B.U.N.K. Volume 3, Number 1 (1986)

Cladistic Statements as Cladistic Formulae or How to Hate Cladistics and Love Mathematics

F.C. Thompson and D.R. Whitehead (1,2,3,4,5,6)
Systematic Entomology Laboratory IIBIII
Agricultural Research Service, USDA
c/o U.S. National Museum
Washington, D.C.
20560 U.S.A.

Interpretive Statement (7,8)

A simple method of presenting cladistic information is given. As cladistic hypotheses provide the basic synthetic building blocks of biology, the use of our method will: result in significant savings in the cost of storage and dissemination of systematic data.

Abstract

Exactly as above, with various modifications and ramifications as required.

Text

Is there a simple, shorthand way to present cladistic hypotheses? Certainly. Now that the cat is out of the bag, we merely belabor the obvious.

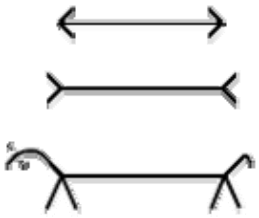
If systematic papers are to have a patina of scientific credibility, one or more games must be played. Cladistics is, de rigueur, a favored game. An elaborate cladogram (a sort of defoliated tree), however, betrays at best a mechanical tedium, and at worst, sheer ignorance. It requires ridiculous amounts of time to prepare, perhaps an artist's services, and space and money to publish.

What substitute will better serve to lend scientific credibility? Why, surely, what can be more elegant and eloquent than a "mathematical" formula? A simple cladistic hypothesis can be stated thus: Group X = Taxa (a+b)+(c+(d+e)). In plain English, a and b as well as d and e are sister groups, with group d+e being the sister of c and that group in turn being the sister of group a+b. ln the equation, the left hand element identifies the group being considered and the right, a cladistic formula, explains the relationships within that group. Even with names spelled out, relationships of a small number of taxa can be clearly expressed in but one or two printed lines. (9)

We restrict further discussion to end-point taxa in dichotomous models. Those who explore mathematical games (or gams, as that is what the typer tried to say (10)) may extend these ideas to more complex models. Cladistic formulae can be used to express any kind of branching relationships.

Our game is to explore perception. For our models to work (11), they must be dichotomous. Who would say that the following two lines are not identical, for instance?

Suppose, however, these lines are presented as: Wagnerian or Wagging networks? Which then is the longest? Or, by rules of parsimony (cheapskatedness), the shortest?



One fault of the standard cladogram is that its visual impact is linear, hence misleading. To overcome this each branching point or node should be considered as fully rotatable. Considered thusly, interesting numbers begin to emerge. For sets of 2, 3, 4, 5, and 6 end-point taxa, there are respectively; 1, 3, 15, 105, and 945 different sets of relationships regardless of rotatability of the nodes, and there are respectively; 2, 4, 8, 16, and 32 different ways to view each single cladogram when rotatability is considered (i.e. a different linear arrangement of the taxa or "isomorph"). Thus for the five-taxon example cited previously, there are 104 other possible hypotheses, and there are also 15 other ways to express the proposed hypothesis; in all, relationships among the five end-point taxa in a strictly dichotomous system may be expressed in 105 x 16 cladograms, or a total of 1680 visually different trees.(12)

The formula given previously for Group X is one hypothesis, and it is one of 16 possible isomorphs. Per se, it has no advantage over a tree (except that it may not be botanical, so therefore it can apply to beetles). Lacking however, is the confusion of lines inherent in the tree; the hypothesis is reduced to symbolic logic, wherein elements shift mentally without recourse to repeated diagrams.

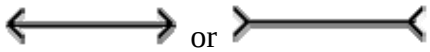
Our point is clear. Smash the optical illusion of linearity. Our reason should also be clear. Accommodate optically contradictory but compatible morphoclines. A few sample morphoclines will illustrate the point. In the formula: Group X = Taxa (a+b)+(c+(d+e)), relationships are indicated by morphoclines in characters 1, 2, and 3, thus:

1. (a, b) → c → d → e
2. (a, b) → (d, e) → c
3. (c, d, e) → a → b

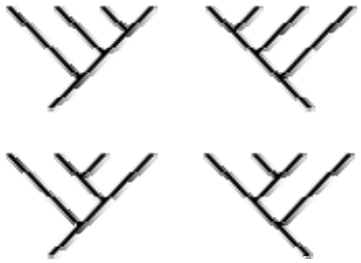
If the hypothesis is drawn directly as a tree, the illusion is that taxa a and e are least related, because they are on opposite sides of the tree. In the isomorphic formula: Group X Taxa (c + (d + e)) + (a + b), rotatability is obvious; in the derivative tree, the illusion is that taxa a and e do not appear to be least related.

Whenever only a bare cladistic hypothesis is intended, the shorthand formula is the ultimate in elegance.

To summarize we ask two questions: (1), which network is longer?

 or

And, (2), which of the following trees are the same?



The answers are (1), neither, and (2), all. If you got both right, good: you are a true cladist, as reality means more than appearance. If you got both wrong, bad: you are a hopeless pheneticist, as appearances (overall similarity) are more important to you than reality. If you got one right, you are confused, uncertain as to whether reality or appearances are more important, hence belong to the eclectic school. And school is out!

---

1. Who is senior and who is junior? For this fantastically important contribution we flipped a coin: The senior author is junior and the junior, senior. At least in age. Beware of illusions: they lead to misconceptions.
2. Anheuser-Busch, Coors, Labatts, Miller, Molson and Schlitz brewed the appropriate intellectual stimuli, not to mention illusions, misconceptions, and eventual delusions. Bob Gordon (the man who studies the beetles that roll him), John Burns, Tom Henry (bugs suck), and every other Tom, Dick, and Harry provided effervescent comment. We apologize to any contributors lost in the foam of time. We do not thank those who thought the topic stupid.
3. Please don't request reprints. At this writing it hasn't been printed yet so there aren't any reprints. Try preprints instead. FCT thought this comment stupid. But as in Penelope Ashe's "Naked Came The Stranger", stupid comments stay in and only good stuff is edited out. So why is the article so long? Illusion?
4. This paper has not been approved for publication. In fact, it was squelched. So, here it is: It is an illusion; misconceived; aborted; miscarried. Justice!
5. 0riginally (1965) prepared for Sitzungsberichte der Historische Investigierungen und Technik, but reassinged to B.U.N.K. by invitation when no S.H.I.T. could be found. It (S.H.I.T.) was illusory.
6. We don't know anything about beetles (FCT, Dipterist and Computerologist; DRW, Diplopodologist, Hemipterist, Weevilologist, ad nauseum), but we assume, trust, hope, and pray that our observations somehow are of interest.
7. One reason these things don't get published is that we seek to avoid being an embarrassment to congressmen. So we seek to avoid perusal by congressmen. This is a "coffee break" article. We have not squandered (spent, maybe- squandered, no) research time nor government funds on it. Hence, it is B.U.N.K.
8. Agricultural Research Service papers receive close scrutiny at all levels. This one defies all levels of acceptance so was merely peeked at. Praise the Willi.
9. lt gets messy when X = (a+(b+c))+(d+(e+((f+g) +((h+(i+j)))))). We probably gummed this up. Too much math. You gotta count the parentheses, and balance the equation (the parentheses define nested sets). At least the left is right, even if the right is wrong
10. Actually we may not be true cladists. because the typer insists on "caldist". Perhaps this is just a tpyo.

11. Free of charge.
12. Since there is only one evolutionary history, there must only be one correct tree. Which of the 1680 do you like?
