

## Functional synonyms and environmental homologues: an empirical approach to guild delimitation

M. B. Dale

*Australian School of Environmental Studies, Griffith University, Nathan 4111, Qld, Australia.*  
Fax: +61 7 3870 5681. E-mail: m.dale@mailbox.gu.edu.au

**Keywords:** Dissimilarity, Function, Functional type, Guilds, Synonymity.

**Abstract:** This paper presents a procedure for identifying guilds using species-in-stand data. Based on a linguistic analogy relating synonymity with functional equivalence, it develops a dissimilarity coefficient for clustering species which is suitable for measuring strength of synonymity. This coefficient combines two distinct aspects of synonymity - lack of co-occurrence of two species and similarity of context of other species where either species does occur. Synonymity is further restricted through a stand clustering to avoid confounding with environmental heterogeneity. The method is applied to data from *Eucalyptus* communities from sand dunes on North Stradbroke Island, Queensland. Some possible extensions are considered.

### Introduction

Finding groups of species has been recognised by Austin and Belbin (1982) as essentially different from grouping stands. Austin (1981), in criticising Dale and Anderson's (1973) two-parameter method, identified one of the sources of such difference with a contrast between diversity in stands and abundance (commonness and rarity) of species. Others, such as Bruelheide (2000), have concentrated on the notion of fidelity, which associates particular species with stands clusters.

Recently there has been emphasis on the functional role of plants and the notion of guilds. Guilds originally seem to have been introduced where a common resource is used differentially by several species, each specialised to exploit a particular range within the resource availability spectrum. Although the notion of a guild thus initially related to partitioning of resources, it has more recently been used in phytosociology to mean a group of species with common functional roles within the community, and hence a group likely to be strong competitors one with another. Thus, it might be expected that the presence of one member of a guild inhibits the occurrence of other members.

Guilds are also one candidate for the alphabet to be used in assembly rules for plant communities - each guild supplying one member to the common pool (Wilson and

Roxburgh 1994). This assumes that every community is formed by selection from the set of available guilds after they have been filtered through the local environment and subject to historical contingency (cf. Austin 1986 and Wisheu and Keddy 1992, for various models of vegetation). Some care is needed here, since a guild of epiphytes could be defined in a tropical rainforest but this is unlikely to be in any way analogous to a guild of epiphytic hepatics within a *Sphagnum* mass except that the plants are physiically dependent for support.

It is an assumption of competition theory that 2 species with essentially the same niche will not co-exist, although various mechanisms have been proposed which do allow co-existence, such as spatial patchiness combined with chance establishment (van der Maarel and Sykes 1993, Ågren and Fagerström 1984). However, identifying similarity of niche is difficult because it is not always obvious what the environmental limitations bounding the potential niche are. Indeed, Liebovitch (1995) has discussed a situation where boundaries are stochastic and not fixed at all!

Plants exist in a realised niche defined in part by competition with other species. Perhaps a more detailed classification of environmental factors as they impinge on plants, such as that of Lubásky (1969), could assist but there may be no environmental causes and constraints where patterns are formed as a result of vegetational proc-

esses (Dale and Hogeweg 1998, Dale 1999). In such cases the environmental differences are a *result* and not a *cause* of vegetation patterns, which makes the whole concept of a niche potentially circular. Does the environmental difference precede or succeed the vegetation difference? If it succeeds then the niche is autopoietic but the causal sequence will usually be difficult to determine.

Members of a guild are a cluster of species with, in part, similar niche requirements. If we could identify guilds without involving environmental attributes then the specification of niches could be obtained as a union of the realised niches of all guild members. But can we identify guilds? Some definitions are based on defining windows in space-time such as the vernal or aestival floras, others pick environmental factors which are associated with certain character traits deemed of importance *a priori*. Thus, Raunkiaer (1934) concentrated on the position of perennating buds, although later work on life forms, such as Ellenberg and Mueller-Dombois (1967) expanded the concept to other environmental concerns.

Such subjectivity in determining ‘important’ environmental factors is unsatisfactory. As Hogeweg and Hesper (1984) put it: “Mind-mediated methods are not suitable [for generating interesting alignments] because minds tend to get stuck in previously recognised patterns and never confine themselves to a set of explicit criteria but use implicit information which renders impossible an evaluation”. This is not to say that the introduction of expert domain knowledge is useless, only that, wherever possible, we should seek to avoid subjectivity.

Attempts have been made to cluster character traits to identify potential functional groups, again called guilds, starting with Lambert (1972). Most of these attempts use standard coefficients of similarity or dissimilarity which reflect co-occurrence of species. It is not obvious that such a procedure is either necessary or sufficient. Regarding the necessity, shared character traits can have phylogenetic rather than functional origins and associated traits may represent a response to more than one environmental problem. Nor can co-occurrence be significant because guild members should inhibit other members. As for sufficiency, Wilson (1999) points out that assembly rules are supposed to be dependent on competition between species and not on common tolerance for environmental factors. As a result, the co-occurrence or otherwise of two species in an environmentally heterogeneous habitat cannot taken as evidence of the existence of guilds. As noted earlier, environmental differentiation may be both a cause and an effect of species disjunctions. In this paper, I propose a method for identifying guilds numerically which overcomes some of these problems. As with Austin and

Belbin (1982), this involves a definition of a particular kind of dissimilarity but with some extensions to deal with the problem of environmental heterogeneity.

### A linguistic analogy and two kinds of cluster

I shall first look at a linguistic analogy that will enable us to establish means for identifying appropriate patterns. Clustering is a model-building procedure which offers a means of identifying structure in data and allowing recoding of data in order to compress it (see Wallace and Dowe 2000). With a text we seek to find means of replacing some parts of that text by symbols capturing the structure.

We can distinguish at least two coding methods. The first is ‘chunking’ in which we take a substring of the data (a group of letters or words) and replace it with a special symbol. In the sentence *Der Hund und der Bulle leben zusammen* we might recognise that the word ‘der’ is repeated, so that substring can be specially coded. The common data compression algorithms such as that of Ziv and Lempel (1977) use precisely this form of coding, examining the probabilities of particular sequences of symbols. Note that the meaning of the string ‘der’ is not important; der can also be a relative pronoun but we can still code it in the same way

Chunking is what most clustering algorithms applied to the usual dissimilarity measures in fact do - they discover chunks which are the same (approximately) and associate these with a code symbol - the cluster name. But this behaviour is a function of the properties of the dissimilarity measure used. If we change what we regard as ‘similar’ then chunking need not result and other forms of structure can be identified. As an example of clustering word co-occurrences, Deerwester et al. (1990) developed Latent Semantic Analysis, which is an efficient indexing scheme for full text documents, using singular value decomposition to first reduce the dimensionality.

The second alternative is a form of clustering where we group words according to their function, that is we derive grammatical classes. Clustering for this purpose, which I shall refer to as *C-clustering*, replaces stochastically equivalent *functional* classes by symbols, that is the members of a class share, approximately, the same function(s). Such classes need not share directly observable symbols but rather share other properties. Linguistically, they include synonyms, antonyms, specialisations, aspects, semantic classes such as domestic animals, phonetic classes such as vowels, graphemic classes such as punctuation and syntactic classes such as modal verbs. The important common feature of all these is that they all require a *contextual* definition. Their association with a

class is dependent on the way they function in language use and this function is exhibited in the words associated with them.

Members of C-clusters thus function in a similar role in a sentence. But such C-clustering is not apparent in a single sentence. Unlike chunking, C-clustering is obtained by examining many sentences and studying the context in which any specific word is used. To find clusters we must examine the interactions with the other components of a sentence. It is as a consequence of observing many sentences and examining the relationships between elements that we recognise the importance of C-clustering. In the sentence *Der Hund, die Katze und das Schwein leben zusammen* C-Clustering might identify a class based on 'Der, die, das' all definite articles in the nominative singular form. In both the exemplary sentences there is further C-clustering based on the concept of 'domestic animals'.

This form of C-clustering is illustrated by Lankhorst and Moddemeijer (1992) who provide an example of capturing combined grammatical and semantic structure, although their method relies entirely on the serial structure of text and is therefore not immediately applicable to vegetation data. Their similarity measure is based on the frequency with which 2 words occur in a particular order separated by some fixed number of other words - including none. The dissimilarity measure is a function of relationships between symbols.

The notion underlying guilds is that their elements have a similar role to play; that is species which function in a similar manner can be regarded as being in the same guild. This is also a definition of synonymy, so finding a set of synonyms is analogous to finding guilds. This is a specialised member of the stochastically equivalent classes and is not found by Lankhorst and Moddemeijer's method.

One possibility of finding synonyms is to use a measure of dissimilarity that is more appropriate to the question being considered. This could then be subject to a normal cluster analysis procedure to identify C-clusters. The linguistic analogy again contributes. If 2 species have the same functional role then they should be interchangeable just as 2 synonymous words are interchangeable; i.e., the species should be semantically equivalent when the semantics assigns the species a specific role in a vegetation. Seeking for synonymous species should provide a means of identifying guilds.<sup>1</sup>

Some words of caution are needed for these are only partial linguistic analogies and it is not clear how far they may be drawn. While the procedure given below does seem to successfully capture some aspects of synonymy, it does not claim to capture them all. As Berry et al. (1993) have pointed out, linguistic meaning can be expounded in many ways and not always in the same words. For example, if I search for information on 'automobiles' I would expect to find also information on 'cars' although the latter word may never appear. Similar difficulties could arise with species. A common environment may be expressed by several different suites of species, indeed historical contingency will make this very likely at a world scale. Words may also be polysemic, having multiple meanings so that while synonymous in one sense they may be unrelated in another; the case of 'der' has already been mentioned. Analogously, a species may have sufficiently differentiated subpopulations that it may serve as an indication of 2 somewhat different environments and synonymy with another species could be restricted to one of these. Michler and Arnold (1996, 1999) have shown that such differentiation can be identified in geographically distinct populations. Words and species are not identical, except in both being human constructions.

### Synonymy and functionality

How can we identify species whose roles are interchangeable. If we examine linguistic data, it is apparent that some words are more or less interchangeable. 'To ascend', 'to scale', 'to clamber up' and 'to climb', for example, all mean much the same; indeed we choose between them less on their meaning than on the pragmatic context of the communication. Exact replacement may not be attainable - it has been said that the only true synonym in English is the pair 'bramble' and 'blackberry'! But it is clear, I believe, that a degree of interchangeability is a property of many words and such words may still be termed synonyms. For example, 'to overcome' has some similarity with 'climbing', as does 'to conquer' but possibly not so much as 'ascending'.

Lewis et al. (1967) proposed a measure for this purpose in linguistics. They argue that synonyms can be characterised by two properties:

- that the 2 species being considered do not co-occur (except in special contexts such as dictionaries).
- that the 2 species being compared are found in a common context of *other* species.

<sup>1</sup> Recently, Das et al. (1998) have also proposed using other attributes to establish similarity but this does not seem to address the synonymy question.

Note that neither of these requirements specifically involves the serial order of words and hence they can be used with vegetation data.

We need to measure the degree with which these goals are met. Sampling fuzziness (Pillar and Orlóci 1993) makes it feasible for the 2 species to co-occur by chance while the contextual species may themselves have synonyms and therefore not always appear. Instead of trying to identify a single measure for both properties simultaneously, we shall instead measure both independently and then seek to combine them in some way.

### Problems

There are some residual problems of importance. The above definition, linguistically, is not restricted to defining synonyms, but would also include such categories as antonyms, specialisations and aspects. Whereas synonyms are alike in meaning, antonyms are opposites. I have not been able to provide any vegetational meaning to such opposites, though there may be such. Specialisation is subsumption where a general category, such as *metal* can be replaced by a specialisation such as *brass, tin* or *steel*. Such might be of interest in savannas if the class 'tree' is recognised, but could be replaced by individual species and it could also be used to represent nested phylogenetic relationships. Aspects are causal associates so that *heat* is an aspect of *friction*. Vegetational analogues of aspect relationships are not obvious, although epiphytism might be an aspect of available hosts.

Another problem is the nature of context - especially where a word can have 2 or more functions e.g., *film* (adjective, verb and noun) against *film* (noun and verb only) which pair also shows distinct meanings - either photographic or surface cover, especially texture. Obviously such overt semantic questions are inappropriate to vegetation studies, but could perhaps be given meaning in cases where a species was named *sensu lato* but in fact occurred in distinct forms. It might also be pertinent if a species changed its contextual meaning as the plant matured; seedling niches may not equate with adult niches. An important question concerns the impact of environmental heterogeneity. Some species have a wide tolerance, others a narrow one. If the environment is heterogenous, two narrow-tolerance species may not co-occur because of environmental differences, but those of wider tolerance could still be present and form a common context. Thus, the identification of synonymy with functional equivalence will only hold within a vegetation type. I shall discuss later how this problem might be avoided.

Finally, there is necessarily scale dependence. As part of the processing we require a limit on the associated 'words' which provide the context within which synonymy can be assessed. Written linguistic material generally provides obvious means of defining the context of co-occurrence in the form of units such as phrases, sentences, paragraphs, sections and chapters, each representing a different scale of observation. Oral linguistic material also provides sensible means of fragmenting the utterances though these do not necessarily correspond exactly with the written markers; possibilities include such as pauses, stress, intonation. Unlike the linguistic case, there is no obvious natural unit equivalent to the 'sentence' for vegetation. Instead the investigator must decide on the size and shape of the sampling area.

In the present example I have used quadrats. The difficulty here is that this is fixing the scale based on some subjective evaluation by the observing ecologist. It is not so much that the scale might be wrong but that no substantive justification for the selection is made. I suspect that faced with choosing a quadrat size, the observer makes some trade-off between the size of the visible local pattern units and the larger plants and the costs of data collection. But I would prefer to know what sort of considerations and constraints were actually used. Yu (1992) suggests that we have habitual ways of thinking, acting, judging and responding, which when taken together form our habitual domain, and sometimes we need to be nudged out of this mindset. Such habits can have far-reaching effects (see Carley and Palmquist 1992). I should add that I personally did not select the quadrat properties in the example.

### The synonymy coefficient

Lewis et al. (1967) describe several coefficients, of which I shall discuss only that they called  $G^{11}$ . This coefficient is defined below for binary (presence-absence) data.

Let  $n_{ij}$  = number of co-occurrences of species  $i$  and  $j$

Lewis et al. (1967) define a window to limit the range within which co-occurrence is possible. For vegetation data, occurrence in the same sampling unit is sufficient.

$n_i$  = number of occurrences of species  $i$ ,

$n_{..}$  = total number of stands,

$\mu_x$  = set of species which co-occur with species  $x$ ,

$r_x$  = set of species which occur at least  $p$  times with species  $x$ ,

$r_i^j$  = set of species which occurs at least  $p$  times with  $i$  and  $<p$  times with  $j$ , and

$C$ ,  $p$  and  $t$  are constants whose values are discussed later.

#### Co-occurrence component

To compare 2 species  $a$  and  $b$ , we examine all other species,  $k$ , and calculate the following:

Let  $h_k^{ab} = [ \text{if } n_{ak} > t \cap n_{bk} > t \text{ then } 0.5 \{ \max(0, n_{ak} - n_a n_k / n_{..}) + \max(0, n_{bk} - n_b n_k / n_{..}) \}$

else if  $n_{ak} > t \cap n_{bk} \leq t$  then  $\max(0, n_{ak} - n_a n_k / n_{..})$

else if  $n_{bk} > t \cap n_{ak} \leq t$  then  $-\max(0, n_{bk} - n_b n_k / n_{..})$

else 0 ]

Now define  $G_{a;b}^{10} = \sum_{k=1 \neq a,b}^u h_k^{ab}$ .

The basis of this measure is a difference between the observed co-occurrence frequency and the expected frequency. However, only certain of these differences are accumulated, specifically positive values of the difference which also meet the threshold conditions expressed in the preceding conditional expression.

#### Context component and combination

This is calculated from the following expression:

$$K_{ij} = C | r_i^j - r_j^i |$$

$r_i^j$  is obviously an asymmetric measure. Such asymmetry also appears in the first stage of Austin and Belbin's (1982) coefficient which might therefore be expected to show some of the properties of the coefficient developed here. In the usual (a,b,c,d) formulation for a 2 x 2 contingency table, Austin and Belbin (1982) used  $a/(a+b)$  and  $a/(a+c)$  as the asymmetric measure of similarity between 2 species. They then used the Manhattan metric to calculate a similarity measure between the rows of this asymmetric matrix. Ronkainen (1998, see also Das et al. 1998) examines this second stage in his discussion of 'external' similarity. With such a measure the similarity of 2 attributes is measured by the differences between their relationship with a probe set, which represents the context. In Austin and Belbin (1982), the probe set is all other species, but Ronkainen argues that this may not be an optimal choice. It might be expected, then, that the Austin and Belbin (1982) coefficient would show some relationship to the synonymity coefficient here defined, since it includes context in two different ways.

We now have our 2 components and combine them so that

$$G_{ij}^{11} = G_{ij}^{10} + K_{ij}$$

which can be converted to a dissimilarity measure if desired by using

$$1 - G_{ij}^{11} / \max(G_{ij}^{11})$$

We have adopted the simplest method of combination, unweighted summation, although this may not be optimal. If the two components are regarded as independent dimensions we might better use  $G_{ij}^{12} = \{(G_{ij}^{10})^2 * K_{ij}^2\}^{1/2}$ .

#### The choice of constants

There remains only the choice of constants. Without much experience, either direct or from the literature, I have accepted the default values suggested by Lewis et al. (1967) for the constants  $C$ ,  $t$  and  $p$ , being 2.4, 0 and 1 respectively.

$C$  is the weight given to contextual information compared to that given to co-occurrence information. The use of such a large value suggests that for the linguistic material studied in Lewis et al. (1967) the context is considerably more important than lack of co-occurrence.

To reduce noise impacts, Lewis et al. (1967) introduce a threshold  $t$  and regard any value of  $n_{ij} \leq t$  as being equivalent to zero. Whether this is a necessary transformation with phytosociological data is not clear, so for these analyses the threshold  $t$  has been set to 0. It would be desirable to base  $t$  on a probabilistic basis involving the expected co-occurrence perhaps along similar lines to those used by Grassle and Smith (1976).

Co-occurrence is again the domain of the constant  $p$ . A higher value would emphasise the commoner members of the contextual group. But while some experiment is certainly needed, it would not seem sensible to reduce the common context to ubiquitous species only! Selectivity in the context group is desirable to more closely bound the 'niche'.

#### Clustering and environmental heterogeneity

Clustering the dissimilarities can be carried out using any clustering procedure which is dissimilarity-based including various linkage methods. I have used Lance-Williams flexible sorting (1967) although it is possible that Podani's (1989) homogeneity methods or Sneath's (1966) r-linkage methods might be more appropriate. Wishart's (1969) mode analysis would provide a non-hierarchical method.

*Stand clustering.* There remains the usual problem of determining the appropriate number of clusters. For stands there are several possibilities and for the analyses here presented it was possible to use the Sandland-Young (1979) test since the quadrats were paired and the stands forming each pair can be regarded as replicates. This test examines the degree to which sample replicates are assigned to the same grouping, but it relies on all replicates of a sample being in the same group. With large numbers of replicates this results in a lack of power since no account is taken of the situation where, for example,  $(n-1)$  replicates are assigned to one group with one replicate assigned elsewhere. Methods described in Critchlow (1985) can avoid this situation. However, with only paired samples there is no benefit to be gained by using the more general approach.

*Species clustering.* It is more difficult to arrive at a suitable method for fixing the number of clusters of species. Obviously methods based on spatial coherency or replication are not available. One possibility is to derive a second dissimilarity matrix based on character traits of species and employ Critchlow's (loc. cit.) method to quantify the correlation between the two matrices. Methods relying on Gaussian distributions seem inappropriate, but Panayirci and Dubes (1983) and Glasbey (1987) provide procedures which might be applicable. In the analyses here presented I have relied on subjective choice.

An objective stopping rule is important. Simply clustering species does not overcome the problem of environmental heterogeneity and further analysis is required. To do this I propose that a clustering of the stands also be carried out to determine homogeneous vegetation types. Each species is then associated with one or more of the stand classes in which it is frequent or assigned to a null class if it is infrequent in all classes. If we have, say, 3 stand classes (A, B, C) a species will be associated with one or more classes, so that it can be labelled as being an A, B, C, AB, AC, CB or ABC depending on its relationships. These class assignments can be regarded as upper bound approximations to the realised niche widths, though admittedly somewhat crude estimates. For each group of species, we regard those which fall into the same class(es) as synonymous. Species in the same C-cluster which fall into different stand groups are regarded as representative of the same guild but under different environmental conditions.

There remains an underlying problem. Boerlijst and Hogeweg (1991) have shown that the dynamics of some chaotic processes can lead to the development of spatial patterns without environmental variation pre-existing. It is arguable, for example, that hummocks in bogs are an

example of this although experimental proof has yet to be provided; I have also seen somewhat similar 'double-spiral' patterns in grassland at Hochsheim, near Vienna. This will inevitably, in the course of time, lead to environmental differentiation and perhaps convergent evolution within the patches so formed. Is this to be regarded as environmentally mediated and hence to be disregarded in a search for guild structures? It is a function of the biota after all! Since I do not have temporal data, I shall perforce ignore such autopoietic pattern for the rest of the paper. However, it does illustrate that there is unlikely to be the neat separation of environmental and competitive aspects that Wilson (1999) envisages.

## Data and analyses

The data used were of species presence/absence in 225 quadrats placed regularly in woodland communities on North Stradbroke Island, a large sand mass east of Brisbane. The data are more fully described in Clifford and Specht (1978) and have been previously employed in studies of character traits by Dale et al. (1984). In this paper, the character information has not been used except as an interpretational aid, although synonymity might be a useful approach to identifying alternative responses to some environmental or disturbance conditions.

Importantly for our present purposes, these quadrats were collected in spatially contiguous pairs, each of 1 m<sup>2</sup> and because of this we can use the Sandland-Young (1979) test to provide an objective means of determining the requisite number of clusters in the stand analysis. I am assuming that environmental variation, and associated vegetation types, will possess spatial coherency, whereas competitive or facilitative effects will result in small patches without coherency at least on the scale of our investigation. The stand clusters will therefore represent different environments and this leaves open the possibility that a common function could be associated with different species in the different stand groups.

The quadrat data were subjected to a hierarchical classification using the Bray-Curtis (1957) coefficient of similarity, a SAHN clustering using the Lance-Williams (1967) flexible sorting algorithm with  $\beta = -0.25$ , followed by a Sandland-Young (1979) test to identify the number of clusters present. The coefficient of similarity is of a common general type. It is the ratio of the amount of overlap to the maximum possible amount of overlap, or how much information is shared by a pair of samples compared to how much could be shared. The maximal shared amount is determined for the specific pair being examined so the universe of discourse is local.

The Sandland-Young (1979) test identifies the number of classes by determining the degree to which replicate samples are placed in the same group. The basic notion is that if large numbers of replicates are being split between groups then we have likely produced too many clusters. The test examines the clusters to determine if the replicates (here pairs) are being segregated more than would be expected by chance. It is thus a measure of local spatial coherency. Of course, it is possible to test for the number of clusters in other ways if replicates are not available - Krishna-Iyer (1949), Critchlow (1985) and Wallace and Dowe (2000) all provide methods which would be applicable for the purpose of objectively determining the number of stand clusters.

The data were further subjected to an inverse analysis using the synonymity coefficient. No formal test for the number of groups was made here but 5 were recognised. They are shown in Table 2, together with their total presence in each of the 3 stand groups. We then examine where the most frequent occurrences of the species are situated within the 3 stand clusters.

In addition, two other inverse analyses were performed, one based on the Austin and Belbin (1982) approach, the other a straightforward application of the Lance and Williams (1968) divisive information analysis. The synonymity and other clusterings could be compared through a contingency table which indicates if there are grounds for assuming that similar results are being obtained by the several methods.

#### *Synonymous species and their traits*

Pillar (1999) has argued that trait-defined plant types are preferable to species and presented a procedure for identifying optimal traits from which the types are constructed. Here we can better proceed in the opposite direction since we identify species clusters which hopefully have a common function, and can use these to identify traits which are associated with such function. This may be preferable because there might be several alternative sets of traits which indicate the same function.

Some species-trait data were available, although very limited in range being dominated by seed properties. These data were clustered using Wallace and Dowe's (2000) Snob program and in addition the synonymity analysis clusters were imposed on the trait data to see if any obvious correlations were apparent. Snob implements a fuzzy non-hierarchical mixture separation procedure with an objective test for the number of clusters present based on the minimum message length principle (see Dale

2000, for a more extended description in an ecological context)

## **Results**

### *Stand grouping*

For the normal analysis of stands, the Sandland-Young Test indicated that 3 groups were present; Clifford and Specht (1978) had previously suggested that 3 or 4 communities were recognisable. Group A represents the ridge top, Group B east- (and to some degree north-) facing slopes and group C west- (and to some degree south-) facing slopes. This tripartite pattern reappeared in several other analyses using different clustering methods. Dale et al. (1984) showed the spatial distribution of the clusters.

### *Species grouping: comparison*

Inspection of the results suggested that 5 clusters were present. Comparisons of the several inverse analyses (Table 1) showed no strong relationship between the synonymity clusters and those of the alternatives. A G-test indicated probabilities of 0.1 and 0.5 with information and Austin and Belbin (1982) results, respectively. Clearly, whatever the synonymity analysis is doing, it is *not* reproducing the results obtained by the other methods. Nor does it seem that whatever context is captured by the Austin and Belbin (1982) coefficient is sufficient to introduce a relationship with the synonymity coefficient. The introduction of the criterion for lack of co-occurrence presumably has a large effect, especially as it is down-weighted by the value for constant *C* used.

### *The two-way table set intersection/Bray-Curtis*

The five species groups are shown in Table 2. We have still to remove environmentally mediated pattern, but it is clear that, in many cases, the species have their highest frequencies in different combinations of the stand groups. I have emphasised the peak frequencies, although this was a somewhat subjective selection in some cases.

In order to remove the environmental effects, I have proceeded as follows. By examining the frequency of each species in each of the stand group, we can identify those groups where it is of high frequency and hence an important component of the vegetation; these have been emphasised in Table 2. In the present example the results can be represented using a kind of Venn diagram showing the overlap of the stand groups and the species associated with all the possible combinations. The results are presented for 4 of the species groups in Figures 1-4. Group 4 has no very frequent species and has been ignored in the interpretation.

**Table 1.** Contingency tables for comparison of inverse analyses. Zero entries omitted.*1a DMINF and Synonymy*

	Synonymy					Total
DMINF	A	B	C	D	E	
A	3	5	1			9
B	3	3	2	1		10
C		2			1	3
D			2			2
E					1	1
Total	6	10	5	1	2	

*1b Austin-Belbin and Synonymy*

	Synonymy					Total
Austin-Belbin	A	B	C	D	E	
A						0
B	4	3	6			13
C	2	7	1		2	13
D				1		1
E						0
Total	6	10	7	1	2	27

**Table 2.** Synonymy species groups: species frequency in stand groups.

Species	RIDGETOP no. Stands = 86	EASTFACING no. Stands = 68	WESTFACING no. Stands = 71
GROUP 1			
<i>Austromyrtus dulcis</i>	15	66	55
<i>Homoranthus virgatus</i>	8	47	30
<i>Monotoca scoparia</i>	75	40	71
<i>Eucalyptus intermedia</i>	40	40	39
<i>Eucalyptus planchoniana</i>	5	30	30
<i>Baeckea linearis</i>	1	50	17
GROUP 2			
<i>Entolasia stricta</i>	54	45	26
<i>Banksia aemula</i>	70	66	63
<i>Leptospermum flavescens</i>	22	59	23
<i>Persoonia linearis</i>	49	43	27
<i>Strangea linearis</i>	22	64	52
<i>Smilax australis</i>	71	5	5
<i>Boronia rosmarinifolia</i>	38	62	64
<i>Xanthorrhoea johnsonii</i>	79	43	67
<i>Lomandra longifolia</i>	62	45	55
GROUP 3			
<i>Pteridium esculentum</i>	85	20	62
<i>Caustis recurvis</i>	6	61	17
<i>Woolsia pungens</i>	5	67	22
<i>Dillwynia pedunculata</i>	10	59	44
<i>Phyllota phyllicoides</i>	8	56	43
<i>Ricinocarpus pinifolia</i>	15	66	55
<i>Brachyloma daphnoides</i>	3	40	15
GROUP 4 no species frequently recorded			
GROUP 5			
<i>Casuarina torulosa</i>	50	2	5
<i>Themeda australis</i>	20	2	29



*Character traits and synonymity groupings.*

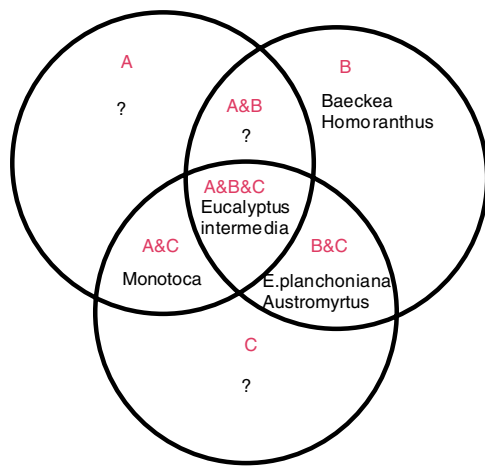
Life form categories are *not* associated with the groupings identified with these data; synonymous species, which presumably have a similar functional status, are apparently not correlated with life form within this small region. The groupings of species from the synonymity analysis were imposed on the character data (Table 3) and a minimal message length evaluation made. This indicated that the null hypothesis of a single cluster was acceptable. The 1-cluster message length = 243.7 nits, the 4 cluster message length 289 nits with a difference of 45.3 nits which indicates odds in favour of the null hypothesis of c.  $4 \times 10^{19}:1!$

The analysis does indicate that some shifts in the

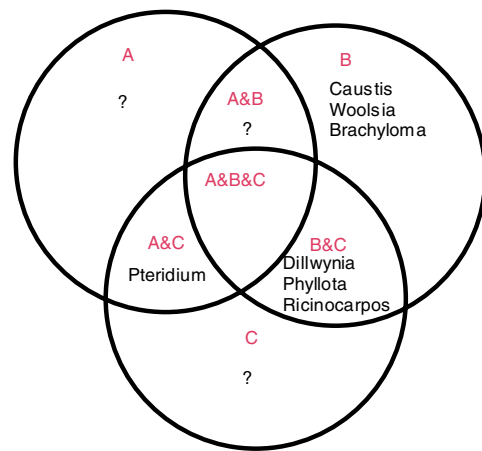
group composition were advantageous and these are shown in Table 4. These might strengthen the ‘myrtaceous’ qualities of the first group, but otherwise provides no further clarification of the nature of the groupings. The traits are not related to the synonymity groups. The small number of both species and traits makes the discovery of any structure extremely difficult, so perhaps the lack of relationship is not surprising even if we are ‘optimistic’ in our induction, as Fisher (1992) recommends.

**Discussion**

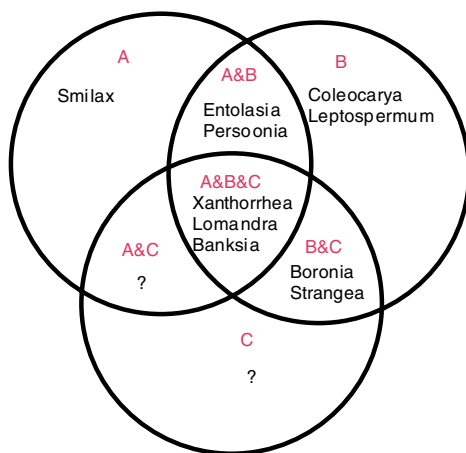
The combination of an appropriate dissimilarity measure and the coupling of normal and inverse analyses appears to succeed in identifying clusters. However,



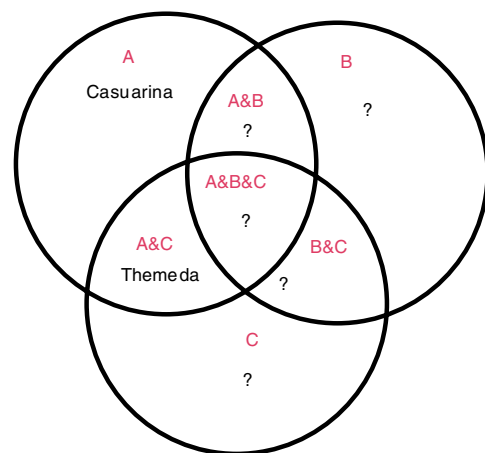
**Figure 1.** Group 1 frequent species and stand groups. Each circle represents one of the 3 stand groups. Species are placed in those groups or intersections of groups in which they have high frequency.



**Figure 3.** Group 3 frequent species and stand groups. See Fig. 1 for explanation.



**Figure 2.** Group 2 frequent species and stand groups. See Fig. 1 for explanation.



**Figure 4.** Group 5 frequent species and stand groups. See Fig. 1 for explanation.

**Table 3.** Character traits.

Trait	Code Alternatives
Leaves	simple; compound
Leaves	scale-like; not scale-like
Leaf Insertion	opposite; alternate; whorled
Leaf Base	Sheathing; not sheathing
Oil Glands	Present; Absent
Stipules	Present; Absent
Latex Ducts	Present; Absent
Root Nodules	Present; Absent
Endosperm	Present; Absent
Germination	phanerocotylar; cryptocotylar
Diaspore type	Seed; Spore
Number of Seeds	1 per fruit; many per fruit
Life Form	small tree; shrub; liane; hemipterophyte; geophyte
Mycorrhizae	Endo-; Ecto-; Absent
Leaf Area	0-2, 2-8, 16+ sq. cms.
Fruit	Dry; Fleshy
Pollination	animal; wind

Austin and Belbin's (1982) coefficient does not appear to be capturing the same relationship information. It is possible that altering the set of species used in defining the context could improve matters. It is possible, too, that the context required is variable and might be better captured in terms of neighbour relations than co-occurrence within quadrats of some fixed size. This needs further investigation.

High synonymity values between species in distinct communities should not be taken as indicative of guild structure but rather of shared or opposed responses to the physical environment. Such environmentally mediated patterns are not evidence for assembly rules (Wilson 1999) where patterns are due to inter-specific interactions. However, where the similarity is shared with species in the same stand group, then we can be confident of some functional equivalence. The stand groups distin-

guish environmental heterogeneity, so within clusters we can expect much greater environmental homogeneity and increasing importance of biotic interactions. The only subjectivity in the analysis is in the selection of important species using the frequency maxima and this could certainly be modified.

The assignment of species to stand types was somewhat *ad hoc*. Harter (1975ab) might provide a better method of determining the stands with which any species is associated. His underlying notion is that a species has 2 distributions, both Poisson. One is for casual occurrences, the other for stands where it is of significance - an 'elite' group. By identifying such elite groupings it might be possible to 'de-noise' the situation somewhat. Byers and Raftery (1997) have demonstrated the effectiveness of such an approach.

#### *Temporal changes in the guild membership*

The present analysis relies on data from a single snapshot, and does not include any information on the ontogeny of the species. It is reasonable that a plant will occupy different niches, and hence pertain to different guilds at various stages in its life cycle. Thus, it may be preferable to use character set types (Lambert 1972, Orłóci 1991, Pillar 1999) in place of species.

One problem with the use of character traits is that they can be the result of a common phylogeny rather than a local functionality. There is some limited evidence for

**Table 4.** Changes to group content indicated by trait analysis.

Species	From Group	To Group
<i>Monotoca scoparia</i>	1	2
<i>Coleocarya gracilis</i>	2	3
<i>Leptospermum flavescens</i>	2	1
<i>Boronia rosmarinifolia</i>	2	1
<i>Pteridium esculentum</i>	3	2
<i>Brachyloma daphnoides</i>	3	2
<i>Casuarina torulosa</i>	5	3
<i>Themeda australis</i>	5	2

this in the results, for one of the groups was primarily composed of members of a single family, the Myrtaceae. Yet other members of that family joined other groups so, whatever the phylogenetic constraints, they are by no means necessarily dominant. There can also be contingent effects, for Myrtaceae are a dominant component of the Australian flora and hence easily available members of any grouping. The introduction of exotic taxa might considerably change the composition of the guilds.

#### *The empty cells*

Why are there empty cells in the Venn diagrams? Two possible reasons can be advanced. First, the required niche may not be available in the vegetation type either because of environmental constraints or because the partition of the environment by other species leaves no gap in the specific location. The other alternative is that no suitable occupant species is available. If this latter alternative is true, then this provides a means of indicating available niches which can potentially be occupied by invading species. It is not obvious how these two situations can be differentiated without detailed specification of the niche boundaries for occupying species and for potential invaders. Our data are a snapshot of the vegetation. We only know that a species is present or absent but we have no knowledge of when it arrived, if present nor if it has arrived but later vanished, or could survive but has not yet arrived, if absent (cf. Anderson and Rann 1995).

#### *Using quantitative data*

Even for binary data the coefficients defined by Lewis et al. (1967) may not be a suitable measure. The authors themselves indicate some dissatisfaction with the coefficient they originally used and developed the  $G^{11}$  coefficient to overcome some of them. It is possible that still better alternatives can be found. For example, Biberman (1994) has also reported a context similarity measure which might prove more appropriate.

One outstanding problem is to expand the coefficient to accept ordered category and numeric data (cf. Heitjan 1993), these being the commonest forms of phytosociological data. Ordered category data can be coded in various ways as binary data, while numeric data could be quantised to give a similar result but it is desirable that the coefficient itself be generalised. I do not propose to investigate this here in any detail. Measures based on fuzzy sets (Roberts 1986, 1987) could substitute for the  $G^{10}$  component and various correlation measures based on subsets defined by single species presence might be manipulated to provide the context similarity measure.

#### *Multiple matrix interpretation*

The coefficient  $G_{ij}^{11}$  is a combination of two distinct measures, specifically  $G_{ij}^{10}$  which represents the co-occurrence excluding the specified pair, and  $K_{ij}$  which measures the commonality of context. It would be possible to maintain these as two separate matrices. Such measures have also been discussed by Tversky (1977), who distinguishes overlap and difference components of similarity. Simultaneous analysis of such multiple dissimilarity matrices has been discussed in Dale and Dale (1994) with special regard to the use of the Pareto alternative (Ferligoj and Batagelj 1992), the present case providing a 10<sup>th</sup> situation where multiple matrices are of potential interest. We might also look toward various Mantel statistics (Smouse et al. 1986) if we seek to relate dissimilarity to external factors.

In fact, most dissimilarity measures involve a combination of individual contributions from several species. This has been noted by Faith and Belbin (1986) and Godehardt and Hermann (1988), and earlier by Weir (1972). In most cases, the components are simply summed although Weir did consider possibilities of weighted summation and Ross et al. (1986) provided a weighted summation option in the TAXON 'MATWAD' program. In fact, combining dissimilarity components is in some ways analogous to voting on decisions, so that methods like those of Kacprzyk (1985, 1986; see also Fuller and Carlsson 1996) might prove useful.

Rather more important is the principle which has been invoked here. We have designed a similarity measure for a specific purpose by combining two different notions of similarity. A final merger into a single number may or may not be necessary. What is clear is that we can extend this methodology to other situations. For example, I have recently found some similar work by Gefeller and Pralle (1997) which introduces a similar notion for situations where multiple outcomes are possible in an experimental situation.

In vegetation studies one possibility is to introduce several definitions of context. If the species are divided into various subsets, then each could supply its own contextual information. This could be an advantage in studies of savannas where arboreal and herbaceous components may need to be separated. A similar separation might be appropriate with bog data, with the Bryophytes distinguished from other plants.

Alternatively, following Ronkainen (1998), several sets of species could be used to define different probe sets and distinct external similarities calculated. This would result in several dissimilarity matrices each reflecting dif-

ferent relationships between the various subsets. The exact vegetational implications of analysing such data remain unexplored, although it should be useful in examining inter-synusial relationships.

## Conclusion

I have shown that by appropriate definition of dissimilarity and by combining normal and inverse analysis, it is possible to derive clusters of species that can reasonably be expected to share a common functionality. Such species may be regarded as part of a guild. The example suggests that links to taxonomic and character trait syndromes may occur but they do not seem strong. By manipulating the multiple matrix approach, various other dissimilarity coefficients can be envisaged which should be useful in several ways. Improvements to the present procedure can be envisaged, but it does seem to provide a viable means of establishing guild structure without undue subjectivity encroaching.

**Acknowledgment:** To H. T. Clifford for assistance with the character trait descriptions.

## References

- Ågren, G. I. and T. Fagerström. 1984. Limiting dissimilarity in plants: randomness prevents exclusion of species with similar competitive abilities. *Oikos* 43: 369-375.
- Anderson, P. K. and B. B. Rann. 1995. A nonparametric test for comparing two samples where all observations are either right or left censored. *Biometrics* 51: 323-329.
- Austin, M. P. 1981. The role of certain diversity properties in vegetation classification. In: A. N. Gillison and D. J. Anderson (eds.), *Vegetation Classification in Australia*. CSIRO and ANU Press, Canberra pp. 125-140.
- Austin, M. P. 1986. The theoretical basis of vegetation science. *Trends Ecol. Evol.* 1: 161-164.
- Austin, M. P. and L. Belbin. 1982. A new approach to the species classification problem in floristic analysis. *Austral. J. Ecol.* 7:75-90.
- Berry, M. T. Do, G. O'Brien, V. Krishna and S. Varadhan. 1993. SVDPACKC (Version 1.0) User's Guide. Tech. Rep. ut-cs-93-194. Dept. Computer Science, University of Tennessee.
- Biberman, Y. 1994. A context similarity measure. In: Proc European Conf. Machine Learning ECML-94, Catalina, Italy. Springer, Berlin. pp. 49-67.
- Boerlijst, M. C. and P. Hogeweg. 1991. Spiral wave structure in prebiotic evolution: hypercycles stable against parasites. *Physica D* 48: 17-28.
- Bray, J. R. and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* 27: 325-349.
- Bruelheide, H. 2000. A new measure of fidelity and its application to defining species groups. *J. Veg. Sci.* 11:167-178.
- Byers, S. and A. E. Raftery. 1997. Nearest neighbour clutter removal for estimating features in spatial point processes. Tech. Rep. 305, Dept. Statistics, Univ. Washington.
- Carley, K. and M. Palmquist. 1992. Extracting, representing and analyzing mental models. *Social Forces* 70: 601-636.
- Clifford, H. T. and R. L. Specht. 1978. *The Vegetation of North Stradbroke Island*. University of Queensland Press, St. Lucia.
- Critchlow, D. E. 1985. *Metric Methods for Analyzing Partially Ranked Data*. Lecture Notes in Statistics 34, Springer-Verlag, Berlin.
- Dale, M. B. 1999. The dynamics of diversity: mixed strategy systems. *Coenoses* 13:105-113.
- Dale, M. B. 2000. Mt. Glorious revisited: secondary succession in subtropical rainforest. *Community Ecol.* 1: 181-194.
- Dale, M. B. and D. J. Anderson. 1973. Inosculate analysis of vegetation data. *Austral. J. Bot.* 21: 253-276.
- Dale, M. B., H. T. Clifford and D. R. Ross. 1984. Species, equivalence and morphological redescription. In: R. Colman, J. Covicovich and P. Davie (eds.), *Focus on Stradbroke; New Information on North Stradbroke Island and Nearby Areas, 1974-1984*, Boolarong Press and Stradbroke Island Management Committee, pp. 195-207.
- Dale, M. B. and P. T. Dale. 1994. Multiple dissimilarity matrices: sources and classification. *Coenoses* 9: 9-13.
- Dale, M. B. and P. Hogeweg. 1998. The dynamics of diversity: a cellular automaton approach. *Coenoses* 13:3-15.
- Das, G., H. Mannila and P. Ronkainen. 1998. Similarity of attributes by external probes. *KDD-98* pp. 23-29.
- Deerwester, S., S. T. Dumais, G. W. Furnas, T. K. Landauer and R. Harshman. 1990. Indexing by latent semantic analysis. *J. Amer. Soc. Inform. Sci.* 41: 391-407.
- Ellenberg, H. and Mueller-Dombois, D. 1967. A key to Raunkiaer plant life forms with revised subdivisions. *Ber. geobot. Inst ETH Stftg. Rübel, Zürich* 37: 56-74.
- Faith, D. P. and L. Belbin. 1986. Comparison of classifications using measures intermediate between metric dissimilarity and consensus similarity. *J. Classif.* 3:257-280.
- Ferligoj, A. and V. Batagelj. 1992. Direct multi-criterion clustering algorithms. *J. Classif.* 9:43-61.
- Fisher, D. 1992. Pessimistic and optimistic induction. Dept. Comput. Sci., Vanderbilt Univ. TR CS-92-12.
- Fuller, R. and C. Carlsson. 1996. Fuzzy multiple criterion decision making: recent developments. *Fuzzy Sets and Systems* 78: 139-153.
- Gefeller, O. and L. Pralle. 1997. NRCCS A nonparametric test for evaluating coherent alternatives in non-randomised studies. Proc Internatl. Conf. Non-randomized Comparative Clinical Studies, Heidelberg April 10-11, Symposium Publishing.
- Glasbey, C. A. 1987. Complete linkage as a stopping rule for single linkage. *J. Classif.* 4:103-109.
- Godehardt, E. and H. Hermann. 1988. Multigraphs as a tool for numerical classification. In: H. H. Bock (ed.) *Classification and Related Methods of Data Analysis*, Elsevier Science, North Holland. pp. 19-228.
- Grassle, J. F. and W. Smith. 1976. A similarity measure sensitive to the contribution of rare species and its use in investigation in marine benthic communities. *Oecologia (Berlin)* 25: 13-22.
- Harter, S. P. 1975a. A probabilistic approach to automatic keyword indexing I. On the distribution of specialty words in technical literature. *J. ASIS* 26: 197-206.
- Harter, S. P. 1975b. A probabilistic approach to automatic keyword indexing II. An algorithm for probabilistic indexing. *J. ASIS* 26: 280-289.

- Heitjan, D. F. 1993. Ignorability and coarse data: some biomedical examples. *Biometrics* 49: 1099-1109.
- Hogeweg, P. and B. Hesper. 1984. The alignment of sets of sequences and the construction of phyletic trees: an integrated method. *J. Mol. Evol.* 20: 175-186.
- Kacprzyk, J. 1985. Group decision making with a fuzzy majority via linguistic quantifiers I. a consensory-like pooling. *Cyber. Syst.* 16: 119-129.
- Kacprzyk, J. 1986. Group decision making with a fuzzy linguistic majority. *Fuzzy Sets and Systems* 18: 105-118.
- Krishna-Iyer, P. V. 1949. The first and second moments of some probability distributions arising from points on a lattice and their application. *Biometrika* 36: 135-141.
- Lambert, J. M. 1972. Theoretical models for large scale vegetation survey. In: J. N. R. Jeffers (ed.), *Mathematical Models in Ecology, Symp. Brit. Ecol. Soc.* 12: 87-109.
- Lance, G. N. and W. T. Williams. 1967. A general theory of classificatory sorting strategies I. Hierarchical systems. *Comput. J.* 9: 373-380.
- Lance, G. N. and W. T. Williams. 1968. Note on a new information statistic classificatory program. *Comput. J.* 11: 195.
- Lankhorst, M. M. and R. Moddemeijer. 1992. Automatic word categorisation: an information-theoretic approach. Dept. Computer Sci., Rijksuniversiteit Groningen, Note CS 9209.
- Lewis, P. A. W., P. B. Blaxendale and J. L. Burnett. 1967. Statistical discrimination of synonymy/antonymy relationship between words. *J. Assoc. Comput. Mach.* 14: 20-44.
- Liebovitch, L. S. 1995. Ion channel kinetics. In: P. M. Iannaccane and M. K. Khokha (eds.), *Fractal Geometry in Biological Systems: an Analytical Approach*, CRC Press, London. pp. 31-56.
- Lubásky, E. 1969. Abiotic ecological factor classification in quantitative analysis of ecosystems. In: T. Frey (ed.), *Conference on Quantitative Methods in Vegetation Study*. Tartu State University, Estonia. pp. 253-256.
- Michler B. and C.-G. Arnold. 1996. Pyrrolizidinalkaloide in Beinwellwurzeln. Ein Beispiel für die Variabilität sekundärer Pflanzeninhaltsstoffe. *Deutsche Apotheker Zeitung* 136, 29: 15-20.
- Michler, B. and C.-G. Arnold. 1999. Predicting presence of proazulenes in the *Achillea millefolium* group. *Folia Geobotanica* 14: 143-161.
- Orlóci, L. 1991. On character-based plant community analysis: choice, arrangement, comparison. *Coenoses* 6: 103-107.
- Panayirci E. and R. C. Dubes. 1983. A test for multidimensional clustering. *Pattern Recognition* 16: 433-444.
- Pillar, V. de Patta. 1999. On the identification of optimal plant functional types. *J. Veg. Sci.* 10: 631-640.
- Pilar, V. de Patta and L. Orlóci. 1993. Taxonomy and perception in vegetation analysis. *Coenoses* 8: 53-66.
- Podani, J. 1989. New combinatorial SAHN clustering methods. *Vegetatio* 81: 61-77.
- Raunkiaer, C. 1934. *The Life Forms of Plants and Statistical Plant Geography*. Oxford Univ. Press, Oxford.
- Roberts, D. W. 1986. Ordination on the basis of fuzzy sets. *Vegetatio* 66: 123-131.
- Roberts, D. W. 1987. An anticommutative difference operator for fuzzy sets and relations. *Fuzzy Sets and Systems* 21: 35-42.
- Ronkainen, P. 1998. Attribute similarity and event sequence similarity in data mining. Dept. Computer Science, Univ. Helsinki, Ser. Publ. C-1998-42.
- Ross, D., M. B. Dale, K. Shields and C. Hulett. 1986. *TAXON Users Manual version P4. Reference Manual 6*, Division of Computing Research, CSIRO, Canberra.
- Sandland, R. L. and P. C. Young. 1979. Probabilistic tests and stopping rules associated with hierarchical classification techniques. *Austral. J. Ecol.* 4: 399-406.
- Smouse, P. E., J. C. Long and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35: 627-632.
- Sneath, P. H. A. 1966. A comparison of different clustering methods as applied to randomly-spaced points. *Classif. Soc. Bull.* 1: 2-18.
- Tversky, A. 1977. Features of similarity *Psychol. Rev.* 84: 327-352.
- van der Maarel, E. and M. T. Sykes. 1993. Small-scale plant species turnover in a limestone grassland: the carousel model and some comments on the niche concept. *J. Veg. Sci.* 4: 179-188.
- Wallace, C. S. and D. L. Dowe. 2000. MML clustering of multi-state, Poisson, von Mises circular and Gaussian distributions. *Statistics and Computing* 10: 73-83.
- Weir, A. D. (update K. Mayo) 1972. *Program DIVINFRE: a divisive classification on binary data including reallocation*. CSIRO Div. Land Research, Tech. Memo. 72/4.
- Wilson, J. B. 1999. Assembly rules in plant communities. In: E. Wehler and P. Keddy (eds) *Ecological Assembly Rules - Perspectives, Advances, Retreats*. Cambridge University Press, Cambridge, U. K. pp. 130-164.
- Wilson, J. B. and S. H. Roxburgh. 1994. A demonstration of guild-based assembly rules for a plant community, and the determination of intrinsic guilds. *Oikos* 69:267-276.
- Wishart, D. 1969. Mode analysis: a generalisation of nearest neighbour which reduces chaining effects. In: A. J. Cole (ed.), *Numerical Taxonomy*, Academic Press, London. pp. 282-234.
- Wisheu, I. and P. A. Keddy. 1992. Competition and centrifugal organisation of plant communities: theory and tests. *J. Veg. Sci.* 3: 147-156.
- Yu, P.-L. 1992. To be a great operations researcher from a MCDM scholar. *Computers and Operations Research* 19: 559-561.
- Ziv, J. and A. Lempel. 1977. A universal algorithm for sequential data compression. *I. E. E. Transactions on Information Theory* IT-23:337-343.