A New Clustering Methodology for the Analysis of Sorted or Categorized Stimuli

Wayne S. Desarbo

Pennsylvania State University
University Park
State College, PA 16802

Kamel Jedidi

Columbia Business School - Marketing
New York, NY 10027

Michael D. Johnson

Stephen M. Ross School of Business at the University of Michigan
701 Tappan Street
Ann Arbor, MI 48109-1234

Abstract

This paper introduces a new stochastic clustering methodology devised for the analysis of categorized or sorted data. The methodology reveals consumers' common category knowledge as well as individual differences in using this knowledge for classifying brands in a designated product class. A small study involving the categorization of 28 brands of U.S. automobiles is presented where the results of the proposed methodology are compared with those obtained from KMEANS clustering. Finally, directions for future research are discussed.

A wide range of methods and procedures exists for the analysis of similarity relations among stimuli. These methods range from traditional spatial multidimensional scaling methods (Shepard 1962; Kruskal 1964) to network clustering methods (Johnson 1967; Shepard and Arabie 1979). The input to such procedures is typically some measure of similarity, whether obtained from paired comparison judgments, or derived from some similarity-based response (e.g., sorting data, classification errors, preference relations, or choice data). One such class of input proximity data is the sorting of stimuli (on the basis of their similarity) into homogeneous subsets or categories. Sorting tasks are employed in a wide range of

basic and applied research contexts. When subjects face a. large number of stimuli (Rao and Katz 1971) or have a limited ability to respond (Horton and Markm 1980), a sorting task may be gainfully employed.

Unfortunately, traditional methods may be ill-suited to this particular type of similarity-based response. Such methods may not capture the category perceptions that underlie the sorting task or the individual differences involved. Our goal is to describe a new clustering methodology designed specifically for the analysis of categorized or sorted stimuli. The methodology describes consumers' common category knowledge, the graded structure of these common categories, and individual differences in the use of the categories in a sorting task. We first describe sorting tasks, how they reveal category perceptions, and the proposed methodology. An illustration of the methodology is then presented with respect to a sort of 28 major U.S. automobile nameplates.

1. Sorting tasks and category knowledge

Sorting tasks provide researchers with a valuable alternative to traditional paired comparison similarity judgments. They are particularly well-suited to studies involving large stimulus sets where exhaustive paired comparison judgments are infeasible (Rao and Katz 1971). Asking subjects to provide excessive paired comparisons may result in fatigue and alter the basis of their judgments (Johnson, Lehmann, and Horne 1990). Having subjects sort the stimuli into piles of similar alternatives greatly simplifies the subjects' task.

Importantly, a sorting task should reveal consumers' product category knowledge. These product categories vary from relatively abstract, superordinate categories (e-g., modes of transportations, to intermediate level categories (e.g., sporty automobiles~~ to more concrete, subordinate level groupings (e.g., expensive Japanese sports cars). Rosch (1975) and Mervis and Rosch (1981) define a category as a group of items that are treated or labeled as equal even though they may be discriminated on the basis of their attributes. Likewise, a consumer who classifies brands into the same pile perceives these brands as relatively equal with respect to some criteria.

Individuals may differ in their perception and use of subordinate level distinctions. At the same time, perceptions should be more consistent across individuals at more intermediate or basic levels of categorization (Rosch 1977). Put differently, there are levels of category knowledge that are relatively universal for a given culture and are typically the first perceptual distinctions made in a sorting task. For example, although consumers may perceive quite different subcategories of sporty cars they should be more consistent in their perception of sporty versus luxury cars. When administered across subjects, a sorting task should reveal this common category knowledge.

Meanwhile, differences from subject to subject should help reveal the structure of these commonly held categories. Category membership is naturally graded. Some category members are relatively prototypical or good examples of a category (Rosch 1975; Medin and Smith 1984). Other members are more atypical of a category and thus may be poor examples of multiple categories. For example, an Escort GT may be both an economy car and a sporty car though not prototypical of either. A sorting task should reveal this graded, overlapping category structure in that prototypical members should be more consistently sorted together across consumers than are more atypical members.

Individuals also differ in their propensity to use common product categories (e.g., sporty versus luxury cars) as a basis for sorting stimuli. A consumer who is more familiar with sports cars may sort these cars into several personally relevant subcategories (e.g., expensive Japanese versus inexpensive Japanese sports cars) while lumping somewhat unfamiliar luxury cars into a single pile. Similarly, a consumer who is more familiar with luxury cars may form several luxury subcategories while lumping somewhat unfamiliar sports cars into a single pile. As a result, the weights placed on common categories in the final stimulus sorts may vary from consumer to consumer.

2. A new clustering methodology for sorted or categorized stimuli

A number of categorization studies in psychology and marketing utilize a sorting task procedure. Basically, the entire set of stimulus objects is placed on cards, one object per card. The subject is then asked to sort the cards into as many or few piles as he/she desires, congruent with the similarities and dissimilarities perceived among the stimuli. The subject is asked to place the cards (stimuli) that they view as similar in the same pile. Given such a data collection procedure, we are proposing a new methodology designed specifically to examine perceptual categorization. The data collected is "subject conditional" in the sense that subjects may utilize differing numbers of piles/groups, as well as differing numbers and types of objects per pile/group. The clustering methodology derives common category perceptions, the graded structure of the categories, and individual differences in the use of these perceptions in a sorting task.

2.1. The model

We ask consumers to subjectively group a designated set of brands into piles on whatever basis they desire, where the number of piles is determined by the individual consumer. Let:

i $= 1, \ldots, I$ consumers;

j,k $= 1, \ldots, N$ brands;

r $= 1, \ldots, R$ latent categories or clusters;

$$\delta_{ijk} = \begin{cases} 1 \text{ if consumer i places brands j and k in the same pile} \\ \qquad\qquad\qquad 0 \text{ else;} \end{cases}$$

$$s_{ijk} = \sum_{R=1}^{R} W_{ir} P_{jr} P_{kr}$$

= the predicted similarity (Carroll and Arabie 1983) between brand j and k perceived by

consumer I;

P_{jr} = the degree of membership of brand j in latent

category r, where $\sum_{R=1}^{R} P_{jr} = 1, \forall j, and \ 0 \leq P_{jr} \leq 1, \forall j, r;$

W_{ir} = the salience/importance weight consumer i has for latent category r. That is Wir indicates

the degree to which consumer i utilizes latent category r in his perceptual categorization process.

Expression (1) is similar in form to the Carroll and Arabie (1983) INDCLUS model devised for the analysis of three-way, two-mode metric proximity data, where expression (1) is used (with additive constants) to model similarity judgments. However, their model is a special case of our since Carroll and Arabie (1983) restrict P_{jr} to be 0 or 1. The INDCLUS methodology is also inappropriate for the analysis of binary data. Finally, as to be shown below, INDCLUS is deterministic while our specification is stochastic in nature.

We define a latent similarity function S_{ijk} as:

$$S_{ijk} = s_{ijk} + e_{ijk} \tag{2}$$

Where:

$$e_{ijk} = error \sim N\left(0, \sigma_i^2\right), \tag{3}$$

$$Cov\left(e_{ijk}, e'_{ijk}\right) = 0, \quad \forall i \neq i',$$

$$Cov\left(e_{ijk}, e_{irs}\right) = 0, \quad \forall (j \neq r, k \neq s).$$

Now, we assume that:

$$P\left(\delta_{ijk} = 1\right) = P\left(S_{ijk} \geq t_i\right), \tag{4}$$

Where t_i is a consumer specific threshold parameter; here, two brands j and k must be perceived as "similar enough" in order to be placed in the same pile or latent category.

Then:

$$P\left(\delta_{ijk} = 1\right) = P(e_{ijk} \geq t_i - \sum_{r=1}^{R} W_{ir} P_{jr} P_{kr}$$

$$= 1 - \varphi\left(\frac{t_i - \sum_r W_{ir} P_{jr} P_{kr}}{\sigma_i}\right)$$

$$\text{WLOG} = 1 - \varphi(t_i - \textstyle\sum_r W_{ir} P_{jr} P_{kr}) \tag{5}$$

Since σ_i can be absorbed into the numerator. Here φ represents the normal cdf function. Similarly,

$$P(S_{ijk} = 0) = \varphi(t_i - \textstyle\sum_r W_{ir} P_{jr} P_{kr}) \tag{6}$$

Given an independent sample of consumers whose categorization judgements are coded into these δ_{ijk} (one obtains upper/lower triangular-half binary matrices) and assuming independence over i, j, and k subscripts, one can form a likelihood expression:

$$L = \textstyle\prod_{i=1}^{J} \prod_{j>k}^{N} \prod (1 - \varphi(\cdot))^{\delta_{ijk}} \varphi(\cdot)^{(1-\delta_{ijk})}, \tag{7}$$

And the corresponding log-likelihood function:

$$Ln\,L = \textstyle\sum_{i=l}^{l} \sum_{j>k}^{N} \sum [\delta_{ijk} \ln(1 - \varphi(\cdot)) + (1 - \delta_{ijk}) ln\varphi(\cdot)] \tag{8}$$

Where:

$$\varphi(\cdot) = \varphi(t_i - \textstyle\sum_r W_{ir} P_{jr} P_{kr}) \tag{9}$$

Thus given the $\Delta = \left((\delta_{ijk})\right)$ binary data and R, we estimate $W = ((W_{ir}))$, $P = ((P_{jr}))$, and $t = (t_i)$ to maximize L or Ln L above. Here, W indicates how consumers weigh the derived latent categories in terms of their own categorization judgements. P indicates the membership of the N brands into R latent categories or clusters where fractional membership is allowed. And, t contains threshold information as to how find discriminations are made between brands for different consumers.

We utilize the conjugate gradient with automatic restarts procedure (Powell 1977) for parameter estimation. We impose the following constraints in the estimation: i) $W_{ir} \geq 0$, ii) $t_i \geq 0$, iii) $P_{jr} \geq 1$, and iv) $\sum_{r=1}^{R} P_{jr} = 1$. The non-negativity constraints are implemented by substituting squared quantities for these parameters (e.g., substituting W_{ir}^2 and t_i^2 for their respective unsquared likelihood expression - See Gill, Murray, and Wright 1981, pp. 268-269). The constraints on P_{jr} in (iii) and (iv) above are operationalized by reparametrizing:

$$P_{jr} = \frac{\theta_{jr}^2}{\sum_{r=1}^{R} \theta_{jr}^2}$$

and estimating θ_{jr} in order to avoid use of a constrained estimation procedure involving additional numerical complexity. An alternating conditional maximum likelihood (ACML) procedure was developed for the estimation of the model's parameters (see DeSarbo, Jedidi, and Steckel 1991).

3. Application

3.1. Study description

A small pilot study was conducted among a screened sample of potential consumers who stated they would be purchasing an automobile within the next six months ("intenders"). The purpose of this, pilot study was to examine the basis of how these potential consumers perceive and categorize domestic U.S. manufactured automobiles. Initially, an exhaustive list of 78 domestic automobile "nameplates" (excluding vans and trucks) were selected for study from the 1989 Consumer Reports New Car Buying Guide. Given that many of these nameplates were similar to other nameplates produced by the same manufacturer (e.g., Ford Escort - Mercury Lynx), this list was reduced down to 28. Table 1 presents this final list of the 28 nameplates included in this study.

Each of ten "intenders" was presented a randomized stack of 28 cards with the picture, name, and manufacturer of each automobile. The respondent was asked to sort these cards into as many or as few piles of similar automobiles as he/she thought necessary to reflect his/her perceptions.

3.2. Traditional analyses

As a means of comparison, six traditional clustering analyses were performed on the aggregated (over consumers) sorting frequency data: single linkage, complete linkage, average linkage, the centroid method, median linkage, and Ward's (1963) method. As one might expect, many of the resulting

Table 1. U.S. automobiles for the sorting task

1. Ford Escort	15. Plymouth Acclaim
2. Ford Festiva	16. Cadillac Seville
3. Geo Prizm	17. Chrysler Fifth Ave.
4. Plymouth Horizon	18. Chrysler LeBaron
5. Pontiac LeMans	19. Eagle Premier
6. Chevrolet Corvette	20. Ford Taurus
7. Ford Mustang	21. Ford Thunderbird
8. Ford Probe	22. Lincoln Continental
9. Plymouth Laser	23. Oldsmobile Cutlass
10. Pontiac Firebird	24. Pontiac 6000
11. Buick Skyhawk	25. Buick Le Sabre
12. Chevrolet Cavalier	26. Cadillac De Ville/Fleetwood
13. Dodge Aries	27. Lincoln Town Car
14. Ford Tempo	28. Pontiac Bonneville

dendrograms were quite different with respect to shape as well as the clusters formed. In addition, Spath's (1980) KMEANS procedure, which operates on dissimilarities, was also performed and will be utilized here to illustrate the results obtained from traditional methods. Table 2 presents the summary results for R=2, . . .,8 clusters. Based upon a visual inspection of the loss function, the sum-of-squared distances

criterion (SSDC), one might select either R = 4 or 6 clusters, given a leveling off of this SSDC at R = 5 and 7. The R = 4 solution renders the following classification:

Luxury: Lincoln Town Car
Chrysler 5th Avenue
Cadillac DeVille
Cadillac SeVille
Lincoln Continental

Sporty: Chevy Corvette
Pontiac Firebird
Ford Mustang
Ford Taurus
Plymouth Laser
Ford Probe
Pontiac Bonneville
Olds Cutlass
Ford Thunderbird

Mid-Size: Eagle Premier
Dodge Aries
Chrysler LeBaron
Plymouth Acclaim
Pontiac 6000
Buick Skyhawk
Chevy Cavalier
Ford Tempo
Buick LeSabre

Compact: Geo Prism
Ford Escort
Plymouth Horizon
Ford Festiva
Pontiac Lemans

These KMEANS results appear to describe four groups of nameplates: luxury, sporty, midsize, and compact. Most would not argue with such a classification, except perhaps with respect to several of the nameplates classified in the sporty category such as the Ford Taurus, the Pontiac Bonneville, and the Oldsmobile Cutlass which might be arguably classified in the midsize category. The corresponding six cluster solution (see Table 2) retains the luxury group intact and creates a more consistent set of five sporty nameplates. However, there is some difficulty in describing the remaining four categories. Note that these KMEANS solutions were different from those obtained with the six hierarchical clustering methods.

Thus, different clustering procedures render somewhat different portrayals of the categorization process among this small set of intenders. In addition, no information concerning individual differences in categorization is provided from such methods, given the use of aggregated frequency data. Finally, all of these clustering solutions fail to capture either the prototypicality of the automobiles within clusters or potential multiple category membership of any of the nameplates.

Table 2. KMEANS cluster assignments

Automobile nameplates:

Number of clusters:	Ford Escort	Ford Festiva	Geo Prizm	Plymouth Horizon	Pontiac LeMans	Chevy Corvette	Ford Mustang	Ford Probe	Plymouth Laser	Pontiac Firebird	Buick Skyhawk	Chevy Cavalier	Dodge Aries	Ford Tempo	Plymouth Acclaim	Cadillac Seville	Chrysler 5th Ave	Chrysler LeBaron	Eagle Primier	Ford Taurus	Ford Thunderbird	Olds Cutlass	Lincoln Continental	Pontiac 6000	Buick Le Sabre	Cadillac De Ville	Lincoln Town Car	Pontiac Bonneville
R = 2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	2	2	2	2	1	2	2	2	1	1	2
R = 3	3	3	3	3	3	2	2	2	2	2	3	3	3	3	3	1	1	3	3	3	1	1	1	3	3	1	1	1
R = 4	4	4	4	4	4	1	1	1	1	1	3	3	3	3	3	2	2	3	3	1	1	2	1	3	3	2	2	1
R = 5	4	4	4	4	4	5	5	5	5	5	4	4	4	4	4	1	2	4	4	3	3	1	3	4	3	2	2	3
R = 6	6	5	6	5	5	2	2	2	2	2	6	6	1	1	1	4	4	1	1	3	3	4	3	1	3	4	4	3
R = 7	5	5	5	5	5	3	2	2	2	3	6	6	4	6	4	1	1	4	4	7	7	1	7	4	7	1	1	7
R = 8	5	5	5	5	5	2	1	1	1	2	7	7	8	7	8	6	3	8	8	4	4	6	4	8	4	3	3	4

Goodness-of-fit statistics

# of clusters (R)	2	3	4	5	6	7	8
SSDC	8.75	7.70	6.44	6.31	4.93	4.41	4.20

SSDC: Sum of Squared Distance Criterion

8

3.3 Analysis with the proposed methodology

An analysis of 2-5 clusters or latent categories (as referred to here) was performed with the proposed methodology. Table 3 presents the various goodness-of-fit measures for these solutions. These various measures appear to indicate that the R = 4 category solution appears "best" in parsimoniously describing the structure in these data.

Table 4 presents the P matrix for these 28 nameplates for the four derived categories. Upon examination of the model values by nameplate, we see evidence of the compact, sport, midsize, and luxury categories similar to that from the KMEANS clustering. However, additional information is to be gained from a careful examination of these "probabilities of membership." An asterisk has been placed in each category (column) to denote the nameplate with the highest probability of membership, i.e., the most prototypical element. For the compact category, Ford Escort (.765) appears to be the most prototypical nameplate. This is not surprising given the traditional popularity of this particular automobile and the large amount of advertising Ford commits to this brand. For the sporty category, the Ford Mustang (.739) appears to be the most prototypical nameplate in the list. For the luxury category, the Lincoln Town Car and the Chrysler Fifth Avenue share the highest values (.743) as the most prototypical nameplates. Note an interesting finding for the midsize automobiles - the highest value in this column is substantially smaller (584 for Eagle Premier) than those for the previous three categories. This is consistent with recent U.S. automobile manufacturer positioning strategies which downplay the "midsize" Label due to the relatively large number of competitive automobiles in this category.

Another interesting aspect of the probabilities (P_{jr}) in Table 4 is that they allow for fractional membership in multiple categories (i.e., fuzzy categories). For example, the Chevy Corvette has fractional membership in both the sporty and luxury car categories. Similarly, the Buick Skyhawk is a fractional member of both the compact and midsize categories. As another example, Ford Thunderbird has almost the same degree of membership in the sporty category as in the midsize category, reflecting the previous history of the automobile as well as its mature, sporty positioning. Note, multiple category membership may also capture a lack of familiarity, as might be the case for the relatively new Plymouth Laser and Geo Prism.

Table 3. Summary goodness-of-fit statistics

Number of categories (R)	D.F.	Ln L	Pbc	Phi	Match	A.I.C.
2	58	− 1824.59	.306	.171	.773	3765.17
3	96	− 1557.76	.482	.387	.817	3307.52
4	134	− 1279.96	.629	.557	.860	2827.91*
5	172	− 1257.57	.648	.577	.865	2859.13

*Minimum A.I.C.

Table 4. The four category solution

	Compact	Sporty	Midsize	Luxury
1. Ford Escort	.765*	.168	.060	.007
2. Ford Festiva	.602	.069	.132	.197
3. Geo Prizm	.437	.128	.259	.176
4. Plymouth Horizon	.602	.069	.132	.197
5. Pontiac Lemans	.450	.108	.255	.187
6. Chevy Corvette	.132	.497	.064	.308
7. Ford Mustang	.107	.739*	.121	.034
8. Ford Probe	.156	.507	.299	.038
9. Plymouth Laser	.161	.473	.326	.040
10. Pontiac Firebird	.131	.472	.165	.232
11. Buick Skyhawk	.333	.110	.489	.067
12. Chevy Cavalier	.329	.175	.446	.050
13. Dodge Aires	.300	.076	.444	.179
14. Ford Tempo	.256	.184	.526	.035
15. Plymouth Acclaim	.309	.077	.439	.175
16. Cadillac Seville	.056	.107	.252	.585
17. Chrysler 5th Ave.	.058	.091	.108	.743*
18. Chrysler LeBaron	.153	.133	.489	.225
19. Eagle Premier	.133	.235	.584	.049
20. Ford Taurus	.072	.346	.546	.037
21. Ford T-Bird	.067	.433	.445	.056
22. Lincoln Continental	.056	.103	.142	.699
23. Olds Cutlass	.083	.177	.448	.292
24. Pontiac 6000	.240	.058	.424	.277
25. Buick LeSabre	.086	.140	.498	.276
26. Cadillac DeVille	.056	.095	.248	.608
27. Lincoln Town Car	.057	.090	.110	.743*
28. Pontiac Bonneville	.083	.183	.401	.332

*Denotes most prototypical nameplate

Finally, Table 5 presents the number of groups sorted by each intender, the normalized W matrix, and the ti threshold coefficients reflecting individual differences among the ten intenders. Intender 2 appears to stress the compact category in his/her categorization scheme, whereas Intender 9 appears to find the luxury category more salient. Other intenders tend to possess higher variance among their set of weights across the four categories. Rougher distinctions among the 28 nameplates appear to be made by Intenders 6 and 9 whose threshold coefficients are largest among the small sample of intenders. Thus, this

methodology renders information concerning the gradient structure of a categorization scheme as well as individual differences in perception.

4. Conclusion

We have presented both a conceptualization of stimulus sorting and a mathematical model for the analysis of sorted stimuli based on this conceptualization. The technical details of the stochastic, individual differences clustering methodology have been provided, as well as a small illustration concerning automobile nameplate sorts. We have shown how this methodology can derive a latent categorization scheme among a set of subjects, including a quantification of the nature of individual differences - an aspect lacking in most traditional clustering techniques.

Future research should be pursued along a number of lines. Rigorous Monte Carlo testing of the proposed methodology should be undertaken to examine the performance of the estimation algorithm as a number of data, model, and error factors are experimentally manipulated. Algorithm performance could be measured in terms of computational effort (e.g., CPU time), parameter recovery (e.g., compare B versus $), and overall goodness-of-fit (e.g., the simple matching coefficient). Generalizations of this methodology to multidimensional scaling can be made where P would represent a continuous space in R dimensions with no constraints imposed on the P_{jr} elements, Finally, more extensive applications should be attempted utilizing the sorting data collection method and this stochastic clustering methodology.

References

Akaike, H. (1374), "A New Look at Statistical Model ldentification," IEEE Transactions on Automatic Control, AC-19 716-723.

Carroll, J. D., and Arabic, I'. (1983). "INDCLUS: An Individual Differences GeneraIization of the ADCLUS Model and the MAPCLUS Algorithm," *Psychometrik*a 48, 157-169.

DeSarbo, W. S., Steckel, J., and Jedidi, K. (1991). "MICROSCALE: A Stochastic Multidimensional Scaling Procedure for the Empirical Determination of Convex Indifference Curves of Preference/ Choice Analysis," *Pyschometrika*, forthcoming.

Gill, P. E., Murray, W., and Wright, M. H. (1981). *Practical Optimization*, Orlando, FL: Academic Press.

Horton, M. S., and Markman, E. M. (1980). "Developmental Differences in the Acquisition of Basic and Superordinate Categories," Child Development 51, 708-719.

Howard, J. A. (1989). Consumer Behavior in Marketing Strategy, Englewood Cliffs, NJ: Prentice Hall.

Johnson, M. D., Lehmann, D. R., and Horne, D. R. 11990). "The effects of fatigue on judgments of interproduct similarity, "*International Journal of Research in Marketing*" 7, 35-43.

Johnson, S. C. (1967). "Hierarchical Clustering Schemes," *Psychometrika* 32, 241-254.

Kruskal, J. B. (1964). "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," Psychometrika 29, I-27.

Medin, D. L., and Smith, E, E. (1984). "Concepts and Concept Formation," A~~~u~ Review of Psychology 35, 113-138.

Mervis, C. B., and Rosch, E. (1981). '*Categorization of Natural Objects," *Annual Review of Psychology* 32, 89-l IS.

Powell, M. J. D. (1977). "Restart Procedures far the Conjugate Gradient Method," *Mathematical Programming* 12,241-254.

Rao, V. R., and Katz, R. (1971). "Alternative Multidimensional Scaling Methods for Large Stimulus Sets," *Journal of Marketing Research* 8,488-494.

Rosch, E. (175). "Cognitive Representation of Semantic Categories*," Journal of Experimental Psychology: General* 104, 192-233.

Rosch, E. (1977). "Human Categorization," in W, Warren (ed.), Studies in Cross-Cultural Psychology: Volume 1, London: Academic Press, l-49.

Sattath, S., and Tversky, A. (1977). "Additive Similarity Trees," *Psychometr*ika 42, 319-345.

Shepard, R. N. (1962). '"The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function, X and II," *Psychometrika* 27, 125-140 and 219-246.

Shepard, R. N., and Arabie, P. (1979). "Additive Clustering: Representation of Similarities as Combinations of Discrete Overlapping Properties, Psychological Review 86, 87-123.

Späth H. (1980). Cluster Analysis Algorithms, New York, NY: Wiley & Sons.

Ward, J. H. (1963). "Hierarchical Grouping to Optimize an Objective Function," Journal of the

    American Statistical Association 58, 236-244.