Comments on Anderson's paper

- The rational analysis (Bayesian)
  - Base rate
  - Coupling probability
  - Independence (of features) assumption
  - Relationship to mechanistic models

- Categories based on overlapping features and based on shared function

- The iterative algorithm
  - Ordering effects

- Hierarchical categorizations
- Possibility of overlapping categories
- A research proposal (optimization along paths rather than partitions) – DF's

Modeling basic level, typicality, and fan effects

Category learning over data through unsupervised learning
See references on slide 8, as well as "
Knowledge Acquisition Via Incremental Conceptual Clustering" (Fisher, 1987)

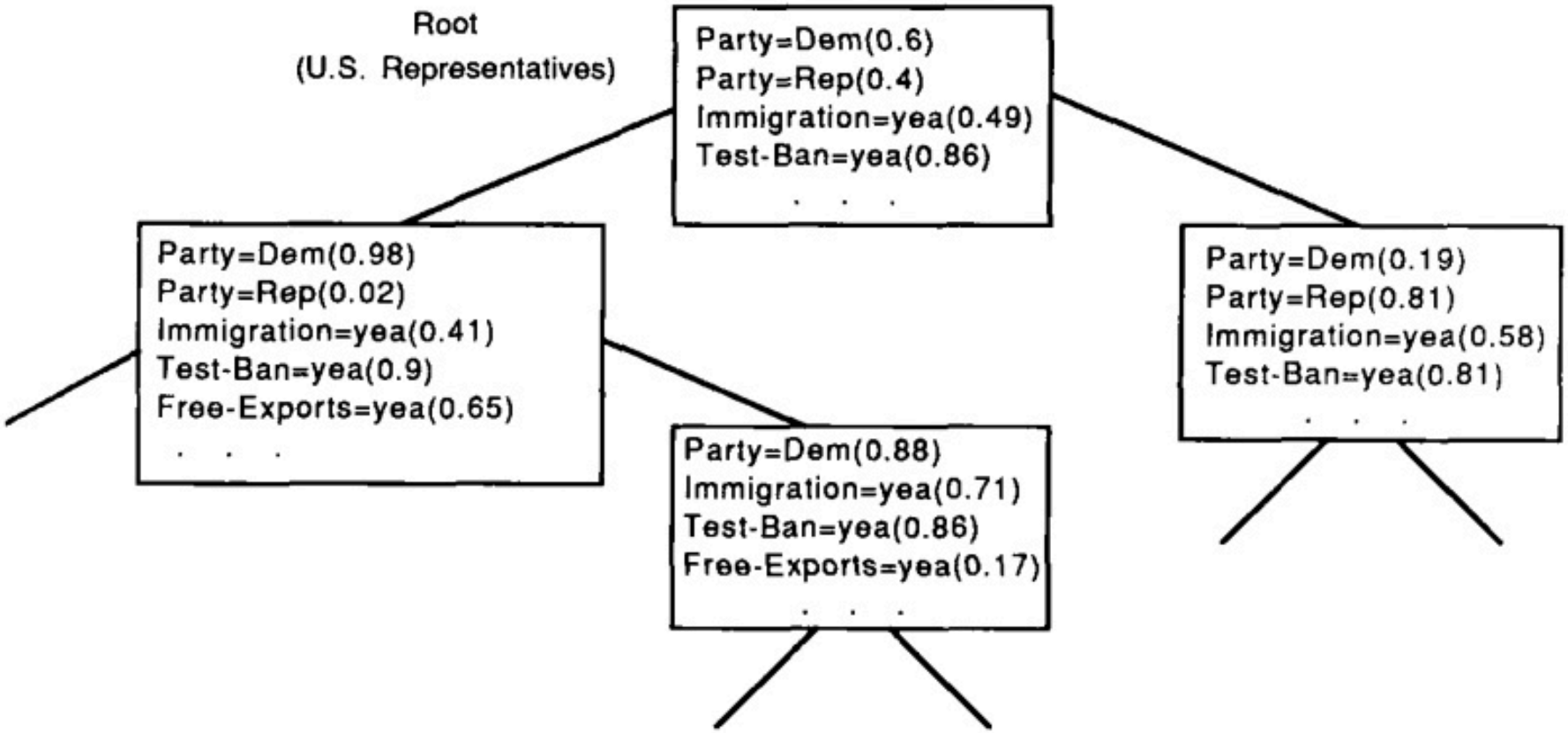

Fig. 2. A sample probabilistic concept tree over congressional voting records.

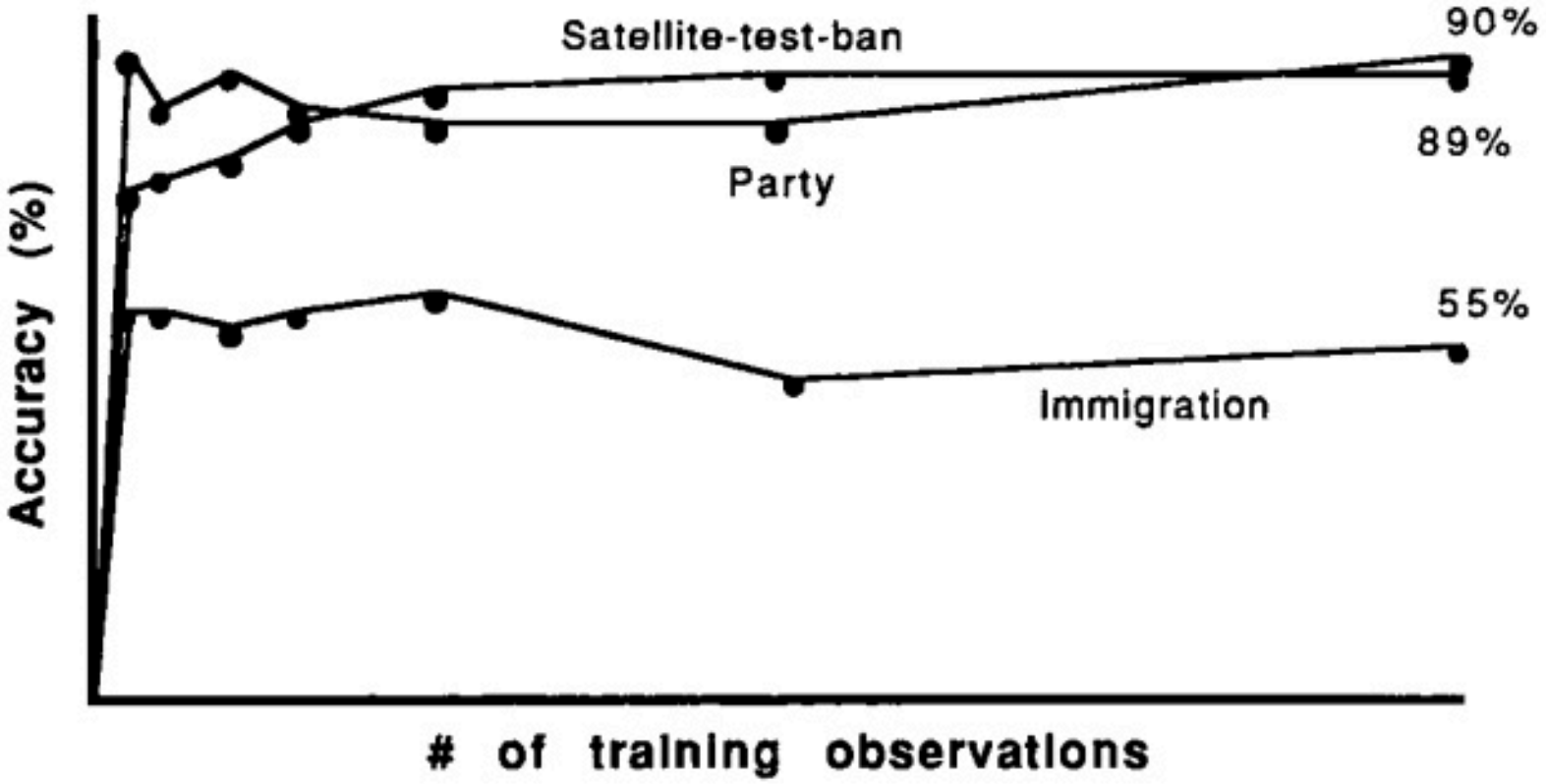

Fig. 3. Learning curves for three attributes in the congressional domain.

Table 5. Descriptions of the 'Charcoal Rot' cluster from the soybean tree. The predictability and predictiveness for each normative value are given in brackets, i.e., [P(value|Charcoal Rot), P(Charcoal Rot|value)].

	N_2 ('Charcoal Rot') $[P(value\|N_2),\ P(N_2\|value)]$
Normative values	Precipitation = below-normal [1.0, 1.0] Temperature = above-normal [0.60, 1.0] Stem-cankers = absent [1.0, 1.0] Fruit-pod-condition = normal [1.0, 0.50] Canker-lesion-color = tan [1.0, 1.0] Outer-stem-decay = absent [1.0, 0.48] Internal-stem-discoloration = black [1.0, 1.0] Sclerotia-internal-external = present [1.0, 1.0]

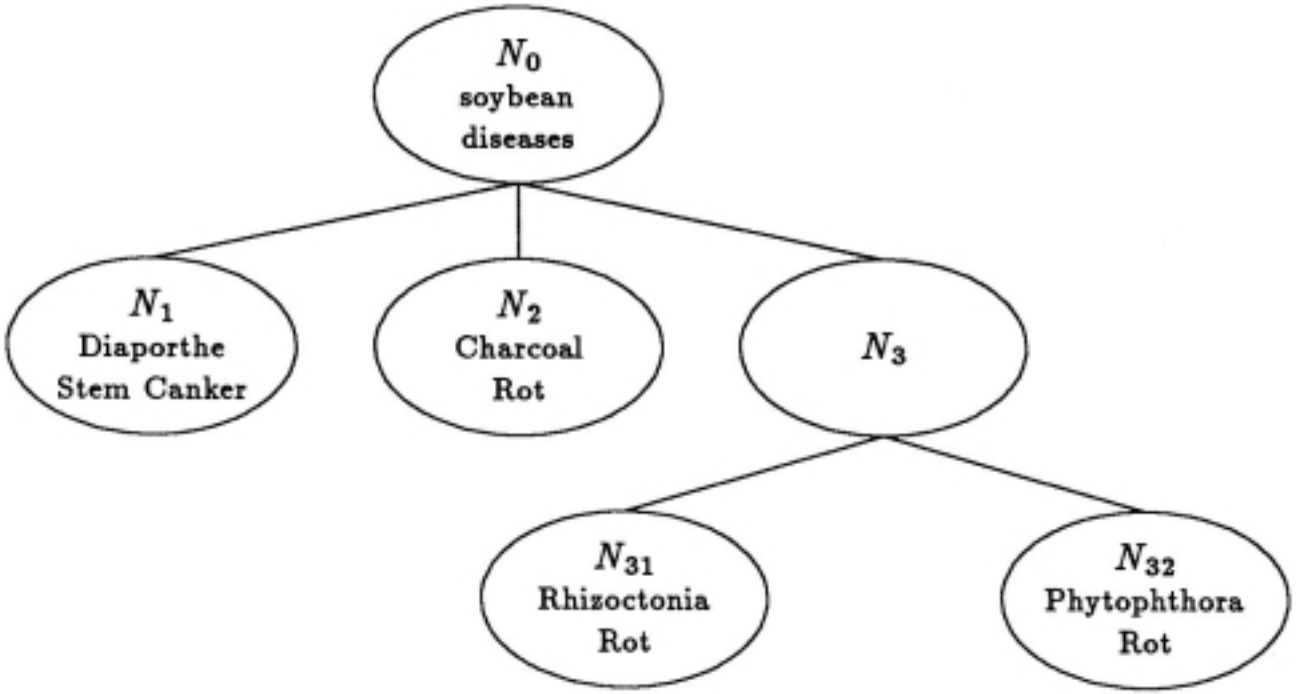

Figure 6. Classification tree of soybean case histories.

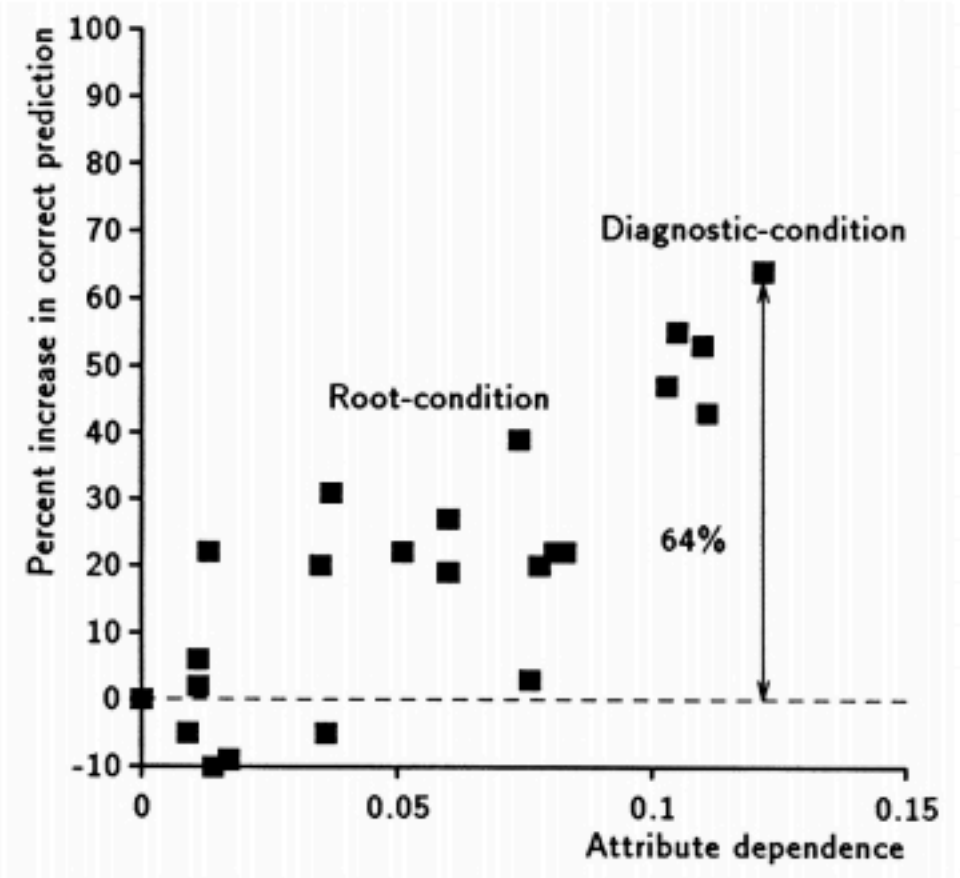

Figure 9. Increase in correct soybean inference as a function of dependence.

Category Match

$$P(C_k)\sum_j [P(V_j|C_k)^2 - P(V_j)^2]$$

- C_k is category k in a set of categories (e.g., a level in a categorization tree)
- V_j is an observation's value along the jth attribute
- The category match score is highly, positively correlated with behavior variables like response time across a large number of studies

Table 2. Probabilistic representation of {fish, amphibian, mammal}.

Attributes	*Values and probabilities*
BodyCover	scales [0.33], moist-skin [0.33], hair [0.33]
HeartChamber	two [0.33], three [0.33], four [0.33]
BodyTemp	unregulated [0.67], regulated [0.33]
Fertilization	external [0.67], internal [0.33]

The incremental (iterative) algorithm

- Classifying an object in an existing class
- Creating a new class
- Combining two classes into a single class
- Dividing a class into several classes

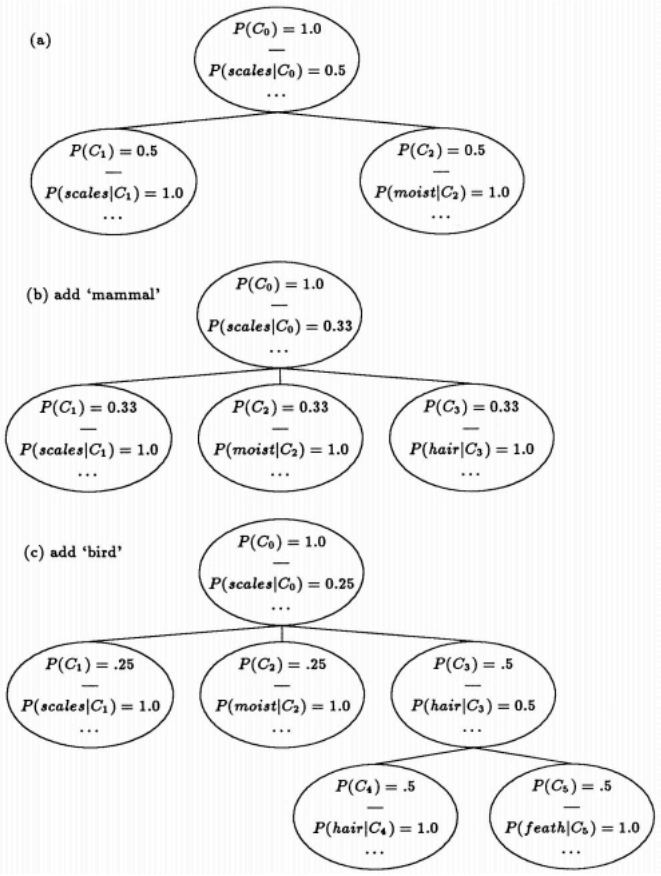

Figure 3. Adding 'mammal' and 'bird' to an existing classification tree. Each node represents an object class, C_i, that is summarized by a set of probabilities, $P(\text{value}|C_i)$.

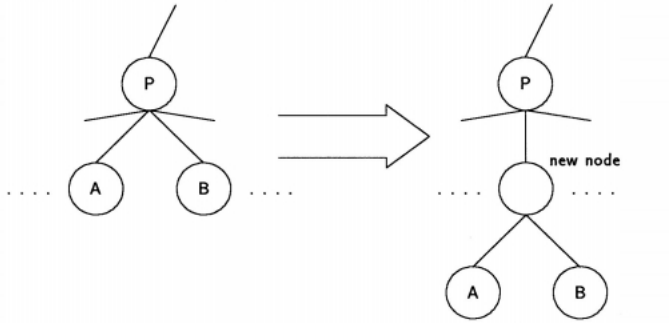

Figure 4. The effect of node merging.

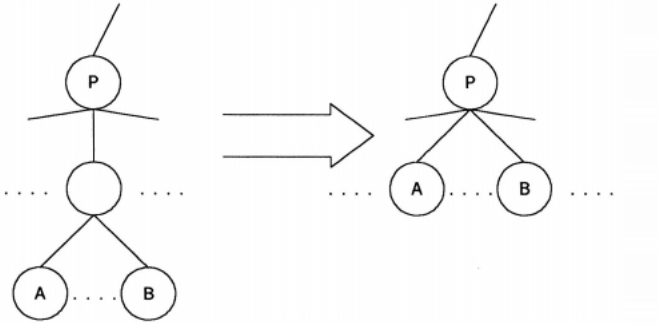

Figure 5. The effect of node splitting.

https://jair.org/index.php/jair/article/view/10162

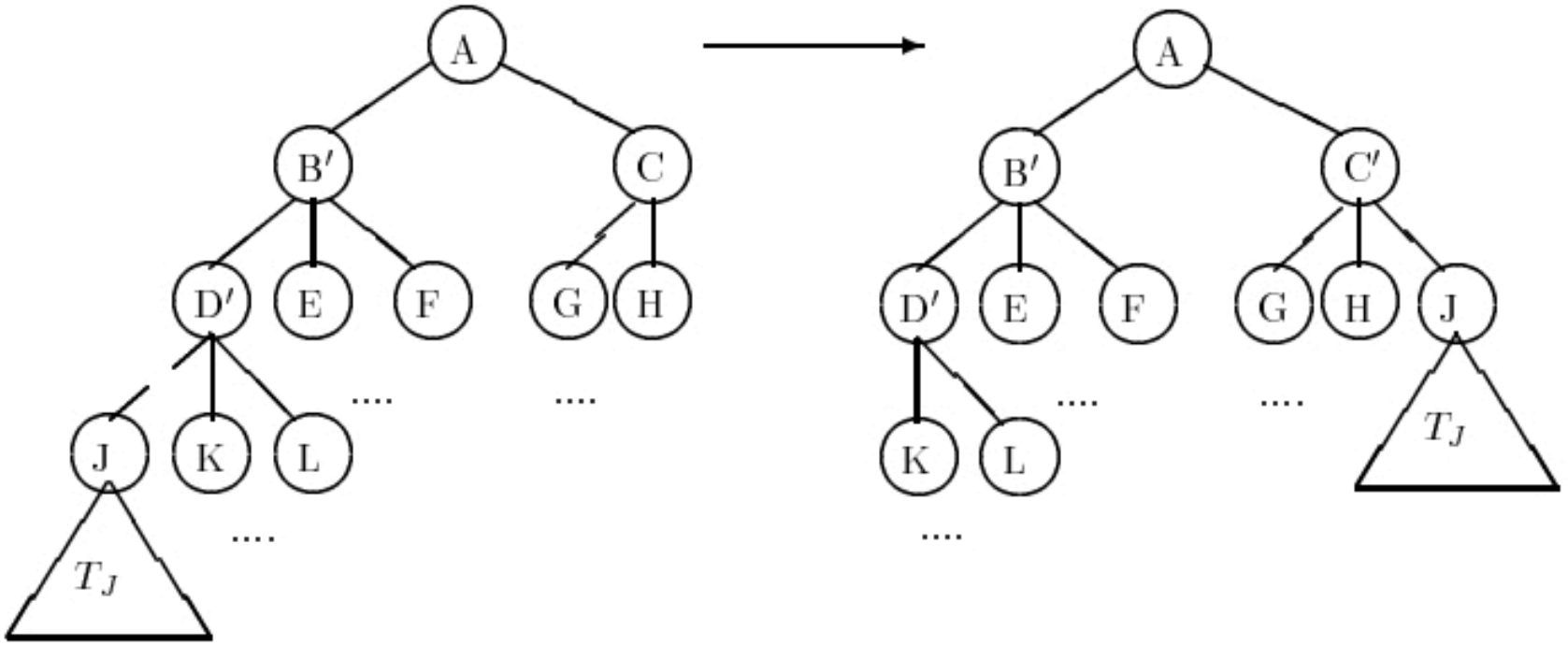

Figure 3: Hierarchical redistribution: the left subfigure indicates that cluster J has just

https://jair.org/index.php/jair/article/view/10162

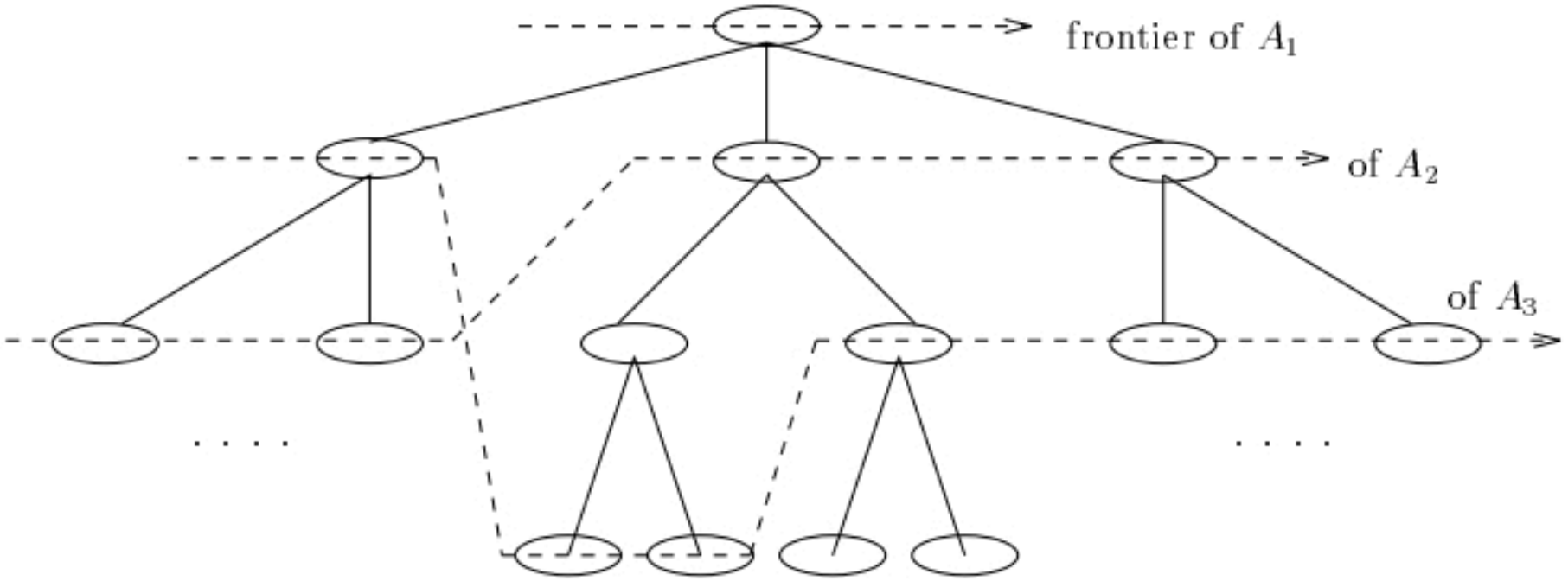

Figure 4: Frontiers for three variables in a hypothetical clustering. From Fisher (1995). Figure reproduced with permission from Proceedings of the First International Conference on Knowledge Discovery in Data Mining, Copyright ©1995 American Association for Artificial Intelligence.

Chapter 1 of Doug's Dissertation