



Taylor & Francis  
Taylor & Francis Group

## Society of Systematic Biologists

---

Cases in which Parsimony or Compatibility Methods Will be Positively Misleading

Author(s): Joseph Felsenstein

Source: *Systematic Zoology*, Vol. 27, No. 4 (Dec., 1978), pp. 401-410

Published by: Taylor & Francis, Ltd. for the Society of Systematic Biologists

Stable URL: <http://www.jstor.org/stable/2412923>

Accessed: 18-10-2017 16:57 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



JSTOR

*Society of Systematic Biologists, Taylor & Francis, Ltd.* are collaborating with JSTOR to digitize, preserve and extend access to *Systematic Zoology*

# CASES IN WHICH PARSIMONY OR COMPATIBILITY METHODS WILL BE POSITIVELY MISLEADING<sup>1</sup>

JOSEPH FELSENSTEIN

## *Abstract*

*Felsenstein, J. (Department of Genetics, University of Washington, Seattle, WA 98195) 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401-410.*—For some simple three- and four-species cases involving a character with two states, it is determined under what conditions several methods of phylogenetic inference will fail to converge to the true phylogeny as more and more data are accumulated. The methods are the Camin-Sokal parsimony method, the compatibility method, and Farris's unrooted Wagner tree parsimony method. In all cases the conditions for this failure (which is the failure to be statistically consistent) are essentially that parallel changes exceed informative, nonparallel changes. It is possible for these methods to be inconsistent even when change is improbable a priori, provided that evolutionary rates in different lineages are sufficiently unequal. It is by extension of this approach that we may provide a sound methodology for evaluating methods of phylogenetic inference. [Numerical cladistics; phylogenetic inference; maximum likelihood estimation; parsimony; compatibility.]

Parsimony or minimum evolution methods were first introduced into phylogenetic inference by Camin and Sokal (1965). This class of methods for inferring an evolutionary tree from discrete-character data involves making a reconstruction of the changes in a given set of characters on a given tree, counting the smallest number of times that a given kind of event need have happened, and using this as the measure of the adequacy of the evolutionary tree. (Alternatively, one can compute the weighted sum of the numbers of times several different kinds of events have occurred.) One attempts to find that evolutionary tree which requires the fewest of these evolutionary events to explain the observed data. Camin and Sokal treated the case of irreversible changes along a character state tree, minimizing the number of changes

of character states required. A number of other parsimony methods have since appeared in the systematic literature (Kluge and Farris, 1969; Farris, 1969, 1970, 1972, 1977; Farris, Kluge, and Eckhardt, 1970) and parsimony methods have also found widespread use in studies of molecular evolution (Fitch and Margoliash, 1967, 1970; Dayhoff and Eck, 1968; see also Fitch, 1973). Cavalli-Sforza and Edwards (1967; Edwards and Cavalli-Sforza, 1964) earlier formulated a minimum evolution method for continuous-character data.

An alternative methodology for phylogenetic inference is the compatibility method, introduced by Le Quesne (1969, 1972). He suggested that phylogenetic inference be based on finding the largest possible set of characters which could simultaneously have all states be uniquely derived on the same tree. The estimate of the phylogeny is then taken to be that tree. While Le Quesne's specific suggestions as to how this might be done have been criticized by Farris (1969), his general approach, which is based on Camin and Sokal's (1965) concept of the compatibility of two characters, has been made rigorous and extended in a series of papers by G. F. Estabrook, C. S. Johnson, Jr., and F. R. McMorris (Estabrook,

---

<sup>1</sup> This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately-owned rights.

1972; Estabrook, Johnson, and McMorris 1975, 1976a, 1976b; Estabrook and Landrum, 1975).

There has been relatively little examination of the properties of parsimony or compatibility methods as methods of statistical inference. Farris (1973, 1977) has shown that a number of different parsimony methods produce maximum likelihood estimates of an "evolutionary hypothesis" consisting of a phylogeny along with the reconstructed states of the characters in a large number of ancestral populations. However, when the object is to estimate only the phylogeny, the Camin-Sokal method has not been proven to give a maximum likelihood estimate except in the case when the probabilities of change in the character states are known to be small (Felsenstein, 1973).

For a given probabilistic model of evolution, one can construct a maximum likelihood estimate of the phylogeny, given the observed data on a set of discrete characters. Phylogenies constructed by the proper maximum likelihood method typically have the property of *consistency*. A statistical estimation method has the property of consistency when the estimate of a quantity is certain to converge to its true value as more and more data are accumulated. The purpose of this paper is to show that parsimony methods (as exemplified by the criterion of Camin and Sokal and by Farris's unrooted tree method) as well as compatibility methods do not possess the property of consistency in all cases. This is done by constructing a particular three-species case in which lack of consistency can be proven, a case in which parallel evolution is relatively probable. In finding such a case, we have thereby also shown that Farris's (1973) maximum likelihood estimate of the "evolutionary hypothesis" can give an inconsistent estimate of the phylogeny, since it always gives the same estimate as a parsimony method. Although it had been suspected that Farris's estimate of the phylogeny might be inconsistent, it was previously known only that it was not the same as

direct maximum likelihood estimation of the phylogeny (Felsenstein, 1973), and no actual proof of its inconsistency had been made.

The result may be regarded as warning us of the weakness of parsimony and compatibility methods. Alternatively, the conditions which must hold in order to have lack of consistency may be regarded as so extreme that the result may be taken to be a validation of parsimony or compatibility approaches. Readers must decide for themselves. In either case the conclusion reached will have the merit of being based on an examination of the properties of phylogenetic methods when considered as methods of statistical inference. Systematists may be tempted to reject this sort of attempt to evaluate phylogenetic methods by the criteria of statistical inference, particularly in view of the oversimplified models of evolution used here. It would seem difficult to take such a reaction seriously if unaccompanied by an attempt to erect a more adequate set of criteria, or to use the present criteria to examine more realistic models of evolution.

To show that a parsimony or compatibility method does not yield a consistent estimate of the phylogeny, it is not sufficient simply to show that it does not yield a maximum likelihood estimate. There are many examples known in statistics of consistent estimation methods which are not maximum likelihood estimates. For example, in samples drawn independently from a normal distribution, the maximum likelihood estimate of the mean of the underlying normal distribution is the sample mean. But the sample median is also a consistent estimator of the true mean. As more and more points are collected, it too will approach the true mean. By analogy to this case it might be argued that, although parsimony and compatibility estimates of the phylogeny are not maximum likelihood estimates, they do provide consistent estimates of the phylogeny. While this will often be the case, we shall see that this conjecture is not always true.

THE EXAMPLE

The example involves characters each of which has two states, 0 and 1. The ancestral state in each character is 0 and the derived state is 1. It is possible for the state of a population to change from 0 to 1, but not to revert from state 1 to state 0. Suppose that we have observed three species A, B, and C and that the (unknown) true phylogeny is as given in Fig. 1. Once a character is in state 1 at the beginning of a segment of the tree, it will not change thereafter, so that all we need to know for each segment is the probability that a character which is in state 0 at the beginning of the segment will have changed to state 1 by the end of the segment. These probabilities are assumed to be the same for all characters in this particular case; they are the quantities P, Q, and R shown in Fig. 1 next to the segments. In this particular case, the probabilities of change are assumed to be the same in segments II and IV of the tree, and the same in segments III and V. This is done purely to make the algebra easier: this assumption could be relaxed somewhat without altering the qualitative conclusions. It is important to realize that the constancy of P, Q, and R from character to character, and the differences between them from segment to segment, amount to strong assumptions about the biological situation. The differences in the probability of change may be due to the segments' being of different length in time (so that the tip species are not contemporaneous). Alternatively, they may be due to differences in the rate of evolution per unit time, differences from segment to segment of a sort which affect all characters. This amounts to the assumption that there are true differences in the overall rates of evolution of different lineages.

In Fig. 1 the segments of the tree are also numbered with Roman numerals. Knowing the probability of 0 → 1 change in each segment, we can easily obtain the probabilities of each of the possible combinations of states in the tips. For example, for the three tip species to be in

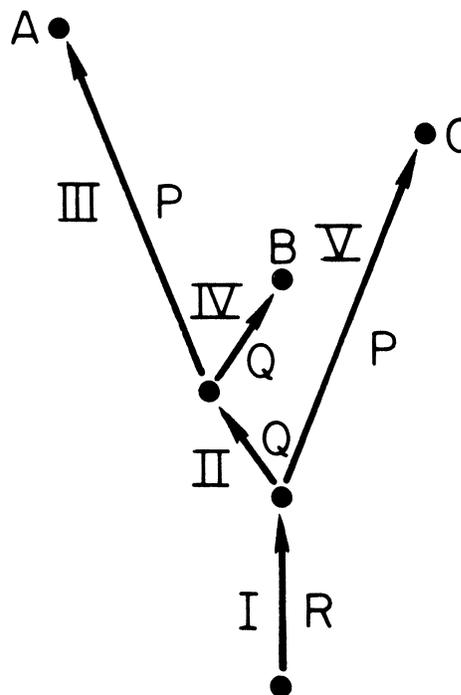


FIG. 1.—An evolutionary tree with three tip species. The segments of the tree are numbered I through V, and next to each is shown the probability of change from state 0 to state 1 in the segment.

states 1, 1, and 0 respectively, there must have been no change from state 0 in segments I and V. There may have been a 0 → 1 change in segment II, or else no change in that segment but 0 → 1 changes in both segments III and IV. The probability of observing states 1, 1, and 0 is thus

$$P_{110} = (1 - R)[Q + (1 - Q)PQ](1 - P) \tag{1}$$

Similarly, we can compute the probabilities of all eight possible configurations of character states:

$$P_{000} = (1 - P)^2(1 - Q)^2(1 - R) \tag{2a}$$

$$P_{001} = P(1 - P)(1 - Q)^2(1 - R) \tag{2b}$$

$$P_{010} = (1 - P)^2Q(1 - Q)(1 - R) \tag{2c}$$

$$P_{011} = P(1 - P)Q(1 - Q)(1 - R) \tag{2d}$$

$$P_{100} = P(1 - P)(1 - Q)^2(1 - R) \tag{2e}$$

$$P_{101} = P^2(1 - Q)^2(1 - R) \quad (2f)$$

$$P_{110} = (1 - P)[Q + (1 - Q)PQ](1 - R) \quad (2g)$$

$$P_{111} = PQ[P(1 - Q) + 1](1 - R) + R \quad (2h)$$

#### RESULTS OF THE CAMIN-SOKAL PARSIMONY METHOD

If we examine  $N$  characters in these three species, we can count how many of the characters are in each of the eight possible combinations: 000, 001, 010,  $\dots$ , 111. Let us call the resulting numbers of characters  $n_{000}$ ,  $n_{001}$ ,  $\dots$ ,  $n_{111}$ . We can use these numbers to discover what will be the result of applying the Camin-Sokal parsimony method to these data. When a character has the configuration 000, then no matter which phylogeny we propose, no changes of character state will be required to explain the evolution of this character along that phylogeny. There are four other configurations of the data which will require only one character state change to be assumed, no matter what phylogeny is postulated. These are 001, 010, and 100, which require one character state change on the segment of the evolutionary tree leading to a single species, as well as 111, which requires a single change at the root of the tree.

The remaining three configurations, 110, 101, and 011, will require different numbers of changes of state on different phylogenies. Let us represent the three possible bifurcating phylogenies as (AB)C, A(BC), and (AC)B, placing parentheses around monophyletic groups. On the phylogeny (AB)C, the configuration 110 requires only one change while the others require two changes. If we let

$$S = n_{001} + n_{010} + n_{100} + n_{111} + 2(n_{110} + n_{101} + n_{011}), \quad (3)$$

then (AB)C requires  $S - n_{110}$  changes of state to be assumed. By similar logic, A(BC) requires  $S - n_{011}$  changes, and (AC)B requires  $S - n_{101}$  changes. Which tree we estimate depends on which requires us to assume the fewest changes of character state. We can immediately see that the Camin-Sokal parsimony

method will estimate the correct phylogeny as (AB)C if and only if  $n_{110} \geq n_{101}$ ,  $n_{011}$ . When  $n_{011}$  is the greatest of these three numbers A(BC) will be the estimate, and when  $n_{101}$  is the greatest the estimate will be (AC)B. When there are ties for the greatest of  $n_{110}$ ,  $n_{101}$ , and  $n_{011}$ , there will be two or more possible estimates.

#### INCONSISTENCY OF THE RESULT

We assume that the  $N$  characters have evolved independently of one another, and have been chosen for study without regard to the configuration of their character states in these three species. Each character may be regarded as falling independently into one of the eight configurations 000,  $\dots$ , 111 with probabilities  $P_{000}$ ,  $\dots$ ,  $P_{111}$ . So the  $n_{ijk}$  are drawn from a multinomial distribution with these probabilities.

In such a case, an elementary application of the Strong Law of Large Numbers (e.g., Feller, 1957:243-244) tells us that as we let  $N \rightarrow \infty$ ,  $n_{ijk}/N \rightarrow P_{ijk}$  for all configurations  $ijk$ . In particular, this implies that as we score more and more characters,  $n_{110}$  will ultimately become larger and remain larger than either  $n_{101}$  or  $n_{011}$  if and only if  $P_{110} > P_{101}$ ,  $P_{011}$ . Whichever of these three probabilities is largest determines which of the three bifurcating phylogenies is certain to be the ultimate estimate as we accumulate more and more characters. Thus the condition for the Camin-Sokal estimate to have the property of consistency is simple: that  $P_{110}$  be greater than or equal to both  $P_{101}$  and  $P_{011}$ . Note in particular that if this condition does not hold, the consequences are striking: if, say  $P_{101} > P_{110}$ ,  $P_{011}$ , then as we accumulate more and more information the Camin-Sokal parsimony method is *increasingly certain to give the wrong answer*, in this case (AC)B.

We now examine the conditions on  $P$  and  $Q$  which are required in order to have inconsistency of the Camin-Sokal parsimony methods. The three quantities  $P_{110}$ ,  $P_{101}$ , and  $P_{011}$  are given by the expressions (2g), (2f), and (2d). Note that all of these

quantities contain a common factor of  $(1 - R)$ . Provided that  $R < 1$  (which we assume), this factor can be dropped. The condition  $P_{110} \geq P_{011}$  then becomes

$$(1 - P)[Q + (1 - Q)PQ] \geq P(1 - P)Q(1 - Q) \tag{4}$$

which simplifies to

$$Q(1 - P) \geq 0. \tag{5}$$

This will always hold, so in the present case it will always be true that  $P_{110} \geq P_{011}$ . Now we need only inquire whether  $P_{110} \geq P_{101}$ . This is the same as asking whether

$$(1 - P)[Q + (1 - Q)PQ] \geq P^2(1 - Q)^2, \tag{6}$$

which is equivalent to requiring that

$$0 \geq P^2(1 - Q) + PQ^2 - Q. \tag{7}$$

Let us view this as a quadratic equation in  $P$  whose coefficients depend on  $Q$ . Since  $1 - Q > 0$  (which we assume), the quadratic in (7) has a minimum at  $P = -Q/(1 - Q)$ . Since this is never positive, the positive values of  $P$  for which (7) is satisfied are those values of  $P$  below the point where the quadratic function is zero:

$$P \leq P_1 = \frac{-Q^2 + [Q^4 + 4Q(1 - Q)]^{1/2}}{2(1 - Q)} \tag{8}$$

$P_1$  is always a real number, so no complications arise. Figure 2 shows  $P_1$  plotted for values of  $Q$  between 0 and 1.  $P_1$  rises from 0 to 1 as  $Q$  goes from 0 to 1. Above the  $P_1$  curve is the region of values of  $P$  for which  $P_{110} < P_{101}$ .

This is the region in which the Camin-Sokal parsimony method is guaranteed to converge to the wrong estimate of the tree as we accumulate more and more data. Note that for every possible value of  $Q$  there is a range of values of  $P$  in which we will encounter this unpleasant behavior. A similar statement holds if we rearrange (7) to obtain limits on the values of  $Q$  as a function of  $P$ , so that for every value of  $P$  there is a range of  $Q$  values in which this unpleasant behavior

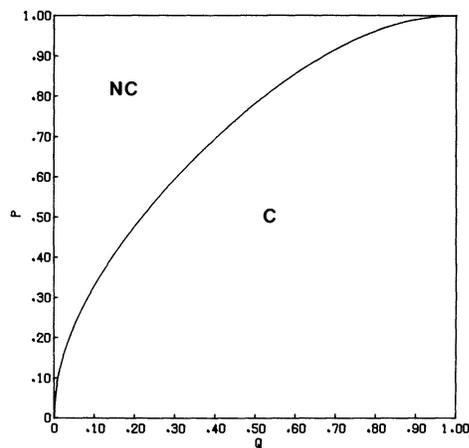


FIG. 2.—Values of  $P$  and  $Q$  for which the Camin-Sokal method fails to be consistent in the present case.  $C$  denotes the region of consistency,  $NC$  the region of inconsistency. Their boundary is the curve relating  $P_1$  to  $Q$ .

occurs. Note that for small  $Q$ , the condition (8) is closely approximated by

$$P \leq Q^{1/2}, \tag{9}$$

and for  $Q$  near 1 it is closely approximated by

$$P \geq 1 - (1 - Q)^2. \tag{10}$$

The effect of (8) is that the Camin-Sokal method will tend to fail when there is a sufficient disproportion between  $P$  and  $Q$ , which is the same as requiring that there be a sufficiently great disproportion between the lengths of the long and the short segments of the tree in Fig. 1.

In a previous paper (Felsenstein, 1973), I showed that for sufficiently small probabilities of evolutionary change, the Camin-Sokal method yields a correct maximum likelihood estimate of the phylogeny, and hence would be consistent. This might appear to be contradicted by (8) and (9), since these show that the Camin-Sokal method can be inconsistent even when  $P$  and  $Q$  are small. But my earlier proof involved holding the lengths (in time) of the segments of the tree constant while letting the rate of change in the

characters become small. This is equivalent to holding the ratio of P to Q constant while letting both approach zero. As will be apparent from dividing both sides of (9) by Q, when this is done the values of P and Q enter the region of consistency for sufficiently small values of P and Q, no matter what the (constant) ratio P/Q. So in this sense, the Camin-Sokal method works for sufficiently small rates of character state change.

#### COMPATIBILITY METHODS

It is a convenient fact that precisely the same three-species example also allows us to find conditions in which the compatibility methods yield an inconsistent estimate of the phylogeny. While the original approach of Le Quesne assumed that the direction of character state change was unknown, and could not be applied to a three-species case, the extensions of the compatibility approach by Estabrook, Johnson, and McMorris do allow us to make inferences in a three-species case when the direction of change on the character-state trees is known. For example, if two binary characters have the states (1,1), (1,0), and (0,1) respectively in the three species, then it is impossible for the transition  $0 \rightarrow 1$  to have taken place only once in each character on a branching phylogeny.

A pairwise consideration of all of the eight possible outcomes of the data will show that the outcomes 110, 101, and 011 are mutually incompatible, but that all other combinations are compatible. If we are trying to find the phylogenies suggested by the largest possible set of mutually compatible characters, these will include (AB)C if and only if  $n_{110} \geq n_{101}, n_{011}$ . Thus, the compatibility method for rooted binary character-state trees will give the same estimate as the Camin-Sokal method in the three-species case. We thus can apply all of the above conditions for inconsistency of the Camin-Sokal method to the compatibility approach. This allows the conclusion that consistency is not a general property of the compatibility methods, but must be proven

for specific probability models of evolution if it is desired.

#### UNROOTED WAGNER TREES

One of the most widely used parsimony methods has been Farris's (1970) method of inferring unrooted evolutionary trees under the assumption that character-state changes are reversible. The consistency of this method can be investigated by an extension of the present approach to a four-species case. This is necessary because there is only one possible unrooted tree in the three-species case, rendering it trivial. Figure 3 shows an unrooted tree with four species, A, B, C, and D. In order to more closely approximate the evolutionary model which underlies Farris's method, we assume that although the characters were originally in state 0, they have the same probability of reversion  $1 \rightarrow 0$  once they are in state 1, as they have of origination  $0 \rightarrow 1$  of state 1 when they are in state 0. Thus each segment of the evolutionary tree is characterized by a probability of character state change which applies equally to both forward change  $0 \rightarrow 1$  and reversion  $1 \rightarrow 0$ . Once again, we assume for simplicity that characters are independently sampled and all have the same probabilities. There are 16 possible data outcomes, 0000 through 1111. Once again, the outcome of applying Farris's parsimony method will depend only on the numbers of characters  $n_{0000}, \dots, n_{1111}$  having each outcome.

It is easy to show, along the same lines as before, that whether the unrooted tree obtained is of form (AB)(CD), (AC)(BD), or (AD)(BC) is determined by which of the three numbers  $n_{1100} + n_{0011}, n_{1010} + n_{0101}$ , and  $n_{1001} + n_{0110}$  is largest. It is not difficult to demonstrate that the exact placement of the root of the true tree will affect only the relative probabilities of obtaining 1100 and 0011, but will leave the total probability  $P_{1100} + P_{0011}$  unchanged, and similarly for  $P_{1010} + P_{0101}$  and  $P_{1001} + P_{0110}$ . Therefore, we need not specify the placement of the root on the (unknown) true tree to compute the probabilities which determine the outcome of

this parsimony method. Suppose that the true phylogeny is one whose unrooted form is given in Fig. 3. We may as well assume that the root is at the left-hand end of the central segment, and that all characters start there in state 0, as these assumptions do not affect  $P_{1100} + P_{0011}$  and the other relevant probabilities.

Considering the two possible character states at the right-hand end of the central segment, we find that

$$P_{1100} + P_{0011} = PQ[1 - Q]^2(1 - P) + Q^2P] + (1 - P)(1 - Q) \cdot [Q(1 - Q)(1 - P) + Q(1 - Q)P] \quad (11)$$

with analogous expressions for the other two relevant probabilities:

$$P_{1010} + P_{0101} = P(1 - Q) \cdot [Q^2(1 - P) + (1 - Q)^2P] + (1 - P) \cdot [Q(Q(1 - Q)P + Q(1 - Q)(1 - P))] \quad (12)$$

and

$$P_{1001} + P_{0110} = P(1 - Q) \cdot [Q(1 - Q)P + Q(1 - Q)(1 - P)] + Q(1 - P) \cdot [(1 - Q)^2P + Q^2(1 - P)]. \quad (13)$$

After some elementary but tedious algebra it can be shown from (12) and (13) that provided that  $Q \leq 1/2$ , which we assume,

$$P_{1010} + P_{0101} \geq P_{1001} + P_{0110}. \quad (14)$$

This establishes that when the true tree is as shown in Fig. 3, our estimate of the unrooted tree topology may converge to either (AB)(CD) or to (AC)(BD), but never to (AD)(BC) as we collect more and more characters. So to establish the consistency of the estimation of unrooted tree topology, we need only enquire whether

$$P_{1100} + P_{0011} \geq P_{1010} + P_{0101}, \quad (15)$$

which will be the condition for consistency. Using (11) and (12) we find after

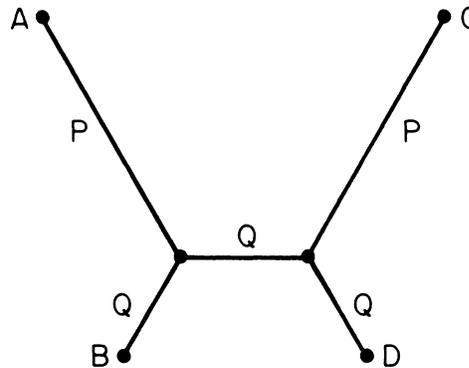


FIG. 3.—True unknown phylogeny (with root omitted) used to find cases in which unrooted Wagner tree parsimony methods will be inconsistent.

further tedious algebra that (15) is simply,

$$2P^2Q - P^2 + 2Q^3 - 3Q^2 + Q \geq 0 \quad (16)$$

which is

$$(2Q - 1)(P^2 + Q(Q - 1)) \geq 0. \quad (17)$$

Since  $Q \leq 1/2$ , (17) is simply

$$P^2 \leq Q(1 - Q), \quad (18)$$

a considerably simpler condition than (8). Note that when  $Q$  is small, (18) reduces to (9). Thus, all the statements about consistency in the Camin-Sokal case when  $P$  and  $Q$  are small are also correct in the case of unrooted Wagner trees.

#### DISCUSSION

We have seen that there are circumstances under which three different estimation methods are not statistically consistent, these being the Camin-Sokal parsimony method, the Estabrook-Johnson-McMorris compatibility method, and Farris's parsimony method for estimating unrooted Wagner trees. For small values of  $P$  and  $Q$ , the condition for inconsistency amounts to requiring that simultaneous changes on two long segments of the tree be more probable a priori than one change on a short segment. This amounts to requiring that parallelism of changes be more probable than unique

and unreversed change in an informative part of the tree (e.g., that simultaneous changes in segments III and V of the tree in Fig. 1 be more probable than a single change in segment II). This certainly seems like a reasonably intuitive condition for inconsistency. The advantage of the argument presented here lies not in leading to a particularly surprising conclusion, one that will cause abandonment of these parsimony and compatibility methods, but as a formal investigation of one of the statistical properties of phylogenetic inference methods.

The models employed here certainly have severe limitations: it will hardly ever be the case that we sample characters independently, with all of the characters following the same probability model of evolutionary change. Extending this analysis to more realistic evolutionary models will certainly be difficult. Yet the task must be undertaken: if inconsistency of a parsimony or compatibility technique is suspected, it does little good simply to point out that the evolutionary models employed here do not apply to the type of data being encountered in practice. That amounts to a confession of ignorance rather than validation of the inference method in question.

#### LIKELIHOOD METHODS

Methods of phylogenetic inference which entirely avoid the problem of statistical inconsistency are already known. Maximum likelihood estimation of the phylogeny is one of them. I have outlined elsewhere (Felsenstein, 1973) how this may be done. In the three species cases maximum likelihood estimation methods can easily be developed. The likelihood of a tree will simply be

$$L = \prod_{ijk} P_{ijk}^{n_{ijk}}, \quad (19)$$

where  $P_{ijk}$  is the probability of data configuration  $ijk$  and  $n_{ijk}$  is the number of characters having that configuration. Estimation is carried out by maximizing (19) over the unknown parameters of the evolutionary model (such as P and Q in equations [2]). This is done for each tree topology, and the final estimate consists of the topology and the evolutionary parameters which yield the highest likelihood. (Note that despite the connotations of the term, the likelihood of a tree is *not* the probability that it is the correct tree.)

When there are larger numbers of species, the number of possible data configurations (the number of terms  $n_{ijk}$ ) in each character becomes so large that it is impractical to use equation (19). I have presented elsewhere (Felsenstein, 1973) an algorithm for evaluating the likelihood of a tree which avoids this difficulty.

Maximum likelihood estimates are not desirable in themselves, but because they have desirable statistical properties such as consistency and asymptotic efficiency. In the case of discrete multistate characters under the sorts of evolutionary model considered here, it can be shown quite generally that the maximum likelihood estimation procedure has the property of consistency. In particular, in the case of the tree shown in Fig. 1, it will be a consistent method whatever the values of P, Q, and R.

The reader familiar with the paper of Farris (1973), which establishes a general correspondence between parsimony methods and maximum likelihood methods may be puzzled at this stage: if parsimony methods are maximum likelihood methods, why have the two been described here as separate methods? Why is one sometimes not consistent while the other is always consistent? This paradox is resolved once one recalls that the maximum likelihood methods used by Farris are different from those described in Felsenstein (1973) and here. Farris used the maximum likelihood method to estimate not only the parameters of the evolutionary tree, but also the states of the characters in a large number of ancestral populations. When this latter kind of maximum likelihood estimate is made, the number of parameters being estimated rises without limit as more characters are examined.

From the point of view of estimating the phylogeny, these extra parameters are "nuisance" parameters. As a result of their presence, the ratio between the number of data items and the number of parameters does not increase indefinitely as more characters are added. It is in situations such as this that maximum likelihood methods are particularly prone to lack of consistency, as I have previously pointed out (Felsenstein, 1973). Indeed, the present results establish that there are conditions under which Farris's likelihood method (giving the same results as a parsimony method) fails to be consistent.

#### PERSPECTIVE

The weakness of the maximum likelihood approach is that it requires us to have a probabilistic model of character evolution which we can believe. The uncertainties of interpretation of characters in systematics are so great that this will hardly ever be the case. We might prefer to have methods which, while not statistically optimal for any one evolutionary model, were robust in that they had reasonable statistical properties such as consistency for a wide variety of evolutionary models. The present results establish that parsimony and compatibility methods can fail to be consistent if parallelism is expected to occur frequently. This helps establish that they do not yield maximum likelihood estimates. However, they pass the test of consistency when parallelism is rare. This leaves them as viable candidates for robust methods. Establishing that robustness (or disproving it) by examining a wider range of models is a daunting task, but it must be undertaken. If phylogenetic inference is to be a science, we must consider its methods guilty until proven innocent.

#### ACKNOWLEDGMENTS

I wish to thank a reviewer for penetrating comments which were of great assistance. This work was supported by Department of Energy Contract No. EY-76-S-06-2225 5 with the University of Washington.

#### REFERENCES

- CAMIN, J. H., AND R. R. SOKAL. 1965. A method for deducing branching sequences in phylogeny. *Evolution* 19:311-326.
- CAVALLI-SFORZA, L. L., AND A. W. F. EDWARDS. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 32:550-570 (also in *Amer. J. Human Genetics* 19:233-257).
- DAYHOFF, M. O., AND R. V. ECK. 1968. Atlas of protein sequence and structure 1967-1968. National Biomedical Research Foundation, Silver Spring, Maryland.
- EDWARDS, A. W. F., AND L. L. CAVALLI-SFORZA. 1964. Reconstruction of evolutionary trees. In Heywood, V. H., and J. McNeill (eds.), *Phenetic and phylogenetic classification*. Systematics Association Publication No. 6, London. pp. 67-76.
- ESTABROOK, G. F. 1972. Cladistic methodology: a discussion of the theoretical basis for the induction of evolutionary history. *Ann. Rev. Ecol. Syst.* 3:427-456.
- ESTABROOK, G. F., C. S. JOHNSON, JR., AND F. R. MCMORRIS. 1975. An idealized concept of the true cladistic character. *Math. Biosci.* 23:263-272.
- ESTABROOK, G. F., AND L. LANDRUM. 1975. A simple test for the possible simultaneous evolutionary divergence of two amino acid positions. *Taxon* 24:609-613.
- ESTABROOK, G. F., C. S. JOHNSON, JR., AND F. R. MCMORRIS. 1976a. A mathematical foundation for the analysis of character compatibility. *Math. Biosci.* 29:181-187.
- ESTABROOK, G. F., C. S. JOHNSON, JR., AND F. R. MCMORRIS. 1976b. An algebraic analysis of cladistic characters. *Discrete Math.* 16:141-147.
- FARRIS, J. S. 1969. A successive approximations approach to character weighting. *Syst. Zool.* 18:374-385.
- FARRIS, J. S. 1970. Methods for computing Wagner trees. *Syst. Zool.* 19:83-92.
- FARRIS, J. S., A. G. KLUGE, AND M. J. ECKHARDT. 1970. A numerical approach to phylogenetic systematics. *Syst. Zool.* 19:172-189.
- FARRIS, J. S. 1972. Estimating phylogenetic trees from distance matrices. *Amer. Nat.* 106:645-668.
- FARRIS, J. S. 1973. On the use of the parsimony criterion for inferring evolutionary trees. *Syst. Zool.* 22:250-256.
- FARRIS, J. S. 1977a. Phylogenetic analysis under Dollo's Law. *Syst. Zool.* 26:77-88.
- FARRIS, J. S. 1977b. Some further comments on Le Quesne's methods. *Syst. Zool.* 26:220-223.
- FELLER, W. 1957. An introduction to probability theory and its applications. Volume I. Second edition. Wiley, New York.
- FELSENSTEIN, J. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Syst. Zool.* 22:240-249.
- FITCH, W. M., AND E. MARGOLIASH. 1967. Construction of phylogenetic trees. *Science* 155:279-284.

- FITCH, W. M. 1973. Aspects of molecular evolution. *Ann. Rev. Gen.* 7:343-380.
- KLUGE, A. G., AND J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18:1-32.
- LE QUESNE, W. J. 1969. A method of selection of characters in numerical taxonomy. *Syst. Zool.* 18:201-205.
- LE QUESNE, W. J. 1972. Further studies based on the uniquely derived character concept. *Syst. Zool.* 21:281-288.
- MORRIS, F. R. 1977. On the compatibility of binary qualitative taxonomic characters. *Bull. Math. Biol.* 39:133-138.

*Manuscript received June 1977*

*Revised August 1978*