Logic in systematics

O. RIEPEL1, M. RIEPEL2 and L. RIEPEL3

Abstract

Systematic biologists attempt to infer the taxonomic relationship of one species to another based on the available evidence. Phylogenetic systematists demand that these species relationships reflect evolutionary history (they expect their taxa to be monophyletic). Just exactly how this is to be achieved remains a subject of debate. There are many different kinds of evidence, and many different ways of inferring taxonomic relationships (plus evolutionary history) from them. In this paper, we argue that one such way of inferring species relationships, the hypothetico-deductive method, proves a bad fit with phylogenetic systematics because it requires an excessively strong assumption of the relationship that obtains between hypotheses of descent and the available evidence.

Key words: Phylogenetic systematics – hypothetico-deductionism – falsification – Karl Popper

Introduction

In a High School paper written at the age of 18, Hennig (1978) lamented the eclipse of systematics by other branches of biology, which claimed to resemble the so-called ‘hard sciences’ such as physics and chemistry more closely. In the same paper, he urged systematists to develop their discipline into a ‘complete and important science in itself’ (Hennig 1978, p. 199). A few years later, Hennig attempted to do just that by publishing his Grundzüge einer Theorie der Phylogenetischen Systematik (1950) as a contribution to Bertalanffy’s (1932) Systematik. In an early review of Hennig’s (1950) book, Günther (1956, p. 41) praised its author for having ‘empirio-critically reconstructed... the natural system... [in a way] that captures the relative relationships of forms’. As such, the review continued, Hennig had succeeded in delivering ‘the proof for the theory of descent’. Günther also praised Hennig’s text for eschewing everything ‘but scientific, i.e. empirio-critically secured insights in its explanation of the enigmatic system’.

A half-century earlier, Lenin (1909 [1959]) famously used the term empirico-critical (originally introduced by Richard H.L. Avenarius: R. Vane-Wright, pers. comm.) in his critique of Ernst Mach’s philosophy. This point owes its significance to the fact that it was Mach who would go on to become the intellectual godfather of logical positivism (Janik and Toulmin 1973). Indeed, the Vienna Circle was originally founded under the name ‘Ernst Mach Verein’ (Joergensen 1970, p. 850; see Stadler 1997, for more detail). In addition, Moritz Schick, who served as the first president of the Mach Verein, was also one of van Bertalanffy’s (1932) advisors. Thus, by complimenting Hennig (1950) on his strictly empirico-critical approach, we may surmise that Günther (1956) was in fact alluding to the logical empiricist background of Hennig’s (1950) phylogenetic systematics.

Hennig’s (1950) systematics was inspired by the methodology previously outlined by Zimmermann (1937, 1943, see Donoghue and Kadereit 1992), whom Hennig (1950, p. 14; Hennig 1966, p. 10) considered ‘one of the best modern theoreticians of systematic work’. Zimmermann (1954, p. 48) also and explicitly characterized his own approach to systematics as ‘empirio-critical’, while he trumpeted its capacity to purge systematics of intuitions, myths and fairytales (Zimmermann 1943, p. 43), ‘Astrology changed into astronomy, alchemy changed to become chemistry. Only in phylogenetics has this change remained incomplete’ (Zimmermann 1943, p. 20). Given their logical empiricist background, it should not come as a surprise that formal logic played a central role in Zimmermann’s (1937, 1943), Hennig’s (1950) and Günther’s (1956) approaches to systematic biology (see, for example, Dupuis 1984, p. 19). After all, no less a leading figure of logical positivism than Rudolf Carnap had characterized the essence of philosophy as the clarification ‘of the concepts and sentences of science by logical analysis’ (cited in Haack 2003, p. 32).

A central concern of logical empiricism was the combination of science and logic (Stroll 2000, p. 54): its focus was on the confirmation or falsification of scientific theories and explanations, as well as the clarification of their logical structures (Kitcher 1993, p. 5). The most widely accepted picture of the scientific method (see Kitcher 1993 for a discussion of this misconception of scientific practice) was the hypothetico-deductive (H-D, also known as deductive-nomological) model of scientific explanation (Hull 1973, p. 30; for a detailed modern discussion of the H-D model see Lipton 2004). The H-D model was perhaps most famously championed by Hempel (1965) and Popper (1976a, 1992). It begins with the postulation of a universal law (or theory) that, in conjunction with a certain set of initial or boundary conditions, logically entails observation statements (instantiations of that law). If the observation statement meets the world of experience, the theory (law) has been confirmed; if not, the theory (law) has been falsified. To allow for the deduction of such observation statements, the laws (theories) of science must take the form of universal statements (Popper 1976a, 1992; see also Hull 1973, p. 31), i.e. they must be so-called ‘all-statements’ (e.g. it follows from the law ‘all ravens are black birds’ plus the initial condition 'here is a raven' that 'here is a black bird'; for further discussion see Rieppel 2004). For logical empiricists, the paradigmatic example of a rapidly maturing science was theoretical physics, which exhibits a very near approximation of the H-D structure (for a more recent, opposing view, see Galison 1987, 1997). In his sketch of a Theoretical Biology,
Bertalanffy (1932, p. 25) set similar goals for the life sciences. He divided biology into three worlds: the ‘first world’ is one of description and simple comparison yielding order (classification), the ‘second world’ is that of biological rules, and the ‘third world’ constitutes a hypothetico-deductively structured ‘theoretical biology’ that delivers the laws of biology. It is against this background that we must read Hennig’s (1950, p. 23) comment on the ‘recently increasing attempts’ to deduce evolutionary theory from universal principles.

Both Zimmermann (1937, 1943, 1954) and Günther (1956) stressed the need for logical (deductive) inference in systematics, while Schmitt (2001, p. 343) praised Hennig for having transformed systematics ‘from a skill or an art to a truly scientific method… which justly takes its place in a hypothetico-deductively structured science corresponding to the picture of science painted by Popper’ (see also Ax 1984; Wiesemüller et al. 2003). Günther (1967, p. 18) found the ‘syllogistic form of argument’ to endow the encephalic hierarchy generated by phylogenetic systematics with the power to issue successful predictions: ‘If that hitherto unknown animal which is presented to me is a bird, then it will be possible to reliably identify a whole series of characters both on it and in it’. Günther’s (1967, p. 18) represents a ‘conditional’ (an ‘if … then’) statement that expresses material implication; it corresponds to what Hempel (2002, p. 53) called an ‘elliptical’ formulation of the H-D model of scientific explanation (it is elliptical in that it does not explicitly state, but rather implicitly assumes the covering law and the boundary conditions under which the logical derivation of those predictions are possible).

The covering law in that example would be a universal statement of the form ‘all birds have c₁, c₂,…’ where c stands for characters such as feathers, beaks, etc. The boundary condition is that a hitherto unknown animal (A) is placed before Günther. The full argument then runs as follows: ‘If A is a bird, A has feathers; A is a bird, therefore, A has feathers’ (modus ponens). Or alternatively: ‘If A is a bird, A has feathers; A has feathers; therefore, A is not a bird’ (modus tollens). Modus tollens, what can be called the ‘mode of denying’, is a form of deductive inference that was widely used by Popper (1976a, 1992) because of its undeniable power and strength: it yields necessary conditions (sufficient conditions are specified by modus ponens; see discussion in Brenner 1993). Necessary conditions can be understood as those that hold in any possible world, i.e. they hold always and everywhere throughout the universe. Note how modus tollens rests on the truth-conditions of material implication: if the consequent is denied (i.e. if it fails to obtain), then, for the sentence to remain true, the antecedent must necessarily be false. One consequence of this is that there cannot be any strict formulation of modus tollens, or of the relation of material implication more generally, in probabilistic terms (Sober 1988). It is because of its strength that Popper (1979) found modus tollens most applicable to the deductively structured science of theoretical physics, whereas it cannot be easily made to fit the historically contingent nature of evolutionary biology (Popper 1976b). The reason for this is that the observable state of affairs that we seek to explain must be logically (i.e. deductively) entailed by the covering law, a condition that is difficult to meet for historically contingent processes. Today, we know that Günther’s (1967) covering law indeed makes demands that evolutionary biology cannot meet; after all, we know of organisms that have feathers but are not birds, viz. the Chinese feathered dinosaurs (of course the covering law can be saved ad hoc by debating the ‘meaning’ of the term ‘feather’ and denying its presence in theropods: New Scientist 12 April 1997).

A similar appeal to the H-D (or deductive-nomological) model, underwritten with reference to Popper, survives in contemporary textbooks on phylogenetic systematics (Wiesemüller et al. 2003, p. 87). The example given by Wiesemüller et al. (2003, p. 87) uses material implication in the expression of the biological covering law: ‘If an animal has a placenta, then it belongs to the taxon Placentalia’. Since we have a boundary condition that specifies there are no animals with a placenta endemic to Australia, we can deduce the conclusion that therefore there are no animals belonging to the taxon Placentalia in Australia. In this example, the covering law again is too strong to accommodate the historical contingency of the evolutionary process: there are endemic animals in Australia that do have a placenta, yet they are not placental mammals, but lizards instead (Weekes 1935; Blackburn 1992; notice, however, that here again the covering law can be saved ad hoc by debating the ‘meaning’ of the term ‘placenta’).

Hennig (1966, p. 128; emphasis added) explored the idea that systematics must be content with deducing phylogenetic relationships from the degree of similarity between species, but found that ‘such a way of reaching conclusions is in fact very problematical’. The characterization of his research program as H-D, i.e. Popperian, by Schmitt (2001) notwithstanding, Hennig (1966, p. 128) concluded that ‘…this is not the true method of phylogenetic systematics’. Hennig (1974, p. 281) considered the appeal to predictiveness a ‘purely formal and essentially futile adherence to a certain group of scientific theories to which does not belong any theory of biological systematics’. The ‘group of scientific theories’ he had in mind was the universal statements required to get the H-D model of explanation off the ground. In spite of these early warnings, some cladists forged ahead with an attempt to set up a system that is as consistent as possible within the H-D system as developed by Popper (Gaffney 1979, p. 85; see also Farris 1983). Since then, Sober (1988) has argued convincingly that Popperian hypothetico-deductivism wedded to the modus tollens form of argument cannot be applied to phylogeny reconstruction (this point was later echoed by Rieppel 2003). But again, some authors continue to maintain that phylogenetic analysis using parsimony is H-D in a Popperian sense (Kluge 2001, 2003, 2004; Crother 2002). Kluge (2003) in particular has defended the deducibility of species relationships. It is the purpose of this contribution to investigate the question of whether phylogenetic systematics (cladistics) does, in fact, correspond to a hypothetico-deductively structured science as the one sketched by Popper (1976a, 1992).

Deduction in systematics

Popper (1976a, p. 2) opened the sixth German edition of ‘The Logic of Scientific Discovery’ with a quote from Kant: ‘Among the deductive inferences, which conclude from consequences to causes, modus tollens delivers proof not only very severely, but also most elegantly. For even if only one single false consequence can be inferred from a statement, then that statement is false’ (our translation). Popper’s falsificationism is built on the modus tollens form of argument (Popper 1976a, section 18; Lakatos 1974), one of many classic deductive inference schemes (see also Ball 1982). Popper’s (1976a, 1992) adoption of modus tollens for his falsificationist programme

© 2006 The Authors JZS 44(3), 186–192
Journal compilation © 2006 Blackwell Verlag, Berlin
highlights two central points: application of *modus tollens* yields necessary conditions; therefore, no probabilification of *modus tollens* is possible (Sober 1988).

Among the most common forms of deductive reasoning (Brenner 1993, p. 6) are *modus ponens*: ‘If P, then Q; P; therefore Q’; symbolized as \([P \rightarrow Q], P : \therefore Q\); and *modus tollens*: ‘If P, then Q; not-Q; therefore not-P’, symbolized as \([P \rightarrow Q], \sim Q : \therefore \sim P\). If deduction is at all applicable to systematics, then *modus ponens* and *modus tollens* should be applicable as well. We therefore propose to test the claim, made in a Popperian context, that species relationships are deducible (Kluge 2003) by examining the applicability of the *modus tollens* form of argument to systematics, or more precisely, to the finite universe of a three-taxon statement. This possibility has previously been explored by Cartmill (1981), whose arguments were criticized by Farris (1983). Ball’s (1982) likewise investigated the use of *modus tollens* in systematics, yet found it too strong to be useful due to the contingency of the historical process that may generate convergence, reversals, and parallelisms. Only if independent criteria were available to reliably distinguish ‘true’ synapomorphy from homoplasy would a Popperian form of falsificationism be applicable to systematics (Ball 1982; see also Rieppel 2003, 2004). Ball’s (1982) conclusion was that whereas such a reliable distinction between ‘true’ synapomorphy and homoplasy may be very difficult in practice, it might not be impossible in principle, such that a Popper-style falsificationism could, indeed, apply to phylogenetic systematics. Here, we offer an alternative formulation of the application of *modus tollens* to a three-taxon statement derived from an argument by Kluge (2003). We will show that this application is logically valid given a limited set of propositions, but that these propositions result in consequences that cannot be reconciled with a falsificationist understanding of systematics, and that would most likely be rejected by all systematists for general reasons.

The *modus tollens* applied to a three-taxon statement

The three-taxon statement (Gaffney 1979) exemplifies the cladistic approach to phylogeny reconstruction in its most simple form. It reduces the analysis to three terminal taxa (A, B and C), and seeks resolution amongst these in terms of relative relationships, e.g. the hypothesis \(h_1\): ‘A is more closely related to B than either is to C’, symbolized as \([A, B] \subset C\). Given that there are only three terminal taxa, there are only three possible logical permutations of relative relationships:

\[
\begin{align*}
h_1 &\equiv \{[A, B] C\} \\
h_2 &\equiv \{[A, C] B\} \\
h_3 &\equiv \{[B, C] A\}
\end{align*}
\]

The number of possible hypotheses of relationships increases dramatically with an increasing number of terminal taxa, but it will always be finite in number. The H-D argument in cladistics takes the finite universe of hypotheses of relative relationships that can be generated for any finite number of terminal taxa as the universe of all possible ‘predictions’, and considers the characters distributed over the terminal taxa as falsifiers of some (or all but one) of these hypotheses.

*Given \(N\) taxa, the number of possible relationships among these \(N\) taxa is given by the formula \((2^{N-3})!/(2^{N-3} \times (N-2)!\) (cf. Felsenstein 1978).

Kluge (2003, p. 236, n. 13) puts forth his falsificationist approach to the three-taxon statement as follows:

When the consequent of an implication is denied in deductive logic it is *modus tollens*, e.g. if \(h\) then not \(e\); therefore, \(h\) is false (Rieppel 2003). In the present case, if \(h_{[A, B]C}\) then not \(e_{1.0.1}\) or not \(e_{0.1.1}/e_{1.0.1}\) is false; therefore, \(h_{[A, B]C}\) is false; if \(h_{[A, C]B}\) then not \(e_{1.0.1}\) or not \(e_{0.1.1}/e_{1.0.1}\) is false; therefore, \(h_{[A, C]B}\) is false.

In his argument, Kluge (2003) employs a notational convention in which each element within the three element list subscribed to evidence statements is assigned to a different taxon depending on what hypotheses this evidence statement is consequent on. In the interest of simplicity, we will adopt a different notation in what follows. The matrix of informative character distributions invoked by Kluge (2003) is the following:

\[
\begin{array}{ccc}
A & 1 & 1 & 0 \\
B & 1 & 0 & 1 \\
C & 0 & 1 & 1 \\
\end{array}
\]

We will uniformly represent the first column as \(e_{1.0.0}\) the second as \(e_{1.0.1}\) and the third as \(e_{0.1.1}\). We then use the following abbreviations:

\[
\begin{align*}
e_{1.0.0} &= e_1 \text{ (first column)} \\
e_{1.0.1} &= e_2 \text{ (second column)} \\
e_{0.1.1} &= e_3 \text{ (third column)}
\end{align*}
\]

We will, again in the interest of simplicity, use the following abbreviations for hypotheses of descent:

\[
\begin{align*} 
   h_{[A, B]C} &= h_1 \\
h_{[A, C]B} &= h_2 \\
h_{[B, C]A} &= h_3 
\end{align*}
\]

Two caveats must be added at this point, however. First, *modus tollens* runs ‘if \(h\), then \(e\); not-\(e\); therefore not-\(h\)’ (\(h \rightarrow e, \sim e : \therefore \sim h\)) or, alternatively, ‘if \(h\), then not-\(e\); therefore not-\(h\)’ (\(h \rightarrow \sim e, e : \therefore \sim h\)). Kluge’s formulation ‘if \(h\) then not-\(e\); not-\(e\); therefore not-\(h\)’ (\(h \rightarrow \sim e, \sim e : \therefore \sim h\)) is not an accurate representation of *modus tollens*, and does not appear in that form in Rieppel (2003). This lapse does not, however, infect the rest of the argument. Secondly, to make the argument suggested by Kluge (2003) an acceptable application of *modus tollens*, Kluge’s natural language ‘or’ must be understood to mean logical conjunction (and) rather than disjunction (or). Given the notational convention suggested above, Kluge’s argument can now be rewritten as:

\[\text{Were we to take ‘or’ as meaning logical disjunction, the suggested argument would of course take a bad turn. If, e.g. } h_1 \rightarrow (\sim e_2 \lor \sim e_3), \sim e_1 \lor \sim h_1 \text{ had to be rendered valid, } \sim e_2 \lor \sim e_3 \text{ would have to be equivalent to } e_1 \text{, or, to put it otherwise, } \sim e_1 \text{ would have to be equivalent to } e_2 	ext{ and } e_3 \text{ [by an application of DeMorgan’s law to } \sim (\sim e_2 \lor \sim e_3)] \text{. This result would, of course, have to be rejected. Given that } e_1 \text{ does not obtain, we should not have license to infer that both } e_2 \text{ and } e_3 \text{ therefore do obtain, since we know that } e_2 \text{ and } e_3 \text{ can never both be the case concurrently. (We would like to thank previous reviewers of this paper for the helpful suggestion that Kluge’s natural language ‘or’ be understood to mean conjunction).} \]
When the consequent of an implication is denied in deductive logic it is *modus tollens*, e.g. $h \rightarrow \sim e$, $e$.

*modus tollens* (Rieppel 2003). In the present case $h_1 \rightarrow (\sim e_2$ and $\sim e_3)$, $e_1 \sim h_1$; $h_2 \rightarrow (\sim e_1$ and $\sim e_3)$, $e_2 \sim h_2$; $h_3 \rightarrow (\sim e_1$ and $\sim e_2)$, $e_3 \sim h_3$.

We have here an interesting suggestion as to how the use of *modus tollens*, and thereby a Popperian scheme of falsification, could proceed in systematic biology. Certain auxiliary assumptions must be invoked, however, to make Kluge’s (2003) argument a valid application of *modus tollens*. In order for $e_1$ to effect falsify $h_1$, we must assume that there exists some relationship between $\sim e_1$ and ($\sim e_2$ and $\sim e_3$) that allows *modus tollens* to gain purchase. Specifically, we must assume that $\sim e_1$ is in fact the negation of ($\sim e_2$ and $\sim e_3$), that, in other words, $e_1$ implies ($\sim e_2$ and $\sim e_3$) and conversely. This follows from the fact that the three evidence statements are mutually exclusive and exhaust the space of possible evidence. That is to say, given the mutually exclusive nature of the three evidence statements (i.e. that at most one is true at a time) and the further fact that exactly one of them can be assumed to be true, we can clearly infer that $e_1$ is true if, and only if, $e_2$ is not true and $e_3$ is not true ($e_1 \leftrightarrow (\sim e_2$ and $\sim e_3)$). This conclusion holds for the following reasons: (a) for a biconditional (an ‘if, and only if’—statement) to be true, both sides of the biconditional must have the same truth value, and (b) given that one (and only one) of $e_1$, $e_2$, $e_3$ is true at a time, the one which is true will appear either to the right or to the left of the biconditional; now (c) if it appears to the right, then the conjunction as a whole will be false, while the left side of the biconditional will be false as well, meaning the entire biconditional is true; if, on the other hand (d) the true proposition appears to the left, then both propositions to the right are false, so that the conjunction as a whole will be true, meaning the entire biconditional is again true. Similar arguments will of course also show (ii) and (iii) below to hold as well:

(i) $e_1 \leftrightarrow (\sim e_2$ and $\sim e_3)$;
(ii) $e_2 \leftrightarrow (\sim e_1$ and $\sim e_3)$;
(iii) $e_3 \leftrightarrow (\sim e_1$ and $\sim e_2)$.

Given these relationships between evidence statements, it is now easy to show how the falsification of any of the three hypotheses can proceed in the sense of Kluge (2003) given the H-D assumptions (iv) through (vi).

(iv) $h_1 \rightarrow (\sim e_2$ and $\sim e_3)$;
(v) $h_2 \rightarrow (\sim e_1$ and $\sim e_3)$;
(vi) $h_3 \rightarrow (\sim e_1$ and $\sim e_2)$.

It is now plain that $h_1 \rightarrow (\sim e_2$ and $\sim e_3)$, $\sim e_1 \sim h_1$ is a correct application of *modus tollens*. Knowing that ($\sim e_2$ and $\sim e_3$) implies $e_1$ (by (i) above), we can derive $\sim (\sim e_2$ and $\sim e_3$) from $\sim e_1$. Since $\sim (\sim e_2$ and $\sim e_3$) is the negation of ($\sim e_2$ and $\sim e_3$), we now get the desired $\sim h_1$ by *modus tollens* from $h_1 \rightarrow (\sim e_2$ and $\sim e_3)$. The same basic line of reasoning will also show the other two argument schemes to be valid: $h_2 \rightarrow (\sim e_1$ and $\sim e_3)$, $\sim e_2 \sim h_2$ is valid given (ii), and $h_3 \rightarrow (\sim e_1$ and $\sim e_2)$, $\sim e_3 \sim h_3$ is valid given (iii).

We have thus seen how the falsification of a hypothesis of descent can logically proceed via *modus tollens* given (i), (ii) and (iii). It should be noted, however, that each of the argument schemes proposed by Kluge (2003) uses $\sim e_1$, $\sim e_2$ or $\sim e_3$ as its starting point (thus, in e.g. $h_1 \rightarrow (\sim e_2$ and $\sim e_3)$, $\sim e_1 \sim h_1$, we take $\sim e_1$ as a starting point to get $\sim (\sim e_2$ and $\sim e_3)$, from which we then get $\sim h_1$.) It must certainly be agreed that it is equally possible to use a positive statement like $e_1$ rather than its negation ($\sim e_1$) as our point of departure. It may in fact even be more appropriate to do this given practical considerations. What can it really mean to make an observation of $\sim e_1$, $\sim e_2$ or $\sim e_3$? If we look at a certain column in a data matrix, what we see is not that some character distribution is not the case; rather, what we see is that a certain character distribution in fact is the case. If, in a three-taxon matrix, a column reads (110), what is known is not just that (110) or (011) is here not represented as being the case, but also, and more importantly, that that (110) is represented as being the case.

If we *do* begin with $e_1$ instead of its negation ($\sim e_1$), we can first off derive ($\sim e_2$ and $\sim e_3$) by *modus ponens* from (i). From the first conjunct, $\sim e_2$, we can, as already demonstrated above, prove $\sim h_2$, while from our second conjunct, $\sim e_3$, we can by the same line of reasoning, prove $\sim h_3$. Given $e_1$ we have, in other words, falsified both $h_2$ and $h_3$. But now that we know both $h_2$ and $h_3$ to be false given $e_1$, we should be in a position to conclude that $h_1$ therefore must be the case. Clearly, given that there is only one way in which history did in fact proceed, the three hypotheses of descent $h_1$, $h_2$ and $h_3$ are just as mutually exclusive as the three evidence statements we considered earlier. An argument analogous to the demonstration used to introduce (i), (ii) and (iii) above therefore allows us to introduce the following propositions as well:

(vii) $h_1 \leftrightarrow (\sim h_2$ and $\sim h_3)$;
(viii) $h_2 \leftrightarrow (\sim h_1$ and $\sim h_3)$;
(ix) $h_3 \leftrightarrow (\sim h_1$ and $\sim h_2)$.

Given (vii) above and the fact that $h_2$ and $h_3$ are both false given $e_1$, we are now in a position to conclude that $h_1$ must therefore be the *true hypothesis*. [Formally expressed, this means that $e_1 \rightarrow h_1$, i.e. that $e_1$ implies $h_1$ given (i) through (ix). We can follow a similar line of reasoning to prove that $h_2$ follows from $e_2$ and that $h_3$ follows from $e_3$ (i.e. that $e_2 \rightarrow h_2$ and that $e_3 \rightarrow h_3$)].

**Discussion**

In order to make Kluge’s (2003) argument a valid application of *modus tollens* given his H-D assumptions (iv)-(vi), we had to

It should be plain that (vii), (viii) and (ix) are uncontroversially true. We know that *some* phylogenetic process in fact took place, so that one of $h_1$, $h_2$, $h_3$ must be true, and since only *one* phylogenetic process took place, at most one of $h_1$, $h_2$, $h_3$ can be true.
introduce the propositions (i)–(iii) formalizing the mutually exclusive character of the three evidence statements. Using these propositions, we showed how it is possible using *modus tollens* to validly prove that the falsification of *h1* follows from \(\sim e1\), the falsification of *h2* from \(\sim e2\) and the falsification of *h3* from \(\sim e3\). We also saw, however, that using a positive evidence statement like *e1* as a starting point rather than its negation, \(\sim e1\), it follows that *h2* and *h3* are both false given *e1*. Given that hypotheses of descent are mutually exclusive along the lines formalized in (vii)–(ix), we then showed that *h1* must be the true hypothesis given *e1*, that *h2* must be true given *e2*, and that *h3* must be true given *e3* – that is, we showed that *e1* implies that *h1* is the case, *e2* implies that *h2* is the case and *e3* implies that *h3* is the case. This follows from the fact that *modus tollens* delivers necessary conditions as discussed in the introductory section.

The result that ‘if *e1*, then *h1*’, etc. (i.e. *e1* → *h1*, *e2* → *h2* and *e3* → *h3*), that is, the possibility of a positive proof of a hypothesis of descent given certain evidence, must be rejected by systematists for quite general reasons (e.g. the possibility of homoplasy). Or conversely, it is only if we could distinguish true synapomorphy from homoplasy that *modus tollens* could gain purchase in phylogenetic systematics (Ball 1982). Falsificationists will further notice that these results dissolve the very asymmetry between falsification and confirmation that originally inspired their position.

The results of this analysis do not rest on any peculiarity of the three-taxon statement under consideration. Given *n* taxa, the hypotheses *h1*, *h2*, ..., *hn* that can be generated remain finite in number. An observation of *en* will still falsify all hypotheses (however, many there may be) except the one that is assumed to imply *en*. The pattern of proof remains the same: if we grant that hypotheses of descent logically entail evidence statements, then (1) a positive evidence statement will, given the mutual exclusivity of evidence statements, always falsify every hypothesis except the one which implies it, and (2) the mutual exclusivity of hypotheses of descent will then allow for the conclusion that this remaining hypothesis must therefore be true. If we want to reject the possibility of positive proof, we clearly have to look to our assumptions for its root. Since propositions (i)–(iii) and (vii)–(ix) ought surely be granted given the mutually exclusive character of hypotheses of descent as well as evidence statements, the problem must lie in the H-D assumptions (iv)–(vi). It is, of course, exactly this thought – namely the thought that a specific hypothesis of descent somehow implies evidence statements – that makes possible the application of *modus tollens* in the first place, and consequently the resulting positive proof of a hypothesis of descent in the second place. Only by linking hypotheses and evidence through a material conditional (an ‘if – then’ statement that expresses material implication as discussed in the introductory section) does *modus tollens* gain any purchase – and only thereby do we reach the final conclusion that a certain hypothesis must in fact be true. Our argument must therefore be seen as a *reductio ad absurdum* of this sort of hypothetico-deductive argument in systematics (see Sober 1988, for a similar conclusion). It shows that if hypothetico-deduction is introduced into systematics in this way, it inevitably leads to untenable results, and that therefore, since Popperian falsification cannot be had outside a H-D framework, Popperian falsification cannot be adopted in systematics in this manner either.

**Conclusions**

That hypothetico-deductionism and the attendant possibility of falsificationism are inapplicable in systematics should come as no surprise (Sober 1988). To say that a hypothesis implies a certain character distribution is to say that if this hypothesis is true, no other character distribution could possibly occur. While a given phylogeny may certainly suggest a greater likelihood of one character distribution over another, to ask for more is simply to ask too much. It is at this juncture that Kluge’s (2003) argument points in another interesting direction, even if away from *modus tollens*, material implication, and Popperian falsificationism. An argument scheme closely similar to Kluge’s (2003) was in fact championed for phylogenetic systematics by Zimmermann (1943).

Given his empirio-critical background, Zimmermann’s (1943, p. 32) approach was again governed by logic, yet not by material implication as underlies *modus ponens* and *modus tollens*, but by a weak rendition of the law of bivalence. Zimmermann (1943) recognized that phylogenetic systematics is to some degree statistical in nature, its results therefore to some degree probabilistic. The starting point for Zimmermann’s (1943) argument is again the assumption of a single unique historical process of phylogeny, which results in the mutual exclusion of alternative hypotheses of relationships. Going from there, Zimmermann (1943, p. 32) requested that alternative hypotheses of relationships be formulated in terms of pairs of contradictory opposites, where only one hypothesis could possibly be true for each pair. On that basis, Zimmermann (1943, p. 32; emphasis added) argued that if one of the alternatives is extremely unlikely, then the other alternative is – according to his own logic – ‘necessarily proven’, even if empirically probable only to a certain degree. This argument may appear incoherent, and requires for its appreciation a proper understanding of the empirio-critical notion of ‘statistical law’ employed by Zimmermann (1943). If of two logically contradictory propositions (e.g. *P* and \(\sim P\)) one and only one can and must be true, and if one of the propositions is rejected, then the other must be necessarily the true one. This is mandated by logic. For Zimmermann (1943), this does not mean that we have practical means that would allow us to know, with absolute certainty, that we have correctly identified the true proposition. In the empirical world, we may approach the true hypothesis with a certain degree of probability only. That our empirical knowledge of the world must remain probabilistic for practical reasons does not mean, however, that propositions of phylogenetic relationships do not have absolute truth-values relative to the real world.

Using again a three-taxon statement, Zimmermann (1943, p. 40) obtained: \([A \land B, C]\) is the case, \([A \land C, B]\) and \([A \land B, C]\) are not the case; or: \([A \land C, B]\) is the case, \([A \land B, C]\) and \([A \land B, C]\) are not the case; or: \([A \land B, C]\) is the case, \([A \land B, C]\) and \([B \land A, C]\) are not the case. On the basis of the uniqueness of the phylogenetic process, the principle of bivalence admits only one alternative of each of these three pairs of contradictory hypotheses of relationships to be true. The problem remains how to select the one hypothesis that is most probably the true one. Given the acknowledged possibility of convergence, the probability to hit on the correct hypothesis of relationships increases with the number of characters considered: ‘If we try to reconstruct the relationships of organisms using singular characters (artificial systematics), the danger of error due to convergence is always big. As soon as we take into account the
Zusammenfassung

Logik in der Systematik

Aufgabe der Biosystematik ist die Rekonstruktion taxonomischer Beziehungen zwischen Arten, die deren evolutiver Geschichte entsprechen (Taxa sollen monophyletisch sein). Wie phylogenetische Verwandschaftsbeziehungen rekonstruiert werden sollen, bleibt nach wie vor umstritten. Es gibt unterschiedliche Formen empirischer Evidenz, und unterschiedliche Methoden phylogenetischer Rekonstruktion. Dieser Beitrag soll zeigen, dass eine dieser Methoden, nämlich die hypothetisch-deduktive, für die phylogenetische Systematik problematisch ist. Der Grund hierfür liegt darin, dass die hypothetisch-deduktive Methode eine zu starke (d.h., nomologische) Beziehung zwischen Theorie (Hypothesen phylogenetischer Verwandtschaft) und Evidenz verlangt.

References


© 2006 The Authors JZS 44(3), 186–192
Journal compilation © 2006 Blackwell Verlag, Berlin