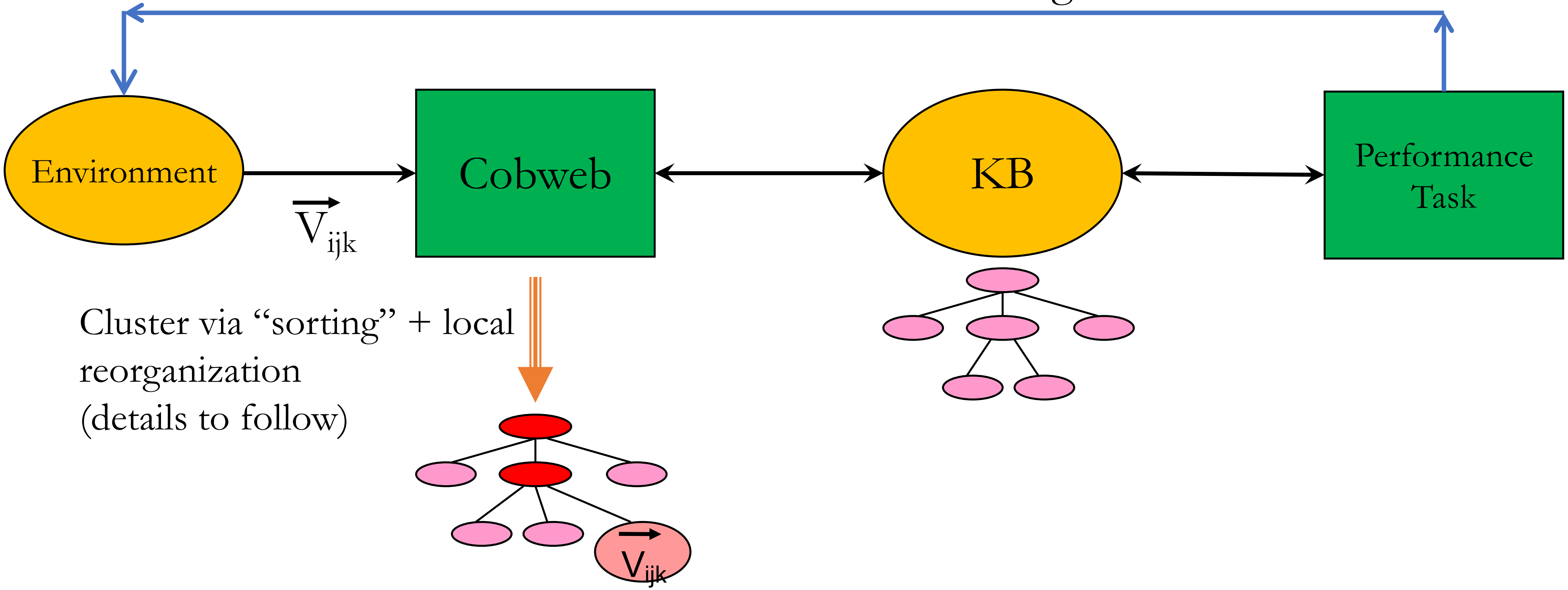
Cobweb Basics

Douglas H. Fisher
Vanderbilt University

- Cobweb as Unsupervised Learning
- The Cobweb Algorithm
- Cobweb Ancestors, and Related Systems and Paradigms
- Cobweb as Agent Memory and Cobweb as Data Mining Tool
- Cobweb Accounts of Psychological Phenomena
- Alternative Quality Measures and Prediction Frontiers

Cobweb as Unsupervised Learning

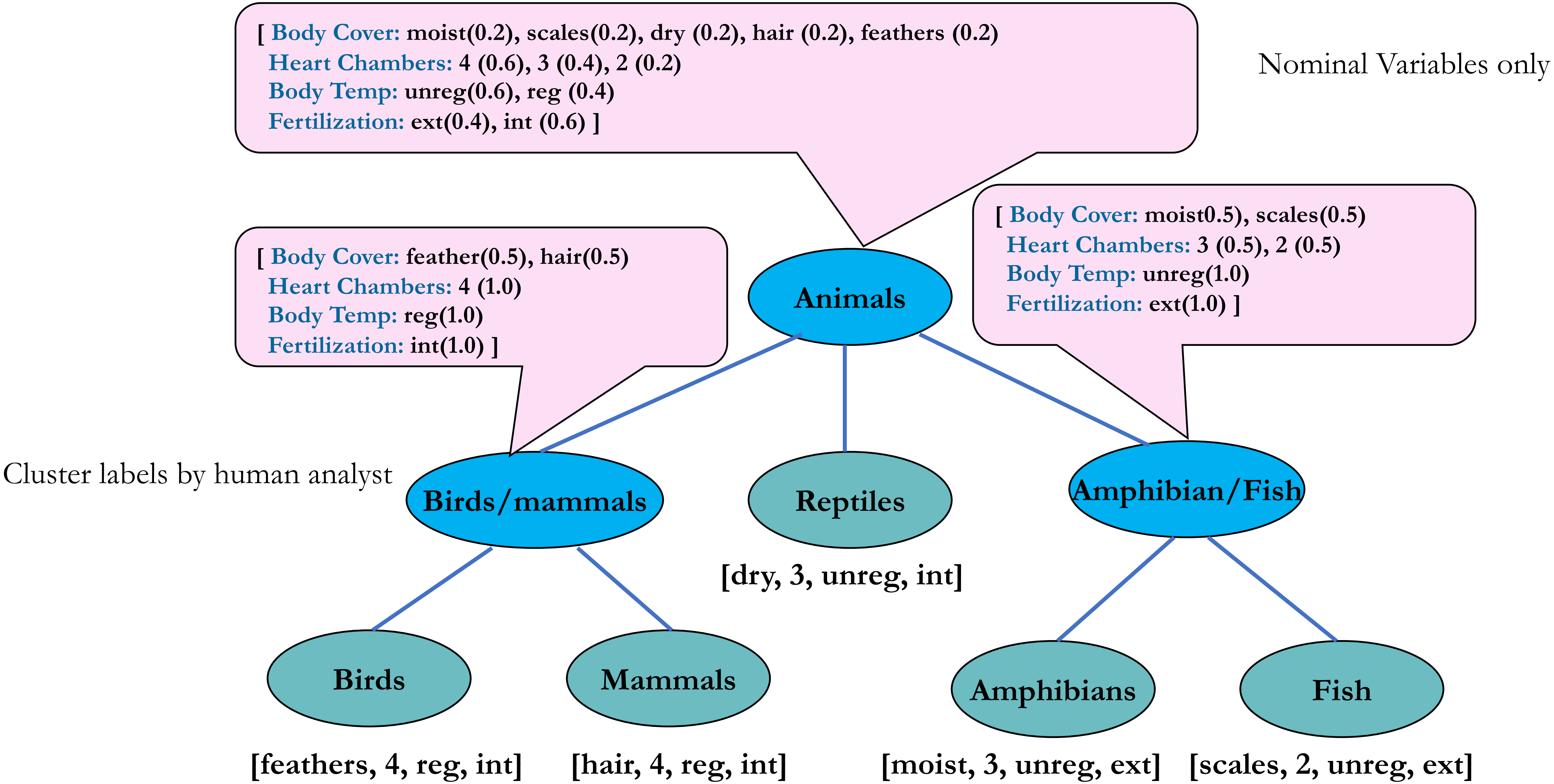
Incremental Hierarchical Clustering



From "[Intelligence in Context]()" (Fisher) Talk to NSF, March 21 2007

- Fisher, D. (1987). "Conceptual Clustering, Learning from Examples, and Inference," *Proceedings of the Fourth International Workshop on Machine Learning*. Irvine, CA: Morgan Kaufmann.
- Fisher, D. (1987) "Knowledge Acquisition Via Incremental Conceptual Clustering," *Machine Learning, 2*, 139–172. Reprinted in Shavlik & T. Dietterich (eds.), *Readings in Machine Learning*, 267–283, Morgan Kaufmann, 1990.

Probabilistic Concept Hierarchies



From "Intelligence in Context" (Fisher) Talk to NSF, March 21 2007

Douglas H. Fisher

The Cobweb Algorithm

FUNCTION COBWEB (Object, Root <of (sub)tree>)

1) Update variable value counts at the Root

2) IF Root is a leaf
   THEN Return expanded leaf to accommodate
      the new object
   ELSE Find that child of Root that **best** hosts
      Object and perform one of the following
      2a) Create a new class if appropriate
      2b) **Merge** nodes if appropriate and call
         COBWEB (Object, Merged node)
      2c) **Split** a node if appropriate and call
         COBWEB (Object, Root)
      2d) IF none of the above (2a,b, or c) then call
         COBWEB (Object, Best child of Root).

Also, **promotion**

From Fisher, D.H. (1987). "Improving Inference Through Conceptual Clustering" Proceedings of AAAI-87, pp. 461-465.

Douglas H. Fisher

Cobweb as Unsupervised Learning

Cobweb's Performance Task of Pattern Completion

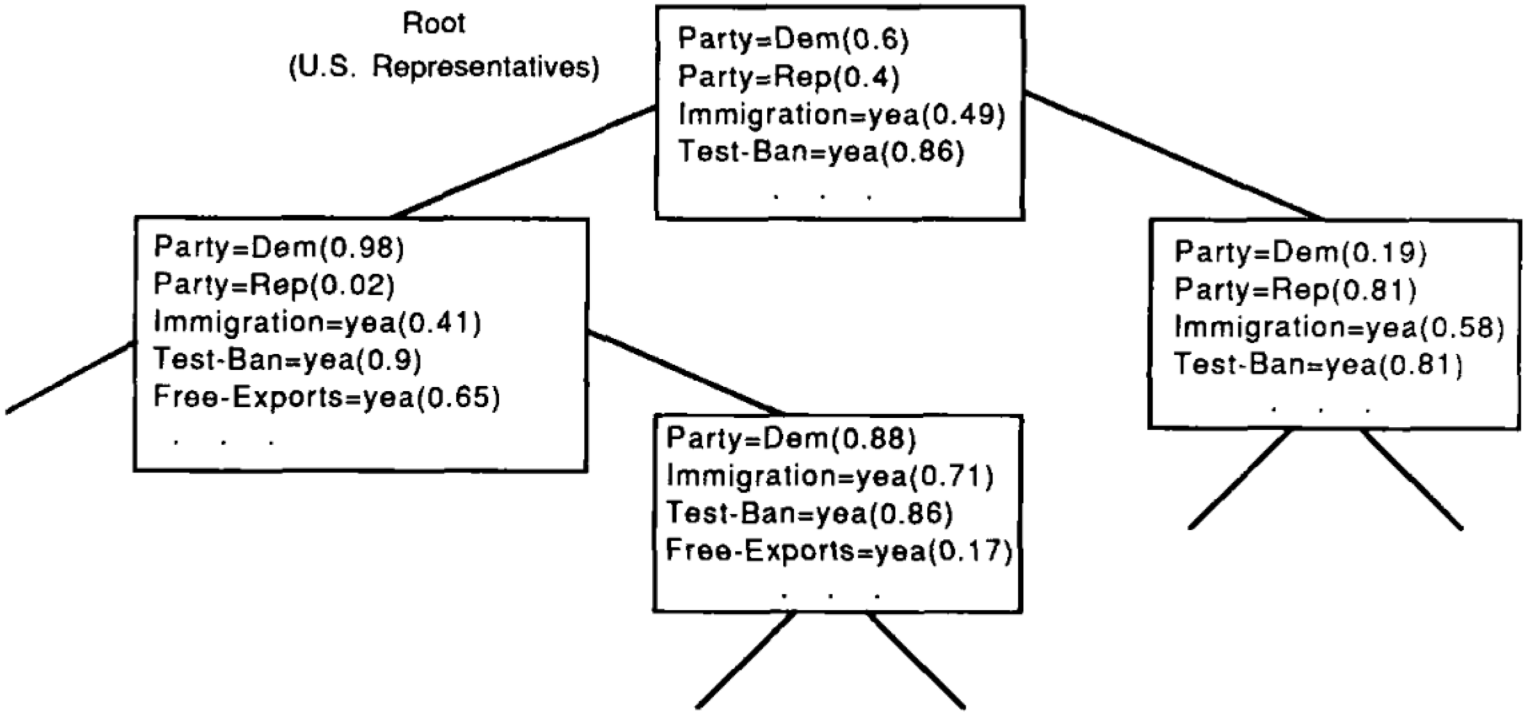


Fig. 2. **A sample probabilistic concept tree over congressional voting records.**

From Fisher, D., & Langley, P. (1990). "The Structure and Formation of Natural Categories," in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

The Cobweb Algorithm

Utility Measures and Complexity

Category Utility: $\mathrm{CU}(\mathrm{C_k}) = P(C_k) \sum_i \sum_j [P(V_i = v_{ij} | C_k)^2 - P(V_i = v_{ij})^2]$

- Gluck, M. A. & Corter, J. E. (1985) Information, uncertainty, and the utility of categories. Proceedings of the Seventh Annual Conference of the Cognitive Science Society (pp. 283-287). Irvine, CA
- Fisher, D. (1987) "Knowledge Acquisition Via Incremental Conceptual Clustering," *Machine Learning, 2*, 139–172.

"… summation over Gini Indices reflected in CU addresses the extent that a cluster predicts the values of all the variables." (Fisher, 1996)

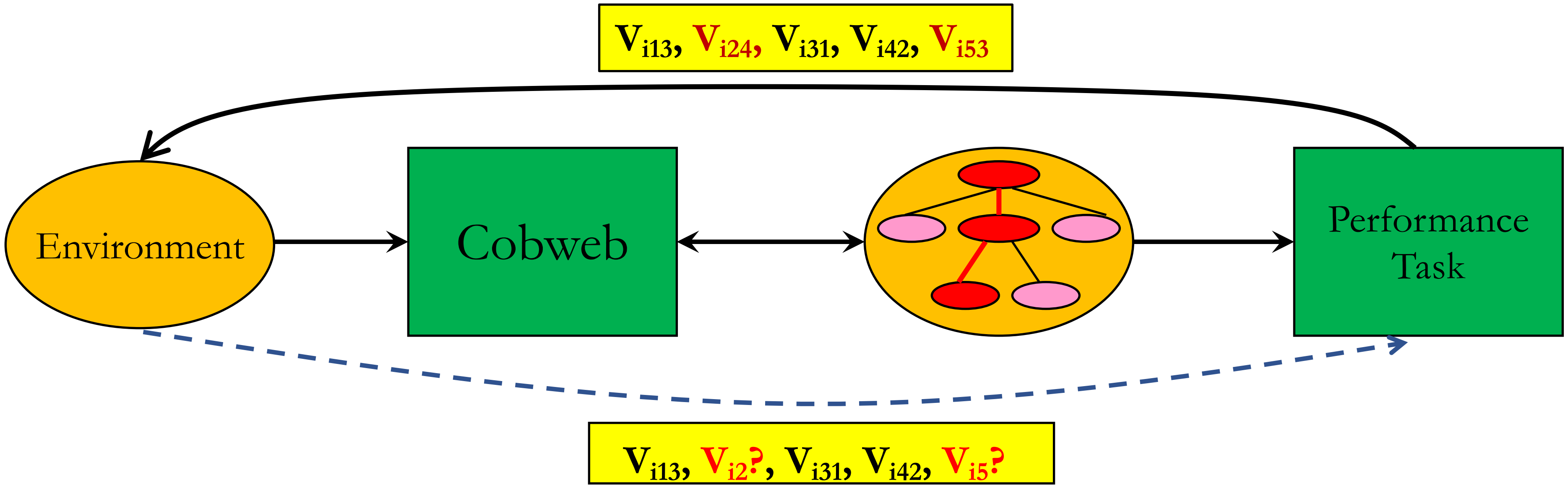
Partition Utility: $\sum_{k=1}^{m} CU(C_k)\ /m$

- Fisher, D. (1987) "Knowledge Acquisition Via Incremental Conceptual Clustering," *Machine Learning, 2*, 139–172.

Complexity: The cost to incorporate a new observation is O(B*D), where B is the branching factor and D is the Depth of the tree. O(B*$\log_B$ N) or simply O($\log_B$ N), where N is the number of previously incorporated observations, and B is considered constant. The cost to incorporate N instances one after the other is O(N log N).

Cobweb as Unsupervised Learning

Cobweb's Performance Task of Pattern Completion

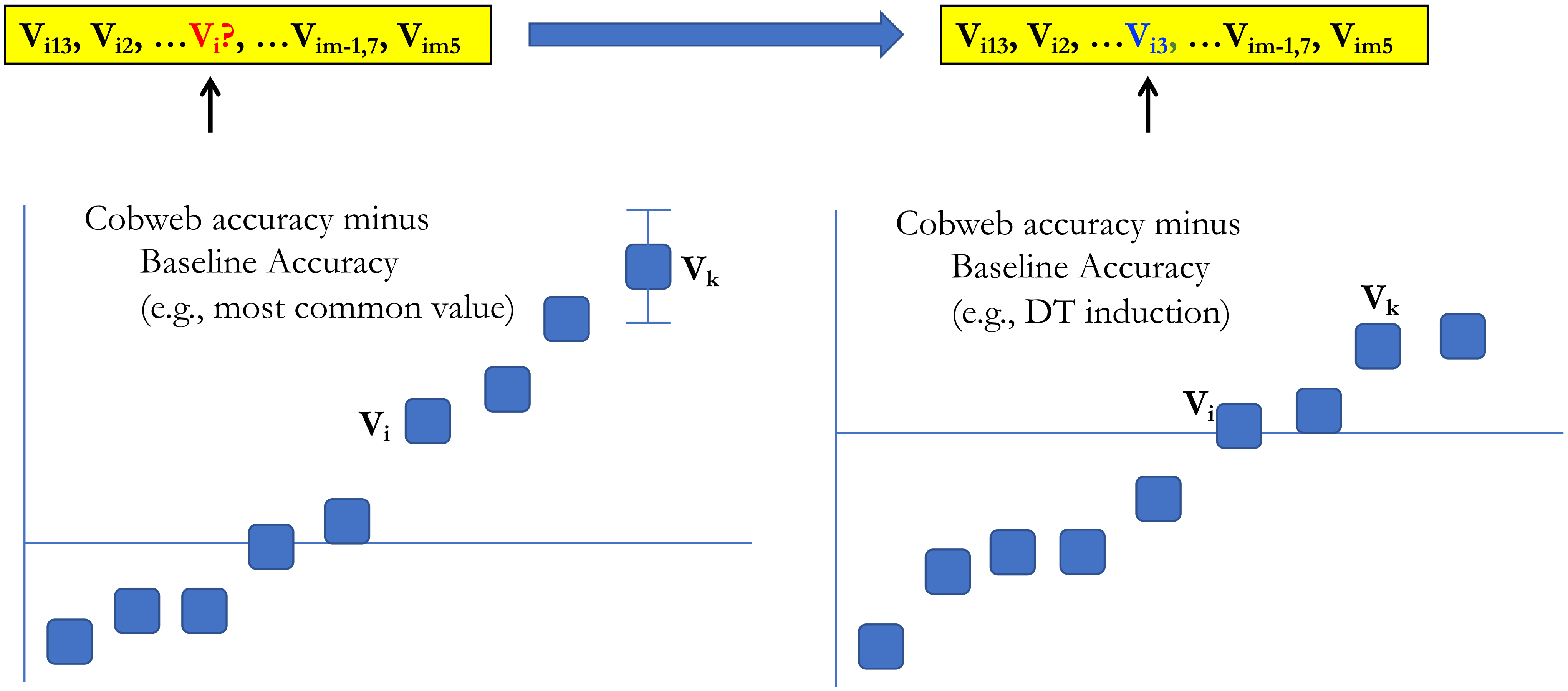


From "Intelligence in Context" (Fisher) Talk to NSF, March 21 2007

- Fisher, D. (1987). "Conceptual Clustering, Learning from Examples, and Inference," *Proceedings of the Fourth International Workshop on Machine Learning*. Irvine, CA: Morgan Kaufmann.
- Fisher, D. (1987) "Knowledge Acquisition Via Incremental Conceptual Clustering," *Machine Learning, 2*, 139–172. Reprinted in Shavlik & T. Dietterich (eds.), *Readings in Machine Learning*, 267–283, Morgan Kaufmann, 1990.
- Fisher, D. (1996). "Iterative Optimization and Simplification of Hierarchical Clusterings," *Journal of Artificial Intelligence Research, 4*, 147–179. (and KDD-95)

Cobweb as Unsupervised Learning

Cobweb's Performance Task of Pattern Completion



Sorted in ascending order of Cobweb's advantage

Douglas H. Fisher

Cobweb as Unsupervised Learning

Cobweb's Performance Task of Pattern Completion

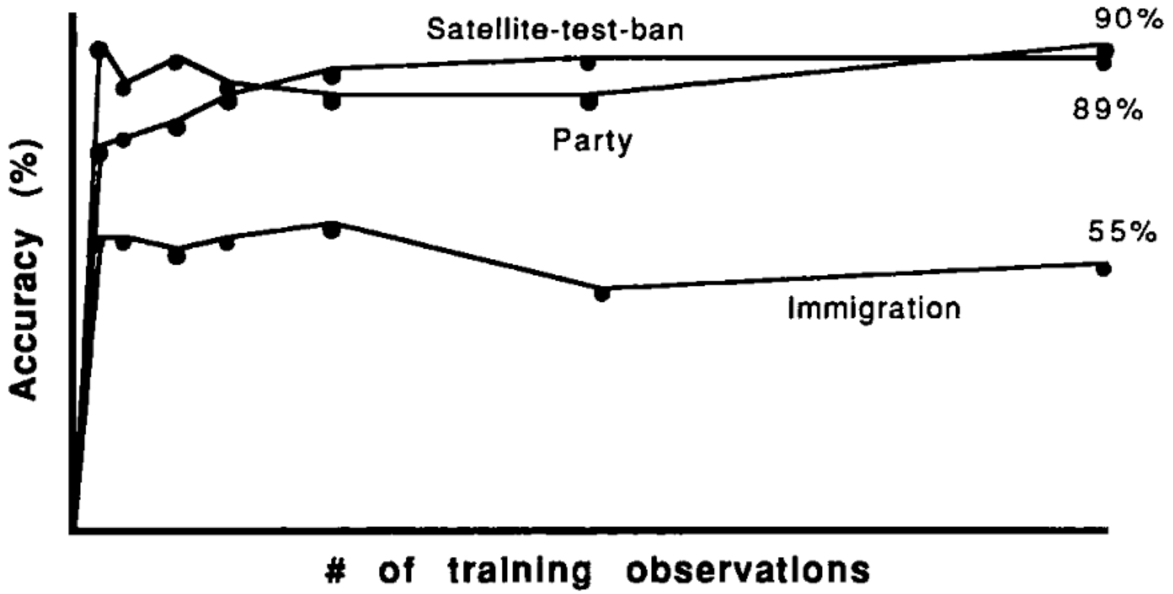


Fig. 3. Learning curves for three attributes in the congressional domain.

From Fisher, D., & Langley, P. (1990). "The Structure and Formation of Natural Categories," in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

Cobweb: Related Systems and Paradigms

Influential Predecessors (Ancestors)

Cobweb resulted from a synthesis of ideas from prior work

- Michalski et al
  - clustering as search
  - conceptual summaries of clusters
- Kolodner and Lebowitz
  - sorting or hill-climbing search
  - motivation for prediction
  - normative, predictable values
- Gluck and Corter
  - a measure for predicting basic levels as an evaluation function)

- Fisher, D. (1987). "Knowledge Acquisition Via Incremental Conceptual Clustering," Technical Report 87-22 (Doctoral Dissertation), Department of Information and Computer Science, University of California, Irvine.

Cobweb: Related Systems and Paradigms

Influential Predecessors (Ancestors)

- Michalski et al

  - Michalski, R. S. (1980). "Knowledge acquisition through conceptual clustering: A theoretical framework and an algorithm for partitioning data into conjunctive concepts" (PDF). International Journal of Policy Analysis and Information Systems. 4: 219–244.

  - Michalski, R. S. & Stepp, R. E. (1983). "Learning from observation: Conceptual clustering" (PDF). In Michalski, R. S.; Carbonell, J. G.; Mitchell, T. M. (eds.). Machine Learning: An Artificial Intelligence Approach. Palo Alto, CA: Tioga. pp. 331–363.

  - Fisher, D.H. & Langley, P. W. (1986). "Conceptual clustering and its relation to numerical taxonomy". In Gale, W. A. (ed.). Artificial Intelligence and Statistics. Reading, MA: Addison-Wesley. pp. 77–116.

Cobweb: Related Systems and Paradigms

Influential Predecessors (Ancestors)

- Kolodner and Lebowitz
  - Kolodner, J. L. (1983). "Maintaining Organization in a Dynamic Long-Term Memory". *Cognitive Science*. ***7*** *(4): 243-280*.
  - Lebowitz, M. (1983). "Generalization from Natural Language Text". *Cognitive Science*. ***7*** *(1): 1–40*.
- Gluck and Corter
  - Gluck, M. A. & Corter, J. E. (1985) Information, uncertainty, and the utility of categories. Proceedings of the Seventh Annual Conference of the Cognitive Science Society (pp. 283-287). Irvine, CA: Lawrence Erlbaum Associates.
  - Corter, J. E. & Gluck, M. A. (1992). "Explaining basic categories: Feature predictability and information". *Psychological Bulletin, 111*(2), *291–303*.

Cobweb Related Systems and Paradigms

Other Related Systems

EPAM

- Feigenbaum, E. A., & Simon, H. A. (1962). A theory of the serial position effect. British Journal of Psychology, 53, 307–320.

- Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. Cognitive Science, 8, 305–336.

- Richman, H. B., (1991). Discrimination Net Models of Concept Formation. In Fisher, D., Pazzani, M.,& Langley, P. (Eds.), Concept formation: Knowledge and Experience in Unsupervised Learning. San Mateo, CA: Morgan Kaufmann. 127–164.

- Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory with EPAM IV. Psychological Review, 102, 305–330.

Cobweb Related Systems and Paradigms

Other Related Systems

Related Systems

- Cheeseman, P., Kelly, J., Self, M., Stutz, J., Taylor, W., & Freeman, D.(1988). Auto Class: A Bayesian classification system. In Proceedings of the Fifth International Machine Learning Conference, pp.54-64. Ann Arbor, MI: Morgan Kaufmann.

- Anderson, J. R., & Matessa, M. (1991). An iterative Bayesian algorithm for categorization. In Fisher, D., Pazzani, M.,& Langley, P. (Eds.), Concept formation: Knowledge and Experience in Unsupervised Learning. San Mateo, CA: Morgan Kaufmann.

Cobweb Related Systems and Paradigms

Related Paradigms

Other approaches to unsupervised learning that can be adapted to pattern completion:

- Learning Association Rule Sets
- Clustering
- Learning Bayesian Networks

- Fisher, D. (2001). Unsupervised Learning (Editorial), *Machine Learning, 45, 1*, 5–7. (Special issue editor on Unsupervised Learning 1).

Related learning paradigms:

- Multi-Task Learning (Caruana, R., 1997, *Machine Learning, 28*, 41-75)
- Data mining clustering
  - Fisher, D. (1996). "Iterative Optimization and Simplification of Hierarchical Clusterings," *Journal of Artificial Intelligence Research, 4*, 147–179. (also KDD-95)
  - Fisher, D. (2002). "Conceptual Clustering," in W. Klosgen and J. Zytkow (eds.), *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press, 388–396, Chapter 16.5.2.

Cobweb as Agent Memory, Cobweb as Data Mining Tool

- Fisher, D. (1988). "A Computational Account of Basic Level and Typicality Effects", *Proceedings of the Seventh National Conference on Artificial Intelligence*. Minneapolis, MN: Morgan Kaufmann, 233–238.

- Carlson, B., Weinberg, J., & Fisher, D. (1990). "Managing Search Using Incremental Conceptual Clustering" *Seventh International Conference on Machine Learning*. Austin, TX: Morgan Kaufmann.

- Fisher, D. (1995). "Optimization and Simplification of Hierarchical Clusterings," *First International Conference on Knowledge Discovery in Databases*, Montreal, Canada: AAAI Press, 118–123.

- Also, see Pat Langley and Chris MacLellan presentations

Cobweb Models of Psychological Effects:

- Basic Level Effects
- Typicality Effects
- Fan Effects
- Linearly Separability

- Fisher, D. (1988). “A Computational Account of Basic Level and Typicality Effects”, *Proceedings of the Seventh National Conference on Artificial Intelligence*. Minneapolis, MN: Morgan Kaufmann, 233–238.
- Silber, J., & Fisher, D. (1989). “A Model of Natural Category Structure and its Behavioral Implications,” Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, MI: Lawrence Erlbaum, 884–891.
- Fisher, D., & Langley, P. (1990). “The Structure and Formation of Natural Categories,” in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

An Algorithmic Model of Psychological Effects

Assumption: Response times and error rates on stimuli are strongly correlated with a measure of match strength; these are not deeply mechanistic accounts.

Category Utility: $\mathrm{CU}(\mathrm{C_k}) = P(C_k) \sum_i \sum_j [P(V_i = vi_j | C_k)^2 - P(V_i = vi_j)^2]$

- Gluck, M. A. & Corter, J. E. (1985) Information, uncertainty, and the utility of categories. Proceedings of the Seventh Annual Conference of the Cognitive Science Society (pp. 283-287). Irvine, CA
- Fisher, D. (1987) "Knowledge Acquisition Via Incremental Conceptual Clustering," *Machine Learning, 2*, 139–172.

Category Match: $\mathrm{C}M(C_{\mathrm{k}}, Obs) = P(C_k) \sum_i [P(V_i(Obs) | C_k)^2 - P(V_i(Obs))^2]$

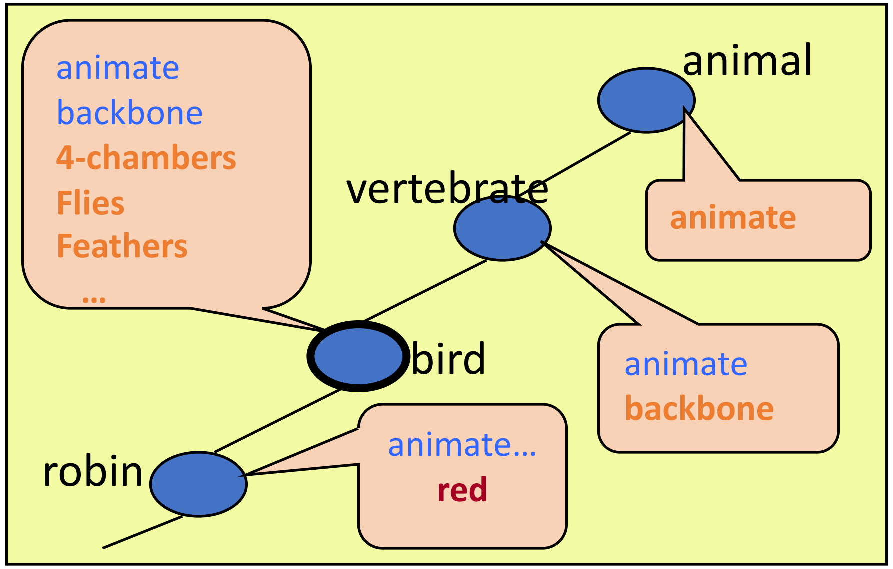
- Silber, J., & Fisher, D. (1989). "A Model of Natural Category Structure and its Behavioral Implications," Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, MI: Lawrence Erlbaum, 884–891.
- Fisher, D., & Langley, P. (1990). "The Structure and Formation of Natural Categories," in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

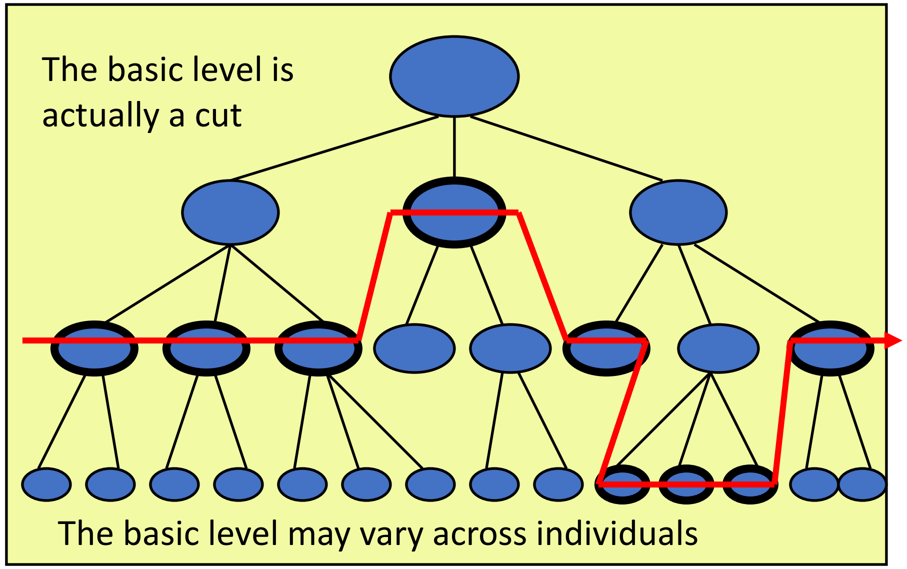
Partition Utility: $\sum_{k=1}^{m} CU(Ck) / m$

- Fisher, D. (1987) "Knowledge Acquisition Via Incremental Conceptual Clustering," *Machine Learning, 2*, 139–172.

Basic Levels: Getting the Most Bang for the Buck

From "Intelligence in Context" (Fisher) Talk to NSF, March 21 1987





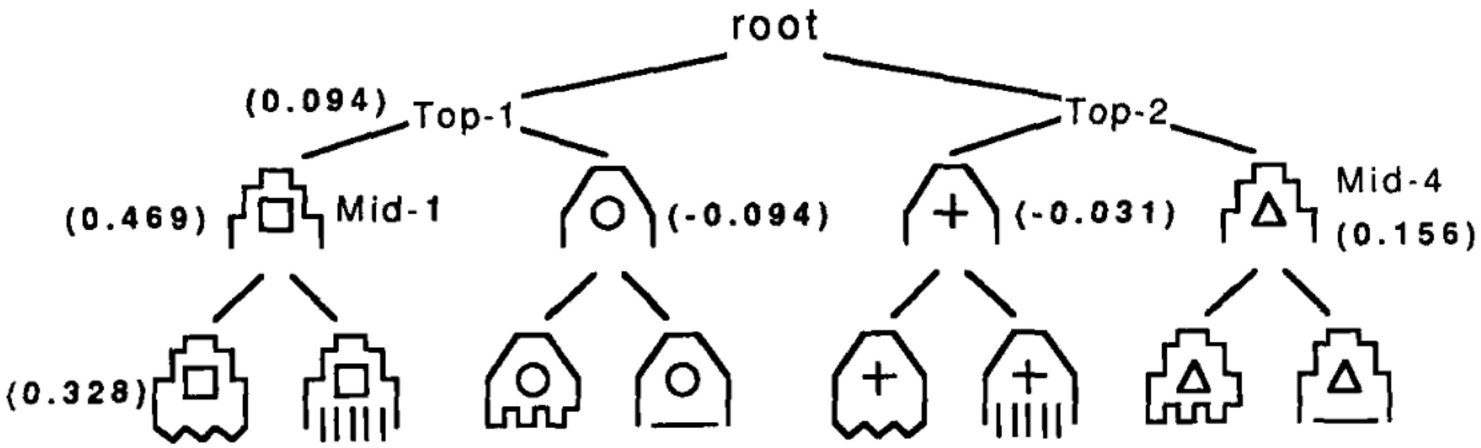


Fig. 4. Approximation of a tree from Hoffman and Ziessler basic-level studies.

Douglas H. Fisher

Typicality and Fan Effects

Typicality effects: members of a category that are most similar to other members of the same category and most dissimilar to members of contrast categories tend to be regarded and treated as more typical of their category.

- Fisher, D. (1988). "A Computational Account of Basic Level and Typicality Effects", *Proceedings of the Seventh National Conference on Artificial Intelligence*. Minneapolis, MN: Morgan Kaufmann, 233–238.

Fan effects: single objects that are more distinct from other objects in a population tend to be recognized more quickly and reliably

- Silber, J., & Fisher, D. (1989). "A Model of Natural Category Structure and its Behavioral Implications," Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, MI: Lawrence Erlbaum, 884–891.

Fan effects restated: a degenerate case of typicality effects, in which intra-category similarity is not a factor, but inter-category dissimilarity is entirely responsible for the behavioral effects

Typicality Effects

	Letter String	Intra-Category Overlap	*Typicality*		Letter String	Inter-Category Overlap	*Typicality*
	JXPHM	low	*low*		HPNWD	low	*high*
	QBLFS	"	"		HPC6B	"	"
A	XPHMQ	medium	*medium*	A	HPNSJ	medium	*medium*
	MQBLF	"	"		4KC6D	"	"
	PHMQB	high	*high*		GKNTJ	high	*low*
	HMQBL	"	"		4KCTG	"	"
	CTRVG				8SJKT		
	TRVGZ				8SJ3G		
B	RVGZK			B	9UJCG		
	VGZKD				4UZC9		
	GZKDW				4UZRT		
	ZKDWN				MSZR5		

Fig. 5. Nonsense strings used to test typicality differences.

- Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Typicality Effects

TABLE VI

AVERAGE RESPONSE TIMES AND CATEGORY-MATCH RANKINGS FOR ROSCH AND MERVIS (1975) DATA

	Response time (msec)	Category match (local)	Category match (COBWEB)
Intraoverlap (a)			
High	560	0.948	0.910
Medium	617	0.823	0.832
Low	692	0.594	0.736
Interoverlap (b)			
Low	909	0.306	0.488
Medium	986	0.196	0.461
High	1125	0.120	0.396

TABLE VII

HUMAN AND PREDICTED RESPONSE TIMES FOR ROSCH AND MERVIS (1975) DATA

	Response time (msec)	Predicted time (local)	Predicted time (COBWEB)
Intraoverlap (a)			
High	560	526	535
Med	617	606	615
Low	692	753	713
Interoverlap (b)			
Low	909	938	968
Med	986	1008	995
High	1125	1057	1062

- Fisher, D., & Langley, P. (1990). "The Structure and Formation of Natural Categories," in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

Fan Effects

Anderson, J. R. (1974). Retrieval of propositional information from long term memory. *Cognitive Psychology, 6* , 451-474.

(1-1) The doctor is in the bank. **(1-2) The fireman is in the park.**
(2-1) The teacher is in the church. **(2-2) The teacher is in the park.**

	Trues		
	1	2	3
1	1111ms (1120)	1174ms (1157)	1222ms (1184)
2	1167ms (1157)	1198ms (1195)	1222ms (1259)
3	1153ms (1184)	1233ms (1259)	1357ms (1321)

	Falses		
	1	2	3
1	1197ms (1168)	1221ms (1240)	1264ms (1306)
2	1250ms (1240)	1356ms (1312)	1291ms (1379)
3	1262ms (1306)	1471ms (1379)	1465ms (1444)

Fig. 8. Human and predicted (in parentheses) response times for Anderson's (1974) fan-effect data.

- Fisher, D., & Langley, P. (1990). "The Structure and Formation of Natural Categories," in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

Linearly Separable and Non-Separable Categories

"Medin (1983) suggests that if independent-cue models are the basis of human conceptual structure, then linearly separable categories should be easier to learn than nonlinearly separable ones." but "Subjects judged the linearly separable set more difficult to learn, and this set also resulted in more recognition errors"

TABLE I

LINEARLY SEPARABLE AND NONLINEARLY SEPARABLE CATEGORIES[a]

	Category C_1					Category C_2			
	V_1	V_2	V_3	V_4		V_1	V_2	V_3	V_4
				Linearly separable objects					
1)	1	1	1	0	5)	1	0	1	0
2)	1	0	1	1	6)	0	1	1	0
3)	1	1	0	1	7)	0	0	0	1
4)	0	1	1	1	8)	1	1	0	0
				Nonlinearly separable objects					
9)	1	0	0	0	13)	0	0	0	1
10)	1	0	1	0	14)	0	1	0	0
11)	1	1	1	1	15)	1	0	1	1
12)	0	1	1	1	16)	0	0	0	0

[a]From Medin (1983).

- Medin, D. L. (1983). Structural principles of categorization. In T. Tighe & B. Shepp (Eds.), Perception, cognition, and development. Hillsdale, NJ: Erlbaum

Linearly Separable and Non-Separable Categories

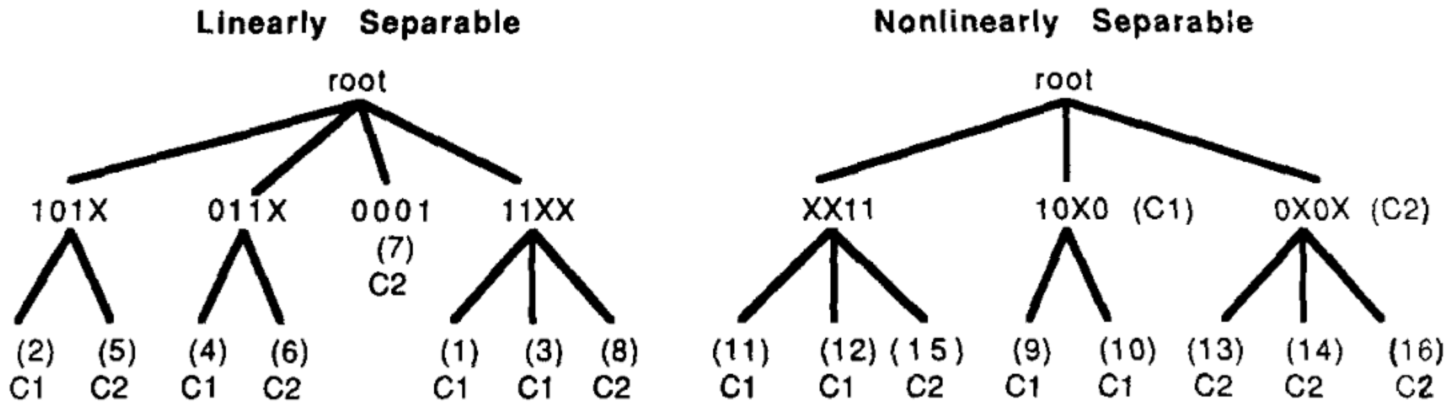


Fig. 1. Concept trees over nonlinearly and linearly separable categories.

"Medin's finding can be explained in terms of the average depth to which observations must be classified before one can perfectly distinguish members of C , from C,. The linearly separable set requires an average depth of 1.87 before reaching a node that contains only members of one category; in contrast, the nonlinearly separable set has 1.37 as its average depth."

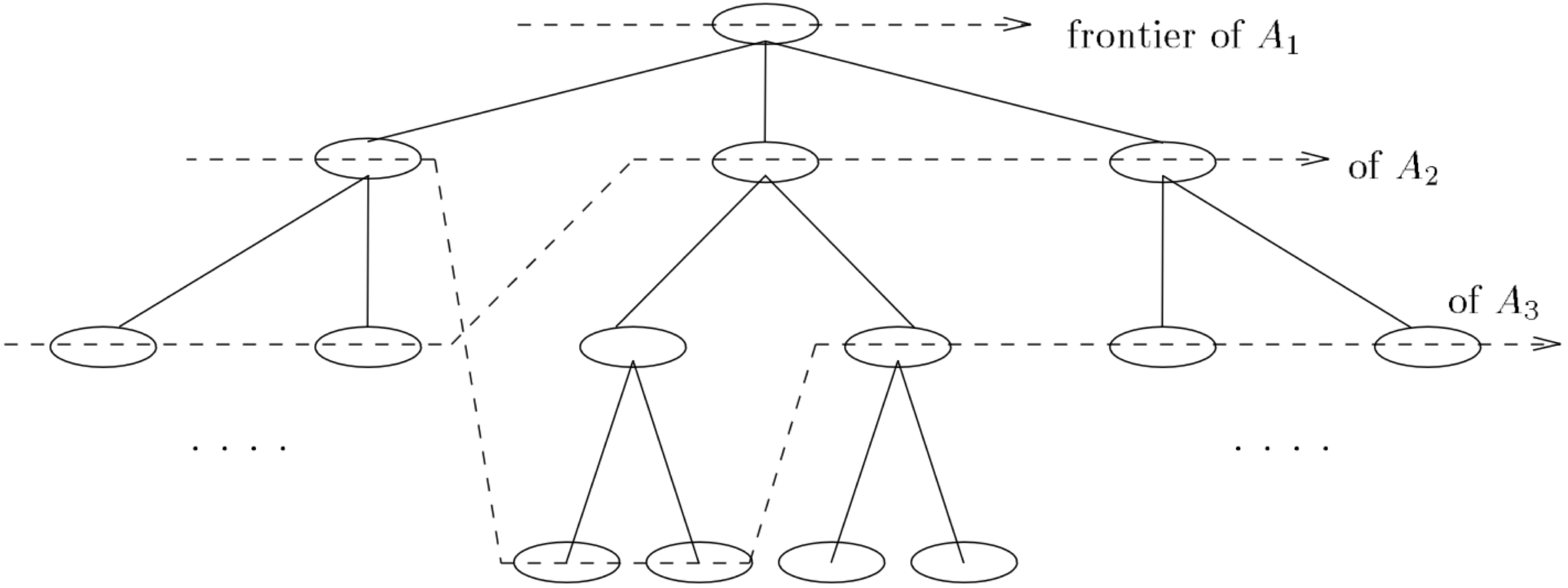
"Our demonstration is simplified, but it nonetheless illustrates that hierarchies or other networks of independent-cue concepts have the same representational power as exemplar and relational-cue models."

- Fisher, D., & Langley, P. (1990). "The Structure and Formation of Natural Categories," in G. Bower (ed.), *The Psychology of Learning and Motivation, 26*, San Diego, CA: Academic Press, 241–284.

Douglas H. Fisher

Alternative Formulations of Basic Cobweb

Identifying Variable Frontiers for Prediction



- Fisher, D. (1996). "Iterative Optimization and Simplification of Hierarchical Clusterings," *Journal of Artificial Intelligence Research, 4*, 147–179.

- Fisher, D. (1989). "Noise-Tolerant Conceptual Clustering" *Proceedings of the International Joint Conference on Artificial Intelligence*, Detroit, MI: Morgan Kaufmann, 825–830. **Incremental approach using "self-supervised" learning.**

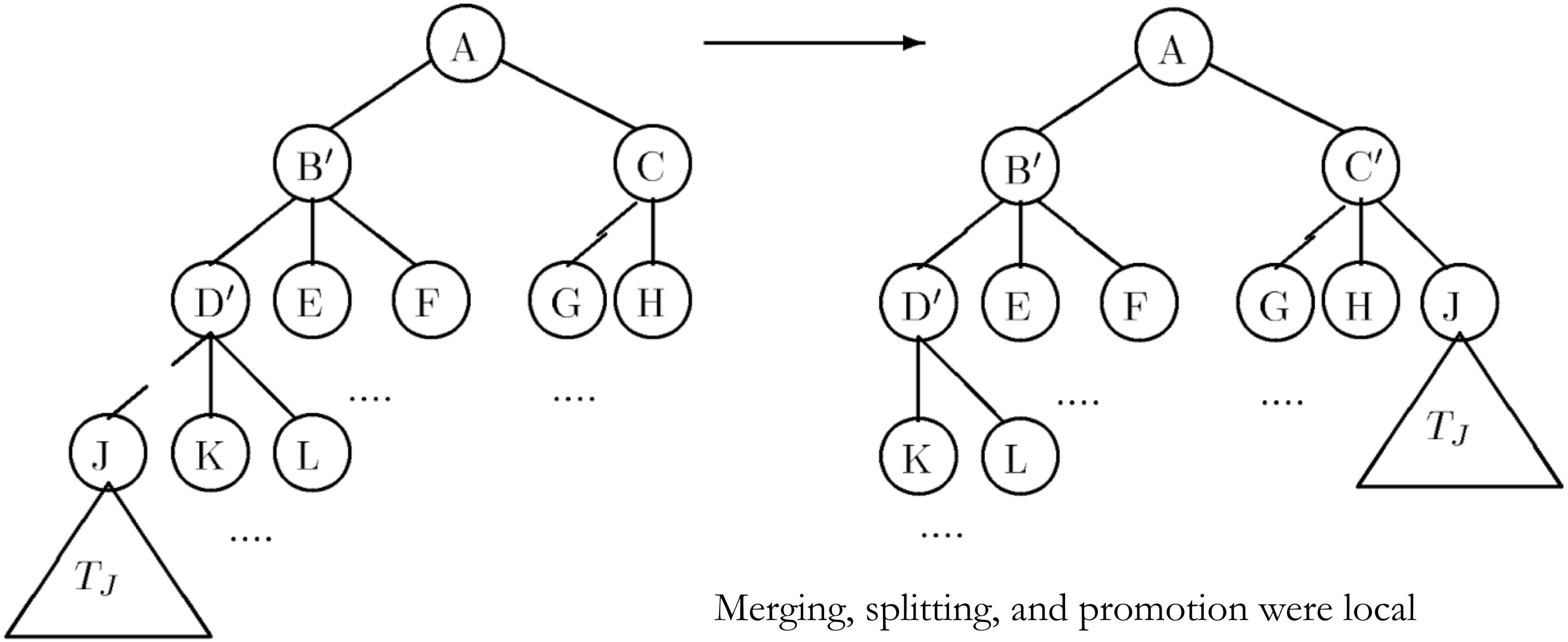
Results with Variable Frontiers for Prediction

	Unvalidated	Validated
Soybean (small)		
Leaves	18.00 (0.00)	13.10 (1.59)
Accuracy	0.85 (0.01)	0.85 (0.01)
Ave. Frontier Size	18.00 (0.00)	2.75 (1.17)
Soybean (large)		
Leaves	122.00 (0.00)	79.10 (5.80)
Accuracy	0.83 (0.02)	0.83 (0.02)
Ave. Frontier Size	122.00 (0.00)	17.01 (4.75)
House		
Leaves	174.00 (0.00)	49.10 (7.18)
Accuracy	0.76 (0.02)	0.81 (0.01)
Ave. Frontier Size	174.00 (0.00)	9.90 (5.16)
Mushroom		
Leaves	400.00 (0.00)	96.30 (11.79)
Accuracy	0.80 (0.01)	0.82 (0.01)
Ave. Frontier Size	400.00 (0.00)	11.07 (4.28)

Table 6: Characteristics of optimized clusterings before and after validation. Average and standard deviations over 20 trials.

Alternative Formulations of Basic Cobweb

global optimization: sorting followed by hierarchical redistribution



Merging, splitting, and promotion were local

- Fisher, D. (1995). "Optimization and Simplification of Hierarchical Clusterings," *First International Conference on Knowledge Discovery in Databases*, Montreal, Canada: AAAI Press, 118–123.

- Fisher, D. (1996). "Iterative Optimization and Simplification of Hierarchical Clusterings," *Journal of Artificial Intelligence Research, 4*, 147–179. **From Cobweb as an agent's memory to Cobweb as a data mining tool**

Douglas H. Fisher

	Unoptimized		Optimized	
	Unvalidated	Validated	Unvalidated	Validated
Soybean (small)				
Leaves	18.00 (0.00)	15.35 (1.81)	18.00 (0.00)	13.10 (1.59)
EPL	40.90 (3.64)	31.90 (6.94)	54.20 (4.74)	34.50 (6.49)
Depth*	2.27	2.08	3.01	2.63
Breadth*	3.57	3.72	2.61	2.66
Cost*	8.10	7.74	7.86	7.00
Soybean (large)				
Leaves	122.00 (0.00)	88.55 (4.46)	122.00 (0.00)	79.10 (5.80)
EPL	437.20 (34.74)	280.40 (28.07)	657.65 (28.38)	380.65 (43.63)
Depth*	3.58	3.17	5.39	4.81
Breadth*	3.82	4.11	2.44	2.48
Cost*	13.68	13.03	13.15	11.93
House				
Leaves	174.00 (0.00)	68.95 (8.15)	174.00 (0.00)	49.10 (7.18)
EPL	664.65 (41.16)	196.20 (35.32)	1005.10 (27.42)	217.25 (39.75)
Depth*	3.82	2.85	5.78	4.42
Breadth*	3.86	4.42	2.44	2.41
Cost*	14.75	12.60	14.10	10.65
Mushroom				
Leaves	400.00 (0.00)	145.50 (20.64)	400.00 (0.00)	96.30 (11.79)
EPL	2238.20 (123.63)	660.90 (117.86)	2608.85 (56.01)	503.40 (72.22)
Depth*	5.60	4.54	6.52	5.23
Breadth*	2.92	3.00	2.51	2.39
Cost*	16.35	13.62	16.37	12.50

- Fisher, D. (1996). "Iterative Optimization and Simplification of Hierarchical Clusterings," *Journal of Artificial Intelligence Research, 4*, 147–179.

Alternative Formulations of Basic Cobweb

Alternative utility measures

"… averaging CU over the clusters of a partition introduces cliffs in the space of partitions; it is likely that better objective functions can be found." (Fisher, 1996).

$$\mathrm{CU}(C_k) = P(C_k) \sum_i \sum_j [P(A_i = V_{ij}|C_k) \log_2 P(A_i = V_{ij}|C_k) - P(A_i = V_{ij}) \log_2 P(A_i = V_{ij})].$$

Partition Utility Alternatives

$$\sum_i \frac{\sum_k P(C_k) \sum_j [P(A_i{=}V_{ij}|C_k) \log_2 P(A_i{=}V_{ij}|C_k) - P(A_i{=}V_{ij}) \log_2 P(A_i{=}V_{ij})]}{-\sum_k P(C_k) \log_2 P(C_k)}$$

Inspired by Quinlan, 1986

$$\sum_i \frac{\sum_k P(C_k) \sum_j [P(A_i{=}V_{ij}|C_k) \log_2 P(A_i{=}V_{ij}|C_k) - P(A_i{=}V_{ij}) \log_2 P(A_i{=}V_{ij})]}{-\sum_k \sum_j P(A_i{=}V_{ij} \wedge C_k) \log_2 P(A_i{=}V_{ij} \wedge C_k)}$$

Inspired by de Mantaras, 1991

- Fisher, D. (1996). "Iterative Optimization and Simplification of Hierarchical Clusterings," *Journal of Artificial Intelligence Research, 4*, 147–179.

Takeaways (at the time, maybe still)

- Probabilistic Concept Hierarchies are representations composed of simple components, and that overcome limitations for other concept models
  - Concept – summary descriptions facilitate inference and serve as proxies for their implied membership in local (or global) optimization
  - Hierarchy – inference and optimization at multiple levels of abstraction,
  - Probabilistic – uncertain inference with rewards for groupings that capture variable inter-dependencies

- Pattern completion as an objective performance task for unsupervised learning (of most any form)

- Incremental (online) learning as the basis of a bounded rational agent

- The interplay of agent modeling and data analysis/mining tools