A NUMERICAL APPROACH TO PHYLOGENETIC SYSTEMATICS[1,2]

JAMES S. FARRIS, ARNOLD G. KLUGE, AND MICHAEL J. ECKARDT

Abstract

Farris, J. S. (Biol. Sci., State Univ., Stony Brook, New York, 11790), Kluge, A. G., and Eckardt, M. J. (Zool., Univ. Michigan, Ann Arbor 48104) 1970. A Numerical approach to phylogenetic systematics. Syst. Zool., 19:172–191.—Principles abstracted from Hennig (1966) are used as axioms to form a quantitative analog of phylogenetic systematics. A close connection is demonstrated between phylogenetics and most parsimonious trees. The compatibility of some existing clustering methods with the principles is discussed, and a new clustering technique, the Weighted Invariant Step Strategy (WISS) is described. Generalization of the axioms to the case where direction of evolution is not assumed is examined, and it is shown that the Wagner Method for estimating evolutionary trees is consistent with the generalized phylogenetic axioms.

The taxonomic philosophy propounded in Hennig's (1966) *Phylogenetic Systematics* is unique among non-quantitative approaches to evolutionary taxonomy for the detail in which its premises have been described. The corresponding "phylogenetic," or "cladist" school of taxonomists have been proportionately influential. Viewpoints similar to those of Hennig (1966) have consequently been incorporated into mathematical models underlying some novel techniques in quantitative evolutionary taxonomy. These facts have led us to study the formal implications of Hennig's treatise.

The aims of this paper are to consider what restrictions are imposed on quantitative evolutionary procedures by the premises of *Phylogenetic Systematics*, to construct quantitative techniques consistent with those premises, and to evaluate existing quantitative techniques with respect to the premises. It is not our primary aim to criticize either the principles or the methods of *Phylogenetic Systematics*, although we shall consider some generalizations of the method. We do not claim to consider all the issues raised by *Phylogenetic Systematics*. We shall restrict ourselves to one central topic: Inference of evolutionary trees.

We shall throughout use "phylogenetic," "phylogeneticist," and "phylogenetics" to refer only to the taxonomic philosophy and methods of Hennig (1966). We shall use the term "monophyletic" in the sense of Hennig.

PHYLOGENETIC AXIOMS

Our formal development will be based on a set of axioms abstracted from Hennig (1966).

AI. A character ("transformation series" of Hennig) is a collection of mutually exclusive states (attributes; features; "characters," "character states," or "stages of expression" of Hennig) which

a) have a fixed order of evolution such that

b) each state is derived directly from just one other state, and

c) there is a unique state from which every other state is eventually derived.

AIa should not be taken to imply that phylogeneticists assume a character to evolve in only one way or in only one direction. We will generally assume that
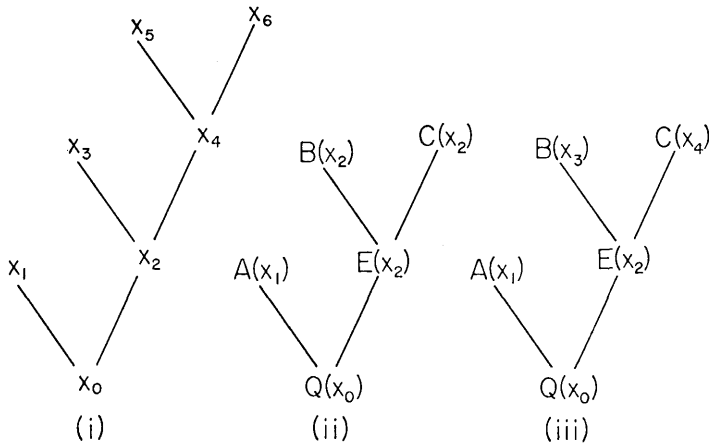
Fig. 1.—A character state tree (i); hypothetical phyletic trees (ii), (iii).

AIa holds only over an appropriately restricted section of the evolutionary tree.

AIb and AIc imply that the evolutionary order of a character can be represented as a *character state tree* (cf. Camin and Sokal, 1965) with a unique root and no closed loops. Such a tree is depicted in Fig. 1(i).

The evolutionary ordering of the states of a character allows us to classify states as ancestral ("plesiomorphous" of Hennig) or derived ("apomorphous" of Hennig). These terms may be used in a relative sense, as "state a is ancestral relative to state b."

AII. "All groups regarded as monophyletic are distinguished by the possession of derived (apomorphous) stages of expression . . . of at least one pair of characters . . ." (Hennig, 1966:91).

AII implies that a minimal condition for a group, G, of OTUs to form a monophyletic group in the phylogenetic system is that there is at least one character with a "stage of expression" (i.e., a state) x, such that [a]all the OTUs in G "possess" x, and [b]no OTU outside G "possesses" x.

The "pair of characters" referred to above results from the fact that every monophyletic group distinct from G—and, in particular,

the "sister group" (Hennig, 1966) of G— must also "possess" a distinctive character state.

The meaning in this context of the condition that an OTU or group of OTUs "possesses" a character state differs from usual numerical taxonomic usage. "It makes no difference whether the synapomorphy [possession of a derived state] consists in the fact that an apomorphous character [state] (a′) is present identically in all species . . . , or whether it is present in different derived conditions (a′ and a″)" (Hennig, 1966:90). This point is exemplified in Fig. 1. Fig. 1(i) specifies the evolutionary order of a set of character states. In Fig. 1(ii) OTUs B and C share a state, x_2, not present in OTU A, and so form a permissible monophyletic group. In Fig. 1(iii) B and C do not have the same state, but the states they have are both derivable from x_2, while state x_1, present in A, is not so derivable. B and C therefore again form a permissible monophyletic group.

To avoid confusion, we will recast AII in terms of some new definitions. A character state will be said to *describe* an OTU (group of OTUs) if and only if the OTU (all members of the group) have just that character state. A group, G, of OTUs will

be said to *share a step*, x, if and only if every OTU in G either (a) has *state* x, or (b) has a state apomorphous relative to *state* x according to the evolutionary ordering assumed for the character. Here we use the concept of a *step*, x, as the *transition* from state y to state x, where y is the most derived state that is ancestral relative to state x. For example, in Fig. 1(iii), B and C share step x_2—that is the step from x_0 to x_2—since that character transition has occurred in the ancestry of both OTUs. On the character state tree of Fig. 1(i) the nodes of the tree correspond to states, while the internodes, or branches, correspond to steps. On a character state tree with no loops, each apomorphous state corresponds to exactly one step; hence we may without confusion use the name of the state to label the corresponding step. Further, we use the name of a character, X, say, to refer either to a collection of states or a collection of steps, there being again no danger of confusion. Thus we will refer to both *state* x and *step* x as belonging to X. It is convenient to consider state x_0, the most ancestral state of X, to correspond to a "null" step, x_0. *We explicitly reserve "sharing states" between OTUs to refer to those OTUs' being described by states in common.* This is done in order to avoid the introduction of multiple meanings of "shared states" in the literature. We will find it convenient to refer to the states *derivable* from a state, x. We define a relation, d, on character states so that y d x (read "y is derivable from x") if and only if either (a) y = x, or (b) y is apomorphous relative to x. A specified evolutionary order on states is again assumed. Hence we would say that two OTUs with states y and z *share a step*, x, if and only if y d x and z d x. The relation d induces a *partial ordering* on the states of a character (cf. Estabrook, 1968). In our development we will assume properties of partial orders in dealing with d. The correspondence between states and steps conserves the ordering d.

We now restate:

AII. For every monophyletic group, G, of OTUs, there is at least one character, X, with a state, x, such that if a state y in X describes any OTU in G then

a) y d x,

while if a state z in X describes any OTU not in G then

b) z $\not{d}$ x.

AIII. In the absence of evidence to the contrary, any state corresponding to a step shared by a group, G, of OTUs is taken to have arisen just once in G.

AIII is a restatement of the "auxiliary principle" of Hennig (1966:121). Because of the correspondence between steps and states, AIII implies that any step (i.e., transition between states) shared by a group, G, of OTUs is assumed—unless there is contrary evidence—to have occurred just once in the ancestral lines giving rise to the OTUs of G. The step is then *unique* in the sense of Wilson (1965), as is the corresponding state. Note that for a multistate character, a unique, or even a *unique and unreversed* (cf. Wilson, 1965), state need not survive in every line possessing it.

AIV. "The more characters certainly interpretable as apomorphous (not characters in general) that there are present in a number of species, the better founded is the assumption that these species form a monophyletic group." (cf. Hennig, 1966:121). This statement is a re-translation of the original German manuscript, which was most graciously made available to us by Dr. Ranier Zangerl. The original text reads, *"Je mehr sicher als apomorph zu deutende Merkmale (nicht: Merkmale überhaupt!) bei einer Anzahl verschiedener Arten vorhanden sind, desto besser ist die Annahme begründet, dass diese Arten zusammen eine monophyletische Gruppe bilden."*

In the terms we introduced above under AII, AIV indicates that a group of OTUs sharing many steps is a group whose monophyly is supported by strong evidence. Only steps corresponding to derived states are counted. In our discussion we will refer to steps corresponding to derived states as *derived steps*.

THEOREMS OF PHYLOGENETICS

Axioms AI-AIV are sufficient to allow us to investigate the mathematics of constructing estimates of evolutionary trees in accordance with phylogenetic principles. In this section we assume that AI holds over the entire tree in question. A more general case is treated later.

TI. We first consider the problem of finding the character states that describe the (usually hypothetical) branching points of a tree. Given that two OTUs, B and C, together with their common ancestor E (cf. Fig. 1), comprise a monophyletic group, and given that for a character, X, B and C are described by states x_B and x_C, we must infer which state of X describes E. We know by AI that just one state of X describes E, since the states of a character are postulated to be mutually exclusive. We also know from AI that there is at least one state, x_i say, such that x_B d x_i and x_C d x_i, since all the states of X are presumed to be derived eventually from an ancestral state. Now for any state, x_i, such that x_B d x_i and x_C d x_i, the corresponding step is by definition shared by B and C. Therefore by AIII, x_i is presumed homologous in B and C, and so step x_i must also be shared by E. Since this is true of any state from which both x_B and x_C are derivable, it must be that E is described by a state, x_E, such that

a) x_B d x_E
b) x_C d x_E
c) If there is another state, x_K, such that x_B d x_K and x_C d x_K, then x_E d x_K.

That is, E is described by the most derived state from which both x_B and x_C may be derived. No state, x_i, such that x_B $\not{d}$ x_i or x_C $\not{d}$ x_i may describe E, since this would violate the ordering of the character states.

The relation between x_B, x_C, and x_E as defined above is a theorem deduced from AI and AIII. We shall denote that theorem TI.

In utilizing TI it is convenient to frame a relation on OTUs analogous to the relation, d, already defined on character states. For OTUs A and B (either or both of which may be hypothetical) and characters $X_1, X_2, \ldots X_N$, we will write A D B (read "A is derivable from B") if and only if for every character X_i, x_{iA} d x_{iB}, $1 \leq i \leq N$.

Since TI holds for every character, we may deduce an immediate corollary:

TII. If OTUs A and B have a most recent common ancestor, E, then E has the following properties:

a) A D E
b) B D E
c) for any OTU, K, A D K and B D K implies E D K.

Thus E must be the most derived hypothetical OTU from which A and B are both derivable.

We denote the most recent common ancestor of a set, G, of OTUs as J(G). TII allows us to convert the AII postulate about the group, G, to a statement about the ancestor, J(G), of the group. We already know from AII that any valid monophyletic group, G, must share at least one step not shared by an OTU outside G. Suppose that for character X there is a step, x_G, that is shared by G and not shared by OTUs outside G, and suppose that x_G is the most derived such step for X. Then by TI, J(G) is described by state x_G. Let F be any group of OTUs not in G. By hypothesis, no OTU, C, in F shares x_G, hence x_C $\not{d}$ x_G. Thus if x_F denotes the state of J(F) in X, then x_F $\not{d}$ x_G. Consequently, J(F) $\not{D}$ J(G). We state this fact as

TIII. For J(.) defined as above, if G is any monophyletic group permissible

under AII, and if F is any group of OTUs not in G, then $J(F) \not\supset J(G)$.

PHYLOGENETICS AND PARSIMONY

Theorems TI-TIII demonstrate a close relationship between the phylogenetic method and the method of most parsimonious trees, as proposed by Camin and Sokal (1965) and formalized by Estabrook (1968). Estabrook showed that TI and TII hold when hypothetical ancestors of groups are constructed to minimize the length of the tree. He also showed that any most parsimonious tree necessarily contains at least one monophyletic group, G, from whose most recent common ancestor, $J(G)$, it is not possible to derive the ancestor of any other possible group of OTUs. Estabrook proposed a combinatoric procedure to generate a tree by selecting groups, G, with ancestors, $J(G)$, such that if F is any group of OTUs not in G, then $J(F) \not\supset J(G)$. The set of trees generated by the Estabrook procedure consequently includes all the trees that are permissible under AI-AIII of the phylogenetic system.

In Estabrook's procedure, most parsimonious trees are identified by measuring the length of each permissible tree. Analogously, in the phylogenetic system, we choose the "correct" tree, at least conceptually, by asking which permissible tree is most in accord with AIV.

Unfortunately, AIV is not sufficiently detailed to allow us to select a unique criterion for choosing a most preferable tree. We know that trees on which the monophyletic groups share many steps are preferable to trees on which this is not so. But AIV deals only with single monophyletic groups and does not tell us how to evaluate a tree consisting of several monophyletic groups. One widely known criterion—parsimony—could be used to select trees. This would be in accord with AIV, since on a most parsimonious tree OTUs that share many steps (this is *not* the same as the OTUs' being *described* by many of the same *states*) are generally placed together. We might argue that the parsimony criterion selects a tree most in accord with AIV by "averaging" in some sense the preferability of all the monophyletic groups of the tree. Other criteria, however, may also agree with AIV. We shall describe a method based on a criterion correlated with, but not equivalent to, the parsimony criterion.

THE WEIGHTED INVARIANT STEP STRATEGY

One method for choosing a tree according to AIV might operate as follows.

1) Select a set, G, of OTUs whose inclusion in a monophyletic group is "best founded" in the sense that G shares at least as many derived steps as does any other possible collection of OTUs. Go to 2.
2) Delete the OTUs in G from the set of OTUs under consideration, replacing them by $J(G)$, the most recent common ancestor of G. Go to 3.
3) If the tree is incompletely specified, return to 1. Otherwise, stop.

Here we assume that the most preferable tree is produced by sequential application of AIV. We find at each stage a cluster of OTUs that are invariant in sharing many derived steps. Such a cluster clearly satisfies AIIa and AIV. The consequences of AIIb will be treated below.

This is a type of "weighted" cluster analysis in that, once $J(G)$ is substituted for G, it is treated just as an OTU (cf. Weighted Pair Group Analysis; Sokal and Sneath, 1963). We shall describe the mathematics needed to program the Weighted Invariant Step Strategy (WISS).

We must first decide how to count the number of shared steps in a group of OTUs. For simplicity in the following discussion, we assume that the groups selected always contain just two OTUs. Suppose that OTUs A and B share a step, x_4 (see Fig. 1(i)). From the definition of sharing, it is clear that A and B then also share x_2 and x_0. In general if two OTUs share a step that

is k steps removed from the ancestral state, then they share at least $k + 1$ steps, including the ancestral ("null") step. Thus in Fig. 1(i) we see that x_4 is $k = 2$ steps removed from x_0, and, as we have already observed, sharing x_4 subsumes sharing x_2 and x_0, for a total of $3 = k + 1$ steps.

To formalize this notion, we assign to each state, x_i, of a character, X, a value, $n(x_i)$ representing the number of steps needed to evolve from the ancestral state, x_0, of X to x_i. Thus in Fig. 1(i), we can see that $n(x_6) = 3$; $n(x_3) = 2$; $n(x_1) = 1$. AIV exhorts us to count only derived steps in choosing groups, so that the ancestral step, x_0, is not counted. Consequently, if two OTUs share a step x_i, the total number of shared derived steps implied by this fact is $n(x_i)$. Note that if only x_0 is shared, the total number of shared steps implied by $n(x_0) = 0$ is still correct.

We denote as $j(x_i, x_k)$ the most derived state, x_j, such that x_i d x_j and x_k d x_j. If OTUs A and B form a monophyletic group, they have most recent common ancestors, C, say, $C = J(\{A,B\})$. For every character, X, the state, x_C, that describes C is $j(x_A, x_B)$, and the number of derived steps shared by A and B in that character is $n(x_C) = n(j(x_A, x_B))$.

We define the *advancement* of an OTU, A, as follows. For characters $X_1, X_2, \ldots, X_N$, let A be described by states $x_{1A}, x_{2A}, \ldots, x_{NA}$. The advancement, $h(A)$, of A is given by

$$h(A) = \sum_{i=1}^{N} n(x_{iA}). \qquad (1)$$

The total number of derived steps shared by OTUs A and B is equal to $h(J(\{A,B\}))$.

An algorithm for WISS clustering would then proceed as follows:

1) Find a pair of OTUs, A and B, such that for any other OTU, E,

   a) $h(J(\{A,E\})) \leqslant h(J(\{A,B\}))$;

   b) $h(J(\{B,E\})) \leqslant h(J(\{A,B\}))$.

Go to 2.

2) Replace A and B by $C = J(\{A,B\})$. Go to 3.

3) If the clustering is incomplete, return to 1; otherwise stop.

It is shown below that

$h(J(\{A,B\})) = (\frac{1}{2})(h(A) + h(B) - diff(A,B))$,

where diff(A,B) is a measure of the phenetic difference between A and B. Hence $h(J(\{A,B\}))$ can be computed without computing $J(\{A,B\})$ itself. This fact renders feasible highly efficient computer algorithms for WISS clustering. Such an algorithm has been developed, and a FORTRAN IV program listing is available from the senior author.

Clustering by either the WISS or Camin-Sokal method produces monophyletic groups that share many derived steps, so that AIIa and AIV are satisfied. AIIb, however, requires that at least one of the steps shared by a monophyletic group G, not be shared by any OTU outside G. It is not always possible to satisfy AIIb with a fixed coding of the data. For example, in the data

O T U	Character	
	1	2
A	1	1
B	1	1
C	0	1
D	1	0,

where 0 is the ancestral state, it is not possible to form any monophyletic groups satisfying AIIb. If OTUs A and B, say, were united, they would share steps "1" in both characters. However, neither corresponding state is restricted to the A,B group, for state "1" of character 1 occurs also in D, while state "1" of character 2 describes C.

One way to resolve this type of difficulty is simply to collect data for more characters until AIIb is satisfied. Other approaches are possible, however. AIIb is technically satisfied if we recode the data, asserting that the "1"'s of D and C are not the same states as those of A and B. Such recoding

is not inconsistent with the phylogenetic philosophy. While AI postulates that the evolutionary ordering of a character is fixed, we still face the possibility of error in determining what the ordering is. With real data, it is often possible to find external evidence that the recoding should have been done anyway.

One interpretation of the spirit of AIIb is that it requires that some of the shared derived steps on which a monophyletic group is based be cladistically reliable in the sense of Farris (1969), in that the corresponding states have not arisen independently in other groups. We would then desire preferentially to create monophyletic groups based on reliable character states. This might be accomplished by combining a WISS clustering procedure with a successive weighting program (Farris, 1969). This approach is now being investigated.

PHENETIC CLUSTERING PROCEDURES

In this section we investigate the compatibility with the phylogenetic axioms of some phenetic clustering methods that have been utilized in evolutionary studies.

Throckmorton (1968) suggested forming an evolutionary tree by performing a complete linkage cluster analysis. His index of similarity between OTUs A and B was the number of derived states describing both A and B. There are two issues here: the form of the cluster analysis and the nature of the similarity measure. The second topic we reserve for later discussion. In this section we assume binary valued characters with one ancestral and one derived state. The effect of this convention is just that *with every character binary valued* the number of derived states describing both OTU A and OTU B is equal to $h(J(\{A,B\}))$. We may thus consider the merits of the clustering procedure independently of the consequences of the similarity coefficient. Under the same restrictions on characters, the number of shared steps between two OTUs is equal to the number of shared states. For con-

venience in comparing methods we will in this section refer to both quantities as number of shared states.

Throckmorton's rationale for utilizing complete linkage analysis is that it effectively estimates the maximum size of the set of derived states invariant within a cluster. It is true that the complete linkage value between two clusters is an upper bound for the number of derived states invariant within the entire group. However, the linkage value computed by the complete linkage procedure is not the least upper bound for the number of invariant derived states. Suppose that each OTU is considered as a collection of derived states. Any derived state invariant within a cluster is present in every OTU of the cluster. Then the set of states invariant within a lineage is the intersection of all the sets representing all the OTUs in the lineage. The number of states in that intersection may not exceed the complete linkage similarity level for that lineage, and will often be smaller. Consider the hypothetical data,

| | Character | | | |
OTU	1	2	3	4
A	1	1	1	0
B	1	1	0	1
C	0	0	1	1

where the ones are taken to be derived states and the zeroes are the ancestral states. A and B share two derived states, and form a cluster. C shares one state each with A and B. If complete linkage cluster analysis were performed, C would be linked to the A,B cluster at the level corresponding to one shared state. In reality, however, there is no derived state common to A, B, and C. C should be linked to the A,B cluster only at the level corresponding to no shared derived states.

The disagreement between the two procedures can in this case be resolved by Throckmorton's (1968) "delete, cluster, delete cycle," according to which derived states describing OTUs in distinct clusters after cluster analysis at step i are deleted

from the analysis for a new cluster analysis step i + 1. Here the "1" states of characters 3 and 4 are deleted because they describe both A and C and both B and C, respectively. At the second cluster analytic step, A and B are united with 2 shared derived states, and C is united to the A,B cluster with no shared, derived states. The deletion procedure has some drawbacks, however. For a more elaborate data set:

OTU	1	2	3	4	5	6	7	8
A	1	0	1	1	0	1	0	1
B	1	0	1	0	1	0	1	1
C	0	1	0	1	1	1	1	0
D	0	1	0	0	0	0	0	0

(Character is the header over columns 1–8)

We can see that the discrepancy between the actual number of shared derived states and the complete linkage estimate can alter the outcome of clustering. Performing complete linkage analysis, A and B form a cluster with 3 shared states. C shares two states each with A and B, so the cluster A,B,C is formed with "two shared states." D shares no states with A or B, so D is united to the A,B,C cluster at the level of no shared states. Now let us unite groups according to the actual number of shared states. Again, A and B are linked with three shared derived states. C shares no derived states with both A and B, but does share one derived state with D. Then a cluster C,D is formed with one shared state. Finally, cluster A,B is united to cluster C,D at the level of no shared states. The results of complete linkage clustering for these data are depicted in Fig. 2(i), those of WISS clustering in Fig. 2(ii).

Application of the deletion cycle procedure in this case produces some interesting results. State "1" of character 2 is lost, since it occurs in both the A,B,C cluster and the D "cluster." States "1" of characters 3 and 6 are lost, occurring in both A and C, as are states "1" of characters 3 and 4, occurring in B and C. The final tree (Fig. 2(iii)) unites A and B with 3 shared states and (A,B), C, and D with no
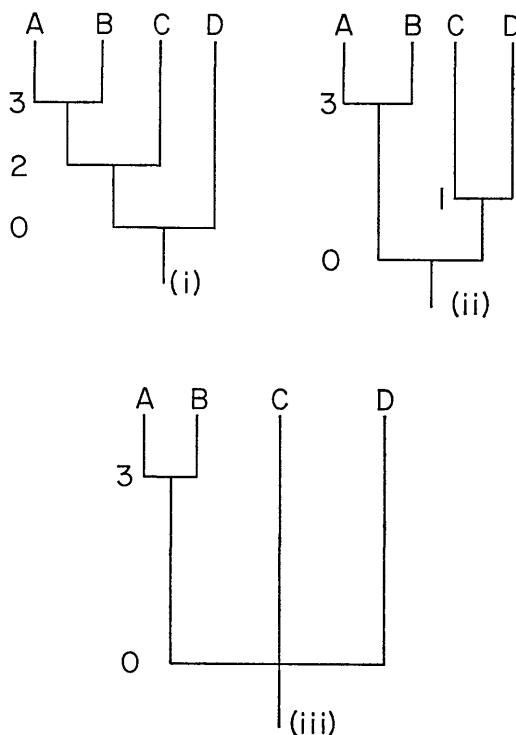

FIG. 2.—Results of clustering four OTUs by complete linkage cluster analysis (i), by counting shared derived states (ii), and by deleting some derived states (iii). Numerals indicate the number of shared derived states assigned to the adjacent linkage.

shared states. The initial complete linkage dendrogram (Fig. 2(i)) for these data contains a cluster ((A,B), C) which has no shared derived steps, hence is not a permissible group under AII. The final dendrogram after deletions (Fig. 2(iii)), on the other hand, lacks the group C,D, for the existence of which there is some phylogenetically valid evidence, the sharing of step "1" of character 2. In either case, it would appear that a complete linkage clustering procedure is in general inconsistent with the principles of the phylogenetic system.

Throckmorton's sole stated reason for using complete linkage analysis was that it would estimate the size of sets of shared, derived character states. Under that rationale, however, it is clearly more appropriate

actually to count the number of shared derived states, as would be done by either a WISS or Camin-Sokal analysis of the hypothetical data above. Since counting shared derived states can lead to clustering different from that produced by complete linkage analysis, we conclude that there does not seem to be any reason to use complete linkage analysis in evolutionary studies. We certainly cannot confirm Throckmorton's conclusion that, "of the obvious clustering methods, *complete linkage of derivative states* is the only one that estimates an operational phylogenetic parameter, and, hence, it is the only obvious method to use for phylogenetic analyses."

Similar criticisms apply to most of the commonly used phenetic clustering procedures. For the data set just discussed, application of Single Linkage Analysis, Weighted Pair Group Analysis, or Unweighted Pair Group Analysis yields a tree with the same cladistic form as that of Fig. 2(i), though the exact values of the linkage levels depend on the method used.

A CODING MODEL

We introduce here a model based on alternative methods of character coding as a means of comparing the properties of similarity coefficients.

Each character, X_i, is a collection of states, $x_{i0}, x_{i1}, \ldots, x_{ip_i-1}$ with p_i elements. We can associate with an array of characters a set of binary variables, v. A binary variable v_{ij} corresponds to state x_{ij} of character X_i. The variables, v, can be used to describe OTUs as can the character states, x. The variables, v, represent, then, an alternative way of encoding the information contained in the original characters. Various methods for recoding x's as v's are possible, and turn out to have implications of interest to evolutionary taxonomy.

We shall discuss three means of binary coding. In *non-additive coding* (Sokal and Sneath, 1963), the value, v_{ijA}, of variable v_{ij} for OTU A is unity if and only if OTU A is described by state x_{ij}. In *additive coding*, v_{ijA} has value unity if and only if $x_{iA} \, d \, x_{ij}$. In *semi-additive coding*, v_{ijA} has value unity if and only if OTU A is described by x_{ij}, except that v_{i0}—the variable corresponding to the ancestral state, x_{i0}, of character X_i—is unity for every character and every OTU. For all three methods a variable, v, that does not have value unity has value zero.

The character states of Fig. 1(i) would have the non-additive coding,

	variable						
state	v_0	v_1	v_2	v_3	v_4	v_5	v_6
x_0	1	0	0	0	0	0	0
x_1	0	1	0	0	0	0	0
x_2	0	0	1	0	0	0	0
x_3	0	0	0	1	0	0	0
x_4	0	0	0	0	1	0	0
x_5	0	0	0	0	0	1	0
x_6	0	0	0	0	0	0	1 ;

the additive coding,

	variable						
state	v_0	v_1	v_2	v_3	v_4	v_5	v_6
x_0	1	0	0	0	0	0	0
x_1	1	1	0	0	0	0	0
x_2	1	0	1	0	0	0	0
x_3	1	0	1	1	0	0	0
x_4	1	0	1	0	1	0	0
x_5	1	0	1	0	1	1	0
x_6	1	0	1	0	1	0	1 ;

and the semi-additive coding,

	variable						
state	v_0	v_1	v_2	v_3	v_4	v_5	v_6
x_0	1	0	0	0	0	0	0
x_1	1	1	0	0	0	0	0
x_2	1	0	1	0	0	0	0
x_3	1	0	0	1	0	0	0
x_4	1	0	0	0	1	0	0
x_5	1	0	0	0	0	1	0
x_6	1	0	0	0	0	0	1 .

The variable v_0 is, of course, constant for additive and semi-additive coding, and hence has no effect on comparisons of OTUs. We have retained v_0 explicitly in these cases so that our notation is homogeneous over coding types.

The additive coding preserves the form and direction of the evolutionary ordering of Fig. 1(i). The semi-additive coding preserves the direction by its special handling of x_0 and v_0. The non-additive coding discards both form and direction of the character-state tree of Fig. 1(i) and preserves only the identity of the character states.

Our usage of the term "non-additive coding" is identical to that of Sokal and Sneath (1963). Our "additive coding" is equivalent to that of Sokal and Sneath except for the interpretation of the ordering relation used to define the binary coding. Sokal and Sneath used the relation "≧" applied to the (possibly arbitrary) numerical scale on which the original character is defined, while we use the relation "d" applied to the evolutionary order of the states of a character. In the case of a character with linearly ordered states coded numerically with the ancestral state having minimal numerical value, "d" is equivalent to "≧." A conceptual—though not formal—distinction between the two usages is that in the phylogenetic system, the ordering relation is significant to the theory of the system itself, while in phenetic practice this need not be.

Additive coding corresponds directly to the operations employed in the phylogenetic system. The reader may verify by inspection of the tree of Fig. 1(i) that, for the v_{ij} additively coded,

$$n(x_i) = \sum_{j=0}^{6} \mid v_j(x_i) - v_j(x_0) \mid , \qquad (2)$$

hence in general, for OTU A,

$$h(A) = \sum_{i,j} \mid v_{ijA} - v_{ijQ} \mid , \qquad (3)$$

where Q is again an OTU with the ancestral state for every character. Now we define the *difference* between OTUs A and B as

$$\text{diff}(A,B) = \sum_{i,j} \mid v_{ijA} - v_{ijB} \mid , \qquad (4)$$

and note that

$$h(A) = \text{diff}(A,Q). \qquad (5)$$

Suppose that OTUs A and B form a monophyletic group with most recent common ancestor C = J({A,B}). As we have already seen, x_{iA} d x_{iC} for every character, x_i. Then for all i and j, $v_{ijC} \leqq v_{ijA}$. Hence,

$$h(A) = h(C) + \text{diff}(A,C). \qquad (6)$$

Similarly,

$$h(B) = h(C) + \text{diff}(B,C). \qquad (7)$$

Now if v_{ijA} and v_{ijB} are both unity, v_{ijC} is also unity; and if v_{ijA} and v_{ijB} are both zero, v_{ijC} is also zero. Consequently, $\mid v_{ijA} - v_{ijB} \mid$ is unity if and only if $\mid v_{ijA} - v_{ijC} \mid$ or $\mid v_{ijB} - v_{ijC} \mid$ is unity; and $\mid v_{ijA} - v_{ijC} \mid$ and $\mid v_{ijB} - v_{ijC} \mid$ cannot be simultaneously positive. Thus

$$\text{diff}(A,B) = \text{diff}(A,C) + \text{diff}(B,C). \qquad (8)$$

Combining (6) and (7),

$$h(A) + h(B) = 2h(C) + \text{diff}(A,C) + \text{diff}(B,C). \qquad (9)$$

Substituting according to (8),

$$h(A) + h(B) = 2h(C) + \text{diff}(A,B). \qquad (10)$$

Rearranging,

$$h(C) = (\tfrac{1}{2})(h(A) + h(B) - \text{diff}(A,B)). \qquad (11)$$

Equation (11) is useful in the WISS procedure, since it allows h(J({A,B})) to be computed from the differences (4) between A, B, and Q, without computing J({A,B}) itself. If the original characters, X, all have linear orders on their states, the common *phenetic difference*, p,

$$p(A,B) = \sum_{i=1}^{N} \mid x_{iA} - x_{iB} \mid , \qquad (12)$$

in the sense of Farris (1967) can be used to compute h:

$$h(J(\{A,B\})) = (\tfrac{1}{2})(p(A,Q) + p(B,Q) - p(A,B)). \qquad (13)$$

Equation (13) is utilized by one version of the WISS computer program. Thus, while the coding variables, v, provide a basis for a useful model, it is not in prac-

tice necessary to resort to binary coding in order to compute h values.

MATCHING COEFFICIENTS

We have seen that when an additive coding is used, the difference of (4) corresponds to operations in phylogenetic methods. When a *non-additive coding* of the v's is used, the statistic,

$$C_A(A,B) = P - (\text{diff}(A,B))\,(\tfrac{1}{2}),\ (14)$$

where P is the number of variables, v_{ij}, effectively counts the number of states describing both A and B. Thus (14) is equivalent to the *concordance of all states* of Throckmorton (1968). Likewise, if a *semi-additive coding* is used, the statistic

$$C_D(A,B) = P - N - (\text{diff}(A,B))\,(\tfrac{1}{2}),\ (15)$$

where N is the number of characters, counts the number of derived states describing both A and B and is equivalent to the *concordance of derived states* of Throckmorton.

In the WISS procedure, OTUs A and B with large values of $h(J(\{A,B\}))$ are clustered together. We can consider $h(J(\{A,B\}))$ as a kind of similarity coefficient. In the procedures proposed by Throckmorton (1968), OTUs A and B are clustered together if $C_A(A,B)$ or $C_D(A,B)$ is large. The correspondence between types of coding and type of similarity coefficient allows us to consider the relative information content of the coefficients. Since $h(J(\{A,B\}))$ corresponds to additive coding, it conserves information on both the form and direction of the evolutionary orders of character states. C_D, corresponding to semi-additive coding, conserves the directionality of the characters. C_A conserves only the state identities.

The information loss induced by concordance can alter the clustering of a set of OTUs. Suppose that for the data

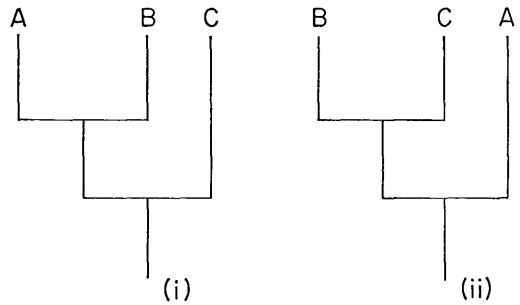|		Character	
OTU	1	2	3
A	3	2	0
B	2	1	0
C	1	1	1 ,

FIG. 3.—Three hypothetical OTUs clustered according to concordance of derived steps (i); the same OTUs clustered by number of shared derived states (ii).

each character is taken to have a linear evolutionary ordering for its states, zero being the most ancestral state. Clustering by shared derived steps, we achieve the tree of Fig. 3(i), the monophyletic groups of which are phylogenetically permissible. Clustering by concordance of derived states, we note that B and C are described by one derived state in common, while A is described by no derived state in common with B or C. Thus, the tree of Fig. 3(ii) results. The group B,C of Fig. 3(ii) does share two derived steps, but these states are also shared by A. Therefore, Fig. 3(ii) depicts a tree that is not permissible by AII.

Concordance cannot generally be used to achieve a valid phylogenetic clustering. In a more general context, we would expect that the loss of information inherent in clustering by concordance would rule out clustering by concordance in any sort of evolutionary application. We may note one exception, however. If all the characters in a study have but two states, a matching method such as concordance can be used without loss of information, since only trivial sequence information is present in the character state trees.

CLUSTERING BY DERIVED STEPS

Equation (11) could be taken as a definition of a similarity coefficient, $s(A,B) =$

$h(J(\{A,B\}))$, which depends only on the number of shared derived steps between two OTUs. (11) indicates, however, that, recalling (5), $s(A,B)$ can be computed from a difference (4) that is calculated on the basis of all character states. It is therefore clear that the aim of clustering by derived steps alone does not at once require that ancestral states be physically deleted from the analysis.

A practical consequence of (11) (or(13)) is that it is not necessary to physically alter a data set in order to accomplish derived-step clustering or in order to achieve several different "derived-step" clusterings corresponding to several hypotheses as to which states are ancestral. Once a matrix of values of p (12) or diff (4) has been computed, the s values corresponding to a hypothesis can be calculated simply by specifying Q. Although diff (4) is calculated on the basis of the binary coding variables, whose values depend on the choice of ancestral states, it can be shown that the matrix of diff itself is invariant over choices of Q.

Some workers have apparently believed that actual deletion of ancestral states is necessary to accomplish clustering by derived steps. This seemingly trivial ignorance has had one significant ramification. Some have concluded that no type of analysis that does *not* physically delete ancestral states can be a truly evolutionary method. Thus, it has been asserted that the most parsimonious tree techniques of Camin and Sokal (1965) or Kluge and Farris (1969) are not valid evolutionary techniques *because* they do not delete ancestral states. Equation (11) indicates clearly that this criticism is unfounded.

REVERSAL AND ITS CONSEQUENCES

Our argument thus far has assumed that AI holds for the entire evolutionary tree under consideration. This is equivalent to assuming irreversibility of evolution. Kluge and Farris (1969) have criticized the presumption of irreversibility, and, as we note above, irreversibility is generally not assumed by phylogeneticists. We shall generalize our methods so that they hold for a weaker interpretation of AI.

We shall assume that the states of a character are related by a character state tree of fixed form, but that

a) evolution is permitted in either direction along any branch of the character state tree, and

b) any state is potentially permissible as the most ancestral state for some restricted region of the evolutionary tree.

We retain the convention that the character state tree has no closed loops. We shall refer to these conditions on character state trees collectively as AI'.

By analogy with the relations d and D introduced earlier, we define two new relations appropriate for discussion of AI'. By $y \; d_X \; z$, read, "y is derivable from z with respect to x," we intend that state y is derivable from state z, taking state x to be most ancestral. Similarly, $A \; D_E \; B$, read, "A is derivable from B with respect to E" implies that A is derivable from B in a frame of reference in which E has (by definition) all ancestral character states. As before, $A \; D_E \; B$ if and only if for any character, X_i, $x_{iA} \; d_{x_{iE}} \; x_{iB}$. Note that neither "$y \; d_X \; z$" nor "$y \; d \; z$" implies the other.

Where earlier we would speak of a step's being absolutely shared by a group of OTUs, here steps may be shared only *with respect to* a given reference point. We will say that a step, x_C, is shared by A and B with respect to E if and only if $x_A \; d_{x_E} \; x_C$ and $x_B \; d_{x_E} \; x_C$.

It is helpful to consider the properties of the conditional relation d_X in terms of the character state tree of X. For the character state tree of Fig. 4, we can see that x_8 and x_9 are both derivable from x_7 with respect to x_1. Thus, OTUs described by x_8 and x_9 could validly be assigned to a monophyletic group if x_1 were taken to
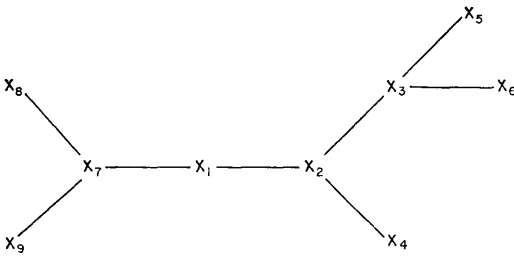
FIG. 4.—A generalized character state tree. Any of the states may be locally most ancestral.

FIG. 5.—A region of an evolutionary tree. OTU E is the local ancestor of cluster A, B, C.

be ancestral. But suppose x_7 were taken to be ancestral. There is no state, x, such that $x_8 \; d_{x_7} \; x$ and $x_9 \; d_{x_7} \; x$. Then OTUs described by x_8 and x_9 could not, for x_7 ancestral, be assigned to a monophyletic group. In terms of the form of the character state tree, we can generalize that if and only if states x_i and x_j can be connected on the tree by a path that does not contain the ancestral state, x_a, then there is a state, x_s, such that $x_i \; d_{x_a} \; x_s$ and $x_j \; d_{x_a} \; x_s$.

Substitution of AI′ for AI necessitates some modifications in our interpretations of AII-AIV. Under AI′ we can no longer guarantee that a particular state is everywhere derived, but only that a state is locally derived in some part of the tree. Conversely, any state that is locally derived in a tree region containing a group, G, is *a priori* equally capable of providing evidence for the validity of G. Then we would consider AIIa satisfied in Fig. 5 if there were states x_C and y_R in characters X and Y such that $x_A \; d_{x_E} \; x_C$, $x_B \; d_{x_E} \; x_C$, $x_H \; \not{d}_{x_E} \; x_C$, $x_K \; \not{d}_{x_E} \; x_C$, $y_H \; d_{y_E} \; y_R$, $y_K \; d_{y_E} \; y_R$, $y_A \; \not{d}_{y_E} \; y_R$, and $y_B \; \not{d}_{y_E} \; y_R$, these conditions implying nothing as to whether x_C and y_R are also locally derived in other regions of the tree. Similarly, in applying AIV to a group, G, we would count any states as derived that were locally derived near G, regardless of whether those states were also derived in other regions of the tree.

Since reversals of evolution are permitted under AI′, it may happen that a state, x, that is somewhere ancestral, may have "x"
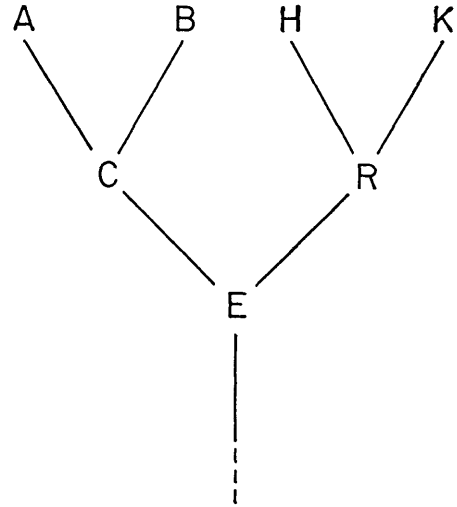
—itself—as an ultimately derived state according to an evolutionary tree consistent with AI′ and AII-AIV. In a sense, then, constructing a tree by AI′ may amount to an implicit recoding of some states in some OTUs. As noted previously, recoding in the process of reaching conclusions is consistent with the phylogenetic philosophy. The "recoding" is, in any event, necessary if irreversibility is not to be tacitly assumed. To take this fact into account, we should conceive of character state trees in two ways. *Primary* (or *a priori*) character state trees describe the relative properties of states as we might initially encode them. Primary state trees should be regarded as having no fixed directionality. This is so even when one particular state may be regarded as actually having been present in the common ancestor of the group under study. *Secondary* (or *a posteriori*) character state trees describe the relations between many "states," some of which may be the "same" on the primary tree. The secondary character state tree for a character differs from the primary tree for that character as a function of the evolutionary conclusions

reached. Constructing a valid secondary tree is consequently partially equivalent to constructing a valid evolutionary tree. In order to avoid a *circulus vitiosus*, of course, we may consider only primary trees as input data in inferring evolutionary relationships.

If we knew—perhaps by Revelation—the true secondary character state trees for a study, we could apply the theory of previous sections directly, "reversals" already having been coded out of existence. But if we admit that our character state trees are primary, only first approximations, we must review our methods.

Under AI′, as we have seen, we may depend only on local properties of character state trees. Several of the results of previous sections can be extended to this case in a logical way.

The nearest common ancestor of a group, G, of OTUs can be constructed under AI′ just as under AI. In Fig. 5, E provides a local set of ancestral states. Hence by AIII, we assign to C of Fig. 5 a state, x_C, such that

a) $x_A \ d_{x_E} \ x_C$
b) $x_B \ d_{x_E} \ x_C$
c) for any OTU, M, $A \ D_E \ M$ and $B \ D_E$ $d_{x_E} \ x_G$ implies $x_C \ d_{x_E} \ x_G$.

The properties are closely analogous to TI; we shall refer to them as LTI.

As a corollary, we have:

LTII. If OTUs A and B have nearest common ancestor, C, with an ancestor,

E, (as in Fig. 5) then

a) $A \ D_E \ C$
b) $B \ D_E \ C$
c) for any OTU, M, $A \ D_E \ M$ and $B \ D_E$ M implies $C \ D_E \ M$.

Earlier we noted that TII provided a connection between the phylogenetic system and the theory of Camin-Sokal most parsimonious trees. Here we can establish a link between phylogenetics—with—reversals and Wagner most parsimonious

trees (Kluge and Farris, 1969; Farris, 1970). The (hypothetical) ancestors constructed by the Wagner procedures are those required by LTII, LTI being equivalent to the *median state property* of the hypothetical intermediates of Wagner Trees (Farris, 1970). This can be readily shown through the use of the binary coding variables introduced above. Suppose that a character, X_i, is expressed as an additive binary coding for the OTUs of Fig. 5, OTU E being taken as having the ancestral state of X_i. Given that state x_{iE} is locally ancestral, then the state, x_{iC}, describing C has binary variables v_{ijC} unity if and only if both v_{ijA} and v_{ijB} are unity. Now if A and B share a step, x_{ij}, with respect to E, then v_{ijA}, v_{ijB}, and v_{ijC} are all unity. If step x_{ij} is not shared with respect to E by A and B, then at least one of v_{ijA} and v_{ijB} is zero, v_{ijC} is zero, and v_{ijE} must be zero. This is an exhaustive classification of events, since any step, x_{ij}, must be either shared or not shared by A and B with respect to E. Under both conditions, v_{ijC} is the median of v_{ijA}, v_{ijB}, and v_{ijE}.

Under AI we were able to devise a similarity coefficient, $s(A,B) = h(J(\{A,B\}))$, that allowed us to cluster according to the number of derived steps shared by A and B. The possibility of reversals under AI′ prevents us from designating any states as invariably derived, so that we may not directly count the number of shared derived steps for two OTUs. We can, however, generalize the statistic by making it conditional on a reference point:

$$s_E(A,B) = (\tfrac{1}{2}) (\text{diff}(A,E) + \text{diff}(B,E) - \text{diff}(A,B)). \quad (16)$$

This is the number of derived steps shared by A and B, given that the states describing E are ancestral—*provided* that E is a sufficiently near ancestor of A and B that no reversals have occurred between E and A and B. That condition is the case, for example, in Fig. 5.

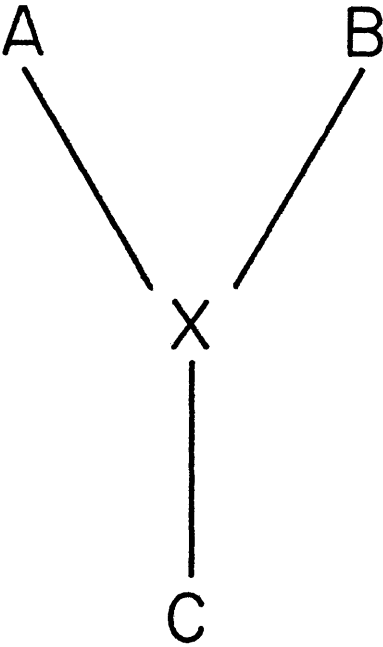The dependence of s_E on the reference-point, E, renders somewhat complex the

FIG. 6.—A neighborhood of a Wagner Network.

problem of constructing clustering schemes consistent with AIV, subject to AI'. Several different extensions of the WISS procedure to the case where reversals are allowed are possible. We might, for example, choose to cluster OTUs A and B with high average values of $s_X(A,B)$, where X ranges over the set of OTUs distinct from A and B. Alternatively, we might cluster according to the criterion

$$\max_{X \neq A,B} \{s_X (A,B)\} \quad .$$

The question of which clustering technique of the WISS type is most preferable remains unresolved at present. The axioms we have taken from Hennig (1966), or, for that matter, any statements in Hennig (1966), seem insufficient to provide a criterion. Choice between methods on the basis of statistical or analytical properties seems a promising possibility and is currently being investigated.

A quite different algorithm for calculating estimated trees under AI' is the Wagner Method (Kluge and Farris, 1969), an approximation technique for finding most parsimonious trees, no assumptions in irreversibility being made. We show here that the placement of OTUs on Wagner Networks (see Farris, 1970) is consistent with AIV subject to AI'.

Fig. 6 depicts a neighborhood of a Wagner Network. Any of the nodes A, B, C may be connected to other nodes not figured. Fig. 6 has a 3-way symmetry about the central node, X. In the figured neighborhood, either A, B, or C must be locally ancestral. Because of the symmetry, we may take C as locally ancestral without loss of generality. Note that the case where X is locally ancestral need not be considered, since A, B, or C may be taken as identical to X. Similarly, if some point lying on an internode of Fig. 6 is regarded as most ancestral, then Fig. 6 can be relabeled so that the relevant point becomes "A," "B," or "C."

Since Fig. 6 is taken as a neighborhood of a Wagner Network, no alteration of the linkages of the network will reduce the length of the network. In particular, we know that the network contains no nodes E, F such that

$$\text{diff}(A,(E,F)) < \text{diff}(A,X),$$

where

$$\text{diff}(A,(E,F)) = (\tfrac{1}{2}) (\text{diff}(A,E) +$$
$$\text{diff}(A,F) - \text{diff}(E,F)). \quad (17)$$

The diff(. , .) function here is equivalent to the diff(. , .) defined above for additive coding. (17) is derived by Farris (1970). Note that (17) is of the same form as (16).

Similarly, there are on the Wagner Network no nodes, Q, R, such that

$$\text{diff}(B,(Q,R)) < \text{diff}(B,X) \quad .$$

The evolution from local ancestor C to node A (or B) is measured by diff(A,C)

(or diff(B,C)); for Y = A or B, these measures can be partitioned exactly:

$$diff(C,Y) = diff(C,X) + diff(X,Y), \quad (18)$$

and diff(X,Y) for Y = A,B is the minimum of values diff(Z,Y) for Z, any possible node on the Wagner Network but outside the neighborhood of Fig. 6. From (16), (17), and (18)

$$s_C (A,B) = diff(C,X). \quad (19)$$

Hence by (18), we partition the evolution of A (or B) into two components, one, diff(C,X), being the number of derived steps shared by A and B in their evolution beyond C, the other, diff(X,Y), Y = A,B, being the number of derived steps unique to Y. Since the number of steps unique to A or B is minimized, the number of steps shared by A or B with other OTUs is maximized. Hence the Wagner Network is consistent with AIV if reversals are permitted.

DISCUSSION

The considerations of this paper lead to one central conclusion: while many numerical taxonomic methods are inconsistent with the phylogenetic approach, it is possible to construct a relatively small class of quantitative techniques that are valid under the premises of phylogenetics.

Classical phylogeneticists (for example Hennig, 1966; Brundin, 1968) have contended that similarity does not reflect phyletic relationship. Two aspects of their arguments are relevant here. First, they have based their conclusions on properties of preexisting measures of similarity, noting, for example, that "similarity" subsumes "synapomorphy," "symplesiomorphy," and convergence. Second, they have explicitly criticized only the notion that magnitude of similarity is indicative of absolute nearness of phyletic relationship, in effect ignoring the possibility that a much weaker correlation between similarity and phyletic relationship may still provide useful information.

We have defined new coefficients of "similarity," s(A,B) and s_E(A,B), through which it is possible to distinguish between similarity owing to derived character states and that owing to ancestral states. We do not contend that the magnitude of s(A,B) directly indicates the closeness of relationship between A and B. We interpret s(A,B) only as a relative measure of the amount of evidence favoring the hypothesis that A and B form a valid monophyletic group. The actual choice of a phyletic tree is left to an algorithm that effectively constructs the evolutionary hypothesis most in accord with available data. Thus only a weak connection between s or s_E and relationship is assumed.

With the exception of the techniques of Camin and Sokal (1965) and of Kluge and Farris (1969), most numerical taxonomic procedures do not measure similarity in a phylogenetically valid way. Indeed, some clustering techniques, such as Unweighted Pair Group Analysis (Sokal and Sneath, 1963), assume a strong correlation between similarity and relationship. Evolutionary interpretations of dendrograms generated by phenetic clustering procedures should, consequently, generally be viewed with skepticism.

The classical theory of phylogenetic systematics is well suited to situations in which character state trees and the identity of ancestral states can be "safely" established without a detailed prior taxonomic analysis. In more general cases it is often necessary to rely on taxonomic analysis to obtain some of the information needed for direct application of classical phylogenetic methods. The initial taxonomic analysis employed in such circumstances may not be the classical phylogenetic methods themselves—or, for that matter, their quantitative equivalents—since a *circulus vitiosus* might result. The initial methods employed must be consistent with the phylogenetic system, but may not make strong assumptions concerning character state trees. Wagner Trees seem particularly promising in this respect, both because of their close

formal analogies with phylogenetic methods and because of their great flexibility. Kluge and Farris (1969) successfully used Wagner Trees to identify character reversals and cases of miscoding. Some Wagner Tree algorithms (Farris, 1970) can construct "rootless" trees (networks) for which no input information on ancestral states is required. We are currently investigating methods through which such undirected networks may be used to infer the direction of evolution.

SUMMARY

Four premises of phylogenetics are abstracted from Hennig (1966) and used as axioms in constructing a quantitative taxonomic system equivalent to the phylogenetic method.

Under the axioms, each possible monophyletic group has a uniquely determined most recent common ancestor. The relation between ancestors and groups is the same as in Estabrook's (1968) formalization of the method of Camin and Sokal (1965).

The class of phyletic trees permissible under the phylogenetic axioms is contained in the set of trees producible by Estabrook's (1968) combinatoric method for finding most parsimonious trees.

There is thus a close connection between the phylogenetic approach and the most parsimonious tree techniques of numerical cladistics.

The Weighted Invariant Step Strategy (WISS) clustering procedure is described and shown to be consistent with phylogenetic principles.

The complete linkage analysis of Throckmorton (1968) should not be used for phylogenetic studies, since it can generate trees that are inconsistent with the premises of the phylogenetic system.

The statistic,

$$s(A,B) = (\tfrac{1}{2})\,(p(A,Q) + p(B,Q) - p(A,B)),$$

where Q is an OTU with all ancestral character states and p is the phenetic difference

between OTUs in the sense of Farris (1967), can be used to compute the number of derived steps shared by OTUs A and B. A *step*, x, is said to be *shared* by A and B if the states describing A and B are both derivable from *state* x.

Since s(A,B) can be computed from overall phenetic differences, it is not necessary physically to delete ancestral states from an analysis in order to achieve clustering by derived steps. Criticisms that most parsimonious tree techniques are phylogenetically invalid because they do not delete ancestral states, are consequently spurious.

The concordance of Throckmorton (1968) —a simple matching coefficient—should not be generally used in evolutionary investigations, since it can indicate incorrect clustering. It may be safely used when each character of a study has only two states.

Quantitative phylogenetic methods can be extended to the case in which no prior assumptions are made concerning direction of evolution. The Wagner method (Kluge and Farris, 1969) is consistent with phylogenetic principles in this case. Wagner Trees may be useful in inferring direction of evolution.

REFERENCES

BRUNDIN, L. 1968. Application of phylogenetic principles in systematics and evolutionary theory. In Tor Orvig, ed., Current problems of lower vertebrate phylogeny. Proc. Fourth Nobel Symposium, pp. 473–495.

CAMIN, J. H., AND R. R. SOKAL. 1965. A method for deducing branching sequences in phylogeny. Evolution, 19:311–326.

ESTABROOK, G. F. 1968. A general solution in partial orders for the Camin-Sokal model in phylogeny. J. Theoret. Biol., 21:421–438.

FARRIS, J. S. 1967. The meaning of relationship and taxonomic procedure. Syst. Zool., 16: 44–51.

FARRIS, J. S. 1969. A successive approximations approach to character weighting. Syst. Zool. 18:256–268.

FARRIS, J. S. 1970. Methods for computing Wagner Trees. Syst. Zool., 19:83–92.

KLUGE, A. G., AND J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool., 18:1–32.

HENNIG, W. 1966. Phylogenetic systematics. Translated by D. D. Davis and R. Zangerl. Univ. Illinois Press, Chicago, 263 pp.

SOKAL, R. R., AND P. H. SNEATH. 1963. Principles of numerical taxonomy. W. H. Freeman and Co., San Francisco, 359 pp.

THROCKMORTON, L. H. 1968. Concordance and discordance of taxonomic characters in *Drosophila* classification. Syst. Zool., 17:355–387.

WILSON, E. O. 1965. A consistency test for phylogenies based on contemporaneous species. Syst. Zool., 14:214–220.

Program in Ecology and Evolution, Division of Biological Sciences, State University of New York at Stony Brook, Stony Brook, New York, 11790, and Museum and Department of Zoology, The University of Michigan, Ann Arbor, Michigan, 48104.