suspicious and 13 as positive under U.S. Department of Agriculture regulations for unvaccinated cattle[1,2].

Unfortunately, only 34 swine serum samples were tested in this survey. Twelve per cent (four samples) of these were found to have titres.

The 153 sheep and goat serum samples showed no evidence of agglutinating antibodies for *Brucella*.

Although the cattle population was the only one in the sample which was large enough to allow any estimation of the incidence of brucellosis infection, the presence of reactors does establish the fact that certain other animals must have been infected. The lack of titres in the smaller ruminants would seem to indicate that the sample of these species has not been infected. Larger samplings of these animals would obviously do much to establish firmer conclusions concerning the actual incidence of brucellosis in these animals.

Table 1. *Brucella* SERUM AGGLUTINATION TITRES IN EASTERN NIGERIAN LIVESTOCK (1962–1963)

	Neg. titres +1 : 25 ±1 : 50		Suspect +1 : 50 ±1 : 100		Positive +1 : 100 or higher		Animals without titres		Animals with titres		Total animals tested
	No.	%	No.	%	No.	%	No.	%	No.	%	No.
Cattle	40	3·7	43	4·0	13	1·2	974	91·1	96	8·9	1,070
Goats	—	—	—	—	—	—	110	100	—	—	110
Sheep	—	—	—	—	—	—	43	100	—	—	43
Swine	2	6·7	2	6·7	—	—	30	88	4	12	34
Totals	42	3·3	45	3·6	13	1·0	1,157	92	100	7·9	1,257

In swine the four animals that had titres showed weak agglutination of the stained antigen. This tends to suggest the possibility of the presence of a strain of *Brucella* other than *B. abortus*. This possibility is supported by the fact that the swine had had no contact with cattle—at least since being in Nigeria.

The implication for public health of the presence of infected livestock is important but does not have quite the same significance as would be the case in Western Europe or North America. The large majority of the local population of Eastern Nigeria does not drink milk except that which comes from a can or the mother's breast. They often dwell with their animals, however, and the meat is sold in the markets immediately after slaughter, increasing the possibility of human contact with the organism.

Although brucellosis has not been reported officially from the Northern region, the presence of titres in the animals (first herd tested) recently imported from this region and the Cameroon Republic strongly suggests that the infection is present in these areas. The presence of the disease in these areas is of greater importance than in Eastern Nigeria because the local populace drinks fresh milk.

JOSEPH W. ADAMS
J. MCKAY

University of Nigeria,
Nsukka.

[1] Hagen, W. A., and Brunner, D. W., *The Infectious Diseases of Domestic Animals*, third ed., 251 (Comstock Publishing Company, Ithaca, New York, 1957).
[2] Hull, Thomas C., *Diseases Transmitted from Animals to Man*, fifth ed., 143 (C. C. Thomas, Springfield, Illinois, 1963).
[3] Lombardini, F., personal communication (1962).

GENERAL

A Generalized Sorting Strategy for Computer Classifications

AGGLOMERATIVE hierarchical methods of computer classification all begin by calculating distance-measures between elements. The hierarchy is then generated by subjecting these measures to a sorting-strategy, which depends essentially on the definition of a distance-measure between groups of elements. In *nearest-neighbour* sorting, this is defined as the distance between the closest pair of elements, one in each group. Macnaughton-Smith has pointed out that much more intense clustering can be produced by taking the most remote pair of elements (*furthest-neighbour* sorting). In *group-average* sorting[1] the distance is defined as the mean of all between-group inter-element distances; in *centroid* sorting it is the distance between group centroids, defined by a conventional Euclidean model. In *median*[2] sorting the distance of a third group from two which have just fused depends on the previous three inter-group distances in the manner of Apollonius's theorem. Although the earlier of these strategies have received some comparative assessment[1,3–5] no attempt seems to have been made to generalize them into a single system. As a result, quite different computer strategies have commonly been used, necessitating a separate computer program for each.

We shall now erect a general system of which all these are special cases. What follows applies rigorously only to measures with the general properties of a Euclidean distance; other measures, for example, the correlation coefficient, require further investigation. The system does not incorporate the method of "information analysis"[5,6], which uses a measure fundamentally different from all others currently in use.

We first assume three groups (h), (i) and (j), containing n_h, n_i and n_j elements respectively and with inter-group distances already defined as d_{hi}, d_{hj}, d_{ij}. We further assume that the smallest of all distances still to be considered is d_{ij}, so that (i) and (j) fuse to form a new group (k), with $n_k(=n_i+n_j)$ elements. The problem is solved if we can always express d_{hk} in terms of any or all of the quantities already defined; they are all, of necessity, known at the moment of fusion. We assume the linear relation

$$d_{hk} = \alpha_i d_{hi} + \alpha_j d_{hj} + \beta d_{ij} + \gamma|d_{hi} - d_{hj}|$$

where the parameters α_i, α_j, β and γ determine the nature of the sorting strategy. It can be shown that the five strategies already mentioned are obtained when the parameters take the following values:

Nearest-neighbour:	$\alpha_i = \alpha_j = +\frac{1}{2}$; $\beta = 0$; $\gamma = -\frac{1}{2}$
Furthest-neighbour:	$\alpha_i = \alpha_j = +\frac{1}{2}$; $\beta = 0$; $\gamma = +\frac{1}{2}$
Median:	$\alpha_i = \alpha_j = +\frac{1}{2}$; $\beta = -\frac{1}{4}$; $\gamma = 0$
Group-average:	$\alpha_i = n_i/n_k$; $\alpha_j = n_j/n_k$; $\beta = \gamma = 0$
Centroid:	$\alpha_i = n_i/n_k$; $\alpha_j = n_j/n_k$; $\beta = -\alpha_i\alpha_j$; $\gamma = 0$

All can therefore be easily provided as optional variants of a single computer program.

The system has other interesting properties. It is desirable that the measures associated with successive hierarchical fusions should be monotonic; nearest- and furthest-neighbour are monotonic by definition, but this is not always true of the strategies with $\gamma = 0$. For these it is in fact easy to show that the system is necessarily monotonic so long as $(\alpha_i + \alpha_j + \beta) \geqslant 1$. The equality holds in the case of group-average, but the condition is not satisfied by median or centroid. Our experience of median sorting is limited, but we have ample evidence[5] that reversals can be troublesome in centroid. It is, moreover, now possible to generate an infinite set of new sorting strategies. We have investigated the system governed by the quadruple constraint ($\alpha_i + \alpha_j + \beta = 1$; $\alpha_i = \alpha_j$; $\beta < 1$; $\gamma = 0$); this defines a set of strategies such that, as β falls from positive to large negative values, the hierarchy changes from an almost completely "chained" system to one with increasingly intense clustering. The extent of clustering is thus not an inherent property of data; a given set of data may now, by varying the parameters, be made to appear as sharply clustered as a user may desire.

G. N. LANCE
W. T. WILLIAMS

C.S.I.R.O. Computing Research Section,
Canberra, A.C.T., Australia.

[1] Sokal, R. R., and Michener, C. D., *Univ. Kansas Sci. Bull.*, **38**, 1409 (1958).
[2] Gower, J. C., *Biometrics* (in the press).
[3] Sokal, R. R., and Sneath, P. H. A., *Principles of Numerical Taxonomy* (Freeman, San Francisco and London, 1963).
[4] Williams, W. T., and Dale, M. B., *Adv. Bot. Res.*, **2**, 35 (1965).
[5] Williams, W. T., Lambert, J. M., and Lance, G. N., *J. Ecol.*, **54**, 427 (1966).
[6] Lance, G. N., and Williams, W. T., *Comp. J.*, **9**, 60 (1966).