), Morgan Kaufmann, Los Altos, CA, 513-517.

[Kolo83] Kolodner, J. Maintaining organization in a dynamic long-term memory. *Cognitive Science* 7 (1983), 243-280.

[Lebo82] Lebowitz, M. Correcting erroneous generalizations. *Cognition and Brain Theory* 5 (1982), 367-381.

[Merv81] Mervis, C., and Rosch, E. Categorization of natural objects. *Annual Review of Psychology* 32 (1981), 89-115.

[Murp82] Murphy, G., and Smith, E. Basic level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior* 21 (1982), 1-20.

[Quin86] Quinlan, J. R. Induction of decision trees. *Machine Learning* 1 (1986), 81-106.

[Rosc75] Rosch, E., and Mervis, C. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7 (1975), 573-605.

[Rosc76] Rosch, E., Mervis, C., Gray, W., Johnson, D., and Boyer-Braem, P. Basic objects in natural categories. *Cognitive Psychology*, 8 (1976), 382-439.

[Rosc78] Rosch, E. Principles of categorization. In *Cognition and Categorization*, E. Rosch and B. Lloyd, Eds., Lawrence Erlbaum, Hillsdale, NJ (1978), 28-49.

[Smit81] Smith, E., and Medin, D. *Categories and Concepts*, Harvard University Press, Cambridge, MA (1981).