

Distinguishing “or” from “and” and the case for historical identification

Arnold G. Kluge*

Museum of Zoology, University of Michigan, Ann Arbor, MI 48109, USA

Accepted 23 August 2002

Abstract

The *adequacy* of a probabilistic interpretation must be judged according to the nature of the event, or thing, being inferred. For example, conditional (frequency) probability is not *admissible* in the inference of phylogeny, because basic statements of common ancestry do not fulfill the requirements of the relations specified by the probability calculus. The probabilities of the situation peculiar to the time and place of origin of species are *unique*. Moreover, according to evolutionary theory, an event of species diversification is *necessarily* unique, because species are parts of a replicator continuum—species arise from ancestral species. Also, these probabilities cannot be *ascertained*, because the relevant situation cannot be repeated—it is unique. Finally, the *applicability* of conditional (frequency) probability is denied, because events of common ancestry have already occurred—there is nothing to predict. However, hypotheses of species relationships can be *identified objectively* according to the degree to which they have survived simultaneous testing with critical evidence, not with generally confirming evidence. The most parsimonious hypothesis of species relationships represents the least disconfirmed, best *supported*, proposition among the alternatives being compared. That hypothesis does not, however, deserve any special epistemological status beyond serving as the focus of the next round of testing.

© 2002 Elsevier Science (USA). All rights reserved.

Uncertainty and probability

Uncertainty is considered one of the important issues in phylogenetic inference, at least by those who induce hypotheses of species relationships probabilistically.¹ Any discussion of uncertainty and probability is bound to be complicated, because two fundamentally different philosophies are available for determining degree of validity, the *subjective* and the *objective*. Another term for the former is “belief” probability. Bayesian (posterior) probability is a personalistic kind of subjective probability, which is widely criticized because of its subjectivity (e.g., Salmon, 1966). The Monty Hall problem provides a clear example of the significance of prior knowledge in Bayesian inference (Howson

and Urbach, 1993). The virtues of recently published phylogenetic studies claiming to be Bayesian (e.g., Huelsenbeck et al., 2001a; Murphy et al., 2001) are misplaced, because the methods employed do not include a truly subjective prior probability of the hypothesis, $p(\mathbf{h})$, as determined from experience. The aforementioned studies are at best pseudo-Bayesian, which leaves their epistemology to be argued elsewhere.

Making matters even more complicated, there are two theories of objective probability to choose from, that which is concerned with a *sequence* of instances or with a *single* instance.² The former is exemplified by the familiar *conditional* probability of statistics, and by inductive inference more generally. Conditional probability can be defined formally as $p(\mathbf{a}, \mathbf{b}) = \mathbf{r}$, the (relative) probability of \mathbf{a} given (under condition) \mathbf{b} ,

* Fax: 734-763-4080.

E-mail address: akluge@umich.edu (A.G. Kluge).

¹ Both the objective and the subjective (epistemic) are said to contribute to the uncertainty of a hypothesis, the determinism/indeterminism of the event, or thing, and the sampling/systematic error of the investigation, respectively.

² The “unique case rule” discussed by Hacking (1965) as the basis for the “law of likelihood” has nothing to do with the probabilification of the kind of unique instance (event) being considered here, because that rule pertains to joint propositions that are consistent with statistical (frequency) data.

where r is some fraction between 0 and 1. Given a frequency interpretation, $p(\mathbf{a}, \mathbf{b}) = r$ becomes the relative frequency of \mathbf{a} within the reference class \mathbf{b} being equal to r , where the condition \mathbf{b} is a random population sample. In the classical interpretation of probabilities, objective uncertainty is attributed to a property *inherent* in the event or thing (see below).

Less familiar is the single-instance type of objective probability, which is usually referred to as *absolute* probability. When absolute probability is interpreted strictly logically it is labeled “absolute logical probability.” In whatever way it is interpreted, this kind of objective probability can be defined formally as $p(\mathbf{a}) = r$, the probability of \mathbf{a} , where r is the absolute value, true or false, as in the names of statements, \mathbf{a} , \mathbf{b} , \mathbf{c} , ... Consider, “its value r is the greater the less statement \mathbf{a} says. Or in other words, the greater the content of \mathbf{a} , the smaller is the value of its absolute logical probability” (Popper, 1983, p. 284). Thus, for nontrivial statements, \mathbf{a} is more probable than \mathbf{ab} , provided \mathbf{b} does not follow tautologically from \mathbf{a} . Objective uncertainty is, in this case, attributed to the situation of which the event, or thing, is a part. Popper (1990) refers to this as the probability of propensities, and he argues that the propensity interpretation is closely related to the classical theory of probability, which simply takes probability as a measure of possibilities (see below). Railton’s (1978) and Popper’s (1959, 1983)³ reviews of the objective theories of probability, and the roles testability and deduction play in the inference of particular instances, should be consulted for details (see also below and Kluge, 1997a,b, 1998, 1999, 2001).

I refer to probabilities as either subjective or objective in the text to follow, with the latter being either conditional or absolute. The term “frequency” probability has also been used in place of conditional probability, a relationship that I acknowledge throughout this paper with the label “conditional (frequency) probability.” All debates concerning subjective and objective probabilities are clearly defined, in just those terms, and it is on this basis that I reject subjective probability, as do many others, because matters of repeatability are not considered (Salmon, 1966). The distinction between conditional (frequency) and absolute probabilities has received less discussion and little refinement (e.g., see Hacking, 2001). The arguments I begin with are intended to underscore the idea that a conditional (frequency) probability is an *inadequate* probabilistic interpretation of a singular historical event. I conclude by pointing out that the degree of corroboration (*sensu* Popper, 1959, 1983) provides an objective probabilistic

interpretation of species relationships, as it can for any basic statement. The burden of objective uncertainty necessarily assumed by conditional (frequency) probability is not an issue when it comes to degree of corroboration.

Basic arguments: examples from human history

The nature of the event, or thing, determines which of the objective probabilities is a satisfactory interpretation. For example, the hypothesized event, or thing, must be well isolated, stationary, and recurrent in the application of conditional (frequency) probability (Popper, 1990). Since many phylogeneticists are conditional (frequency) dogmatists (e.g., Felsenstein, 2001),⁴ or seem to be unfamiliar with the nature of historical events (e.g., de Queiroz and Poe, 2001; Huelsenbeck et al., 2000),⁵ I begin my discussion with well-documented examples from human history (see also Carpenter, 1992).

Consider the following:

Statement: *William defeated Harold at Hastings in 1066.*

Question: *What is the mean?*

Wenzel and Carpenter (1994) combined this statement and question in order to draw attention to a contradiction; the statement concerns a single event, the Battle at Hastings (B_H), while the question pertains to a sequence of events ($E_1, E_2, E_3, \dots, E_n$). A mean is an abstraction; it does not pertain to any thing, like the stated historical event, B_H . Further, a conditional (frequency) probability of a historical event, like $p(E) = 1/2$, is not a *testable* statement, because (1) the limit of the relative frequency for a series of independent random observations is realized as the event approaches infinity, but from which it follows that (2) the statement $p(E) = 1/2$ is empirically meaningless, because a hypothesis cannot be proven false by something proven to be true. This simple exercise, as provided by Wenzel and Carpenter’s basic statement and question, suggests that there is much to consider when ascertaining the probability of a historical proposition.

The situation relevant to a historical event clearly exposes the inadequacy of a conditional (frequency) probability interpretation of these kinds of basic statements. In the case of the Battle at Hastings, as stated

³ For example, Popper’s (1959) Appendix *IX, Corroboration, the Weight of Evidence, and Statistical Tests, provides several formal proofs of the distinction between absolute and conditional (frequency) probability (*contra* de Queiroz and Poe, 2001; see Kluge, 2001).

⁴ Both dogmatic and pejorative, as the following quote from Felsenstein reveals: “After coping with taxonomists, who tend to dismiss statistical inference and adopt arbitrary and bizarre ‘hypothetico-deductive’ philosophical frameworks, it is refreshing to deal with statisticians, who are not tempted to replace the hard work of inference by philosophical quotation-mongering” (Felsenstein, 1987, p. 208).

⁵ Siddall (2001) emphasized the subjectivist error of nonindependence ignored by de Queiroz and Poe, not their having overlooked the distinction between conditional (frequency) and absolute probability, which is one of the issues focused on in the present paper.

above, we may conclude that the outcome of the event happened or it did not; either William defeated Harold or he did not. That objective uncertainty of that battle, like any battle, is determined by the *situation* pertaining to the event. While the written record indicates only the coarsest of these details, they are sufficient for the point I wish to make. To dismiss such records, for whatever purpose, on the grounds that they might have been fabricated or are incomplete, is ad hoc. As for the facts, according to the record, the Crown of England was disputed upon the death of King Edward the Confessor, January 5, 1066. The record further indicates that the claimants were the Saxon earl, Harold Godwine, and William the Bastard, the illegitimate son of Robert I, Duke of Normandy. Further still, the issue is recorded as having been resolved in a daylong battle, October 14, 1066, on a ridge 10 miles northwest of Hastings, Sussex County, England. Last, the Battle of Hastings is said to have ended when the mailed Norman cavalry killed many of the light-armored Saxon axe-men, including Harold, who died with an arrow through his eye. As characterized by just these few parameters pertaining to the situation, without speculating on unrecorded physicalities, like the improbabilities of the weather and other local circumstances, it is fair to say that there is no population of that *particular* battle, (B_H); there is no sequence from which repeated samples might be drawn, in order to evaluate its objective uncertainty, say, according to mean, variance, and skew. Likewise, there is nothing to which a Monte Carlo routine can be applied, that which would provide a meaningful set of pseudoreplicates and from which departure from random might be estimated.

Other battles are not repetitions of the one in question. For example, it is recorded that the Battle of Waterloo (B_W) took place on June 18, 1815, 3 miles south of Waterloo, near Brussels, Belgium. Thus, the battles of Hastings and Waterloo involve different situations—the times and places were not the same. The causes (reasons) for those two battles and the consequences resulting therefrom are also different. In the case of the Battle of Hastings, Edward the Confessor promised the Crown of England to both Harold and William.⁶ Harold assumed the Crown upon Edward's death, and ruled for 9 months, until William, then the Duke of Normandy, “took” it from him on the field of battle, at Hastings. William became King William I the Conqueror, which ended the Anglo-Saxon phase of English history, and his Norman policies dramatically transformed the secular and ecclesiastical life of England. In the case of Waterloo, the expansionist policy of France, under Napoleon Bonaparte, had become intolerable to the rest of Western Europe, including

England, which was represented at this battle by the army of the Duke of Wellington. Napoleon's “Waterloo” was the rest of Western Europe's economical and political gain.

Even though each of these two battles is objectively unique, and therefore $B_H \neq B_W$, some conditional (frequency) probabilists still might argue that they are instances of some homogeneous construct, e.g., a class concept of battle (“B”), for which a conditional (frequency) probability or likelihood might then be calculated. For example, the average relative number of opposing troops involved in Western European battles before the 20th century (“Bx”), or the probability of defeat given the relative sizes of the opposing forces throughout recorded history (“By”), might be estimated statistically. Thus, at least for those committed to conditional (frequency) probability, B_H and B_W might serve as instances of both “Bx” and “By.” However, the objective uncertainties of B_H and B_W are different, and it is clear that neither battle is a replicate of the other event. These battles cannot inform on each other.

Insights provided by the theoretical requirements of probabilities

Further insight into the kind of objective probability that is consistent with the inference of unique historical events can be obtained from the theoretical requirements of two kinds of conditional (frequency) probabilities, what some call *classical* (or theoretical) and *historical* probability. Requirements of the classical model include: (1) all events E (outcomes) can be enumerated (e.g., all of the values a measurement or a count can possibly take, as in the case of dice), (2) the probability of each event is known, and (3) a corresponding formula for each event may be written as

$$p(E) = \frac{\text{number of times } E \text{ will occur}}{\text{total number of } E\text{'s that occur.}}$$

The conditional (frequency) kind of historical probability differs from those requirements in important, if subtle, ways: (1) only events that have occurred can be enumerated, which means that new information on events can change the probability; (2) the probability of each event is based on how often it occurred in the past; and (3) there is no formula (in the technical sense), but the data can be compared nonetheless, such as

$$p(E) = \frac{\text{number of times } E \text{ did occur}}{\text{total number of } E\text{'s that occurred.}}$$

Essentially, the $p(E)$ of this kind of historical probability refers to the instances of a class that is a subset of E , whereas the $p(E)$ of classical probability refers to instances of the class of all E . It is clear that there is a frequency kind of probability calculus at work in both

⁶ That either, or both, promises may have been “vague” has no effect on the dispute itself.

classical and historical interpretations, so that: (1) every probability is a single-valued real number, $0 \leq p(E) \leq 1$, and (2) the probabilities of all possible events, $E_1, E_2, E_3, \dots, E_n$, in a set add up to 1, $\sum p(E_i) = 1$.

The fact that both of these kinds of probabilities are based on a conditional (frequency) interpretation of objective uncertainty means that they are inadmissible to the interpretation of a single historical event, not withstanding the “historical” probability label given to one of these probabilities. As argued above, the Battle of Hastings is unique, and the objective uncertainty of the outcome of the event (William would, or he would not, defeat Harold) was to be determined by the properties peculiar to the situation, of which the battle was a part. Reiterating, the basic statement to which the Battle of Hastings pertains concerns a particular event (B_H), not a nontrivial set of two or more events instantiating some class concept (“B”), and the uniqueness of that event is a function of the properties peculiar to the situation of B_H . Thus, the kind of probability that may be applied to a basic statement of uniqueness, like that pertaining to the Battle of Hastings, or to some event in species history, must lie elsewhere in objective probability theory.

The origin of species: relevant contingent facts

According to the contingent facts of Darwinian theory (Darwin, 1859), the origin of species provides an especially strong argument for the inadmissibility of conditional (frequency) probability in the inference of phylogenetic relationships. Like the aforementioned battles, each species (S_i) is the outcome of a unique event—there is a singular point of common ancestry for any pair of species.⁷ That being the case, we still might ask why cannot there be a concept of sister species relationships (“S”), like the alternative common ancestral events that constitute the *logically* closed hypothesis set of species relationships, the three-taxon case being (A,B)C, (A,C)B, (B,C)A, for which conditional (frequency) probability has been used to predict which of those alternatives is more likely to be true? Effectively, we are asking why there cannot be instances S_i of “S”. There can be no such relationship, because each point of common ancestry is not just a unique event, but one that is *necessarily* unique, and that being the case there cannot be, as I have argued elsewhere (Kluge, 1990), a reference class (“S”) from which to sample. Species are *unique*, because of the probabilities of propensity, each certainly having its own time and place of origin, and extent (like battles), and they are *necessarily* so (unlike

battles), because species are, according to Darwinian theory, parts of a *replicator continuum*, in which unique species originate from other unique species (Lidén, 1990).⁸ A battle does not necessarily replicate another set of battles, which in turn replicate more battles, and so on. While the inductively committed might argue that battles may be enumerated as members of some kind of reference class, logically there can be no such reference class when it comes to the history of species, beyond the trivial case of a single instance. There is but one, and only one, history of any S_i , or of all S_i , $p = 1/1$ in either case. To emphasize this important distinction, the terms *historical individual* and *individual* are used for these two kinds of historical things, those that are “necessarily unique” (e.g., S) and those that are just “unique” (e.g., B), respectively.

The calculus of conditional (frequency) probability is inconsistent with the nature of a species origin, and so conditional (frequency) probability is inadmissible in the probabilification of the objective uncertainty of historical individuals. As for the absolute probability of any statement of sister species relationships, the diversification can be interpreted as having happened or not. As for the probability of life itself, creationists use a conditional (frequency) interpretation to claim that it is so improbable as to be impossible, while the absolute probabilist only observes that it happened!

More from the probability of propensities

As noted in the introduction, objective uncertainty has been attributed to a property inherent to the event, or thing, or to the situation of which the event, or thing, is a part. The latter interpretation, Popper (1990) emphasizes, is important, because the probability of those propensities generalizes and extends the idea of forces throughout the universe. It is important as well, as I argued above, because it extends to the necessarily unique singular case of common ancestry—the ancestral, historically contingent, situation.

Propensity probability may also generalize to evolutionary theory. For example, Stamos (2001) makes the case for the quantum-indeterminism of point mutations. That is, the “nucleotide-level of molecular evolution is in all likelihood indeterministic (obeying the ‘laws’ of quantum physics), and the necessary uniqueness of the relevant aspects of the situation means that *every* evolutionary event has its own propensity or objective probability” (T. Grant, personal communication; my italics). Coupling this explanation of the micro-evolutionary process with the historically contingent/

⁷ The ancestor, or common ancestral species, of any monophyletic group, such as (A,B), “is *identical* with *all* of the species that have arisen from it” (Hennig, 1966, p. 71; my italics); it is not one of the parts of the more inclusive taxon being defined ostensibly.

⁸ “A unique event is one that happens to be one of a kind. A necessarily unique event is defined in terms that preclude any other instance of that event” (Hull, 1974, pp. 47, 97, 98, 1980).

necessarily unique condition of species makes a strong case for the indeterminism of evolutionary theory (see also Lewontin, 1966; Sarkar, 1998).

This view of evolution calls into question a great deal of published research involving conditional (frequency) probability. It also argues against the lawfulness so often assumed in particular fields, like comparative biology. For example, Losos et al. (1998) claim that the diversification of *Anolis* lizards in the Greater Antilles constitutes a set of adaptive “replications,” which assumes evolutionary theory is fundamentally deterministic and statistical in character. However, those lizard case histories are better interpreted as evidence of indeterminism, because the cladistic hypotheses Losos et al. provide as independent evidence of their thesis represent substantially different ecomorph histories, i.e., there is no intensionally defined set instantiated by repeated, independent observations. Moreover, it can be argued that Losos et al. did not, in the first place, provide a necessary and sufficient (intensional) definition of the concept of adaptive radiation.

To justify the adequacy of a conditional (frequency) probability calculus when interpreting the events of nucleotide evolution requires, as noted above, that they be well isolated, stationary, and recurrent, requirements denied in the system of nucleotide evolution (Elsasser, 1966), just as in the case of the species lineage system. Attributing law-like generalities, a deterministic kind of explanation, to species history and character evolution, like adaptive radiation and homoplasy being evidence of adaptation, can be nothing more than exercises in *historicism* (Popper, 1957).

Despite the fact that historical contingency limits evolutionary options, one has to admit that the number and variety of those possibilities that have been realized are truly staggering. Furthermore, there is absolutely no reason to believe that the possibility of further extraordinary diversification of unique lineage systems is closed as the result of phylogenetic constraint (Popper, 1990).

Without having to face the burden of probabilifying objective uncertainty

As for making inferences of unique historical events, Popper pointed out that “propensities cannot be measured because the relevant situation changes and cannot be repeated. This would hold, for example, for the different propensities of some of our evolutionary predecessors to give rise to chimpanzees and to ourselves. Propensities of this kind are, of course, not measurable, since the situation cannot be repeated. It is unique. Nevertheless, there is nothing to prevent us from supposing that such propensities exist, and from estimating them *speculatively*” (Popper, 1990, p. 17; my italics).

Still, as Hull emphasized, the uniqueness of past events does not mean evidence no longer exists. Indeed, all that is required in the inference of common ancestry is that these kinds of “natural phenomena contain traces of their past and thus can serve as records” (Hull, 1974, pp. 95, 96).

And how to proceed objectively, given such “traces,” as evidence, is straightforward according to Watkins (1984, p. xi): “to submit our knowledge of the external world to an ordeal by skepticism and then, with the help of the little that survives, to explain how scientific rationality is still possible.”

Thus, phylogenetic inference is still possible, where sister group relationships are severely tested for their ability to explain synapomorphies as homologues, as evidence of common ancestry (see below; Farris, 1983; Kluge, 1997a, 1999). In this scientific paradigm, there is no reason to assume the burden of having to speculate on the objective uncertainty of past events, as demanded of conditional (frequency) probability.

That an informed kind of *speculation* is even possible is debatable (Farris, 1983). For example, if phylogeny is modeled with gene sequence data, as required of maximum likelihood and Bayesian methods, then science is impossible, because the probability of a nucleotide event having occurred cannot be known, given that knowledge of the relevant history has to be presupposed (Siddall and Kluge, 1997). To “reciprocally illuminate” the evidence from the phylogenetic hypothesis(es) that the evidence is used to test, as required by most gene sequence alignment protocols, and as employed in the estimation of nucleotide substitution parameters of the models, is viciously circular (*contra* Mindell, 1991), because no independent test is involved when verifying one hypothesis with another (Hull, 1967; Walton, 1985). To assume law-like generalities, like a common mechanism, in order to eliminate the groping, only engenders a lack of realism, because model assumptions are counterfactual conditionals (Grant, 2002).

Clearly, there can be no advantage to Bayesian inference, in which each clade is supposed to be definable in terms of a cumulative probability, because there is no frequency class on which to base that kind of estimation—there is no basis for using a conditional (frequency) probability to characterize the objective uncertainty of any clade. Moreover, it is meaningless in Bayesian inference to provide a *subjective estimate* of the prior probability of competing hypotheses, $p(\mathbf{h})$. Since each lineage (i.e., outcome of cladogenesis) is a historical individual, as exemplified by the three-taxon case, it follows that if $p_{(A,B)C} = 1$, then $p_{(A,C)B} = 0$ and $p_{(B,C)A} = 0$; if $p_{(A,C)B} = 1$, then $p_{(A,B)C} = 0$ and $p_{(B,C)A} = 0$; or if $p_{(B,C)A} = 1$, then $p_{(A,B)C} = 0$ and $p_{(A,C)B} = 0$. In other words, as argued above, a historical hypothesis is inferred to be true, $p(\mathbf{h}) = 1$, or is inferred to be false, $p(\mathbf{h}) = 0$; it cannot be determined to be some real

number in between, $0 < p(\mathbf{h}) < 1$. Likewise, it is irrational to assume that the prior probabilities of the hypotheses are flat (uninformative), as proposed by Huelsenbeck et al. (2001a,b), in which in the case of three taxa $p(\mathbf{h}_1) = 0.333_{(A,B)C}$, $p(\mathbf{h}_2) = 0.333_{(A,C)B}$, and $p(\mathbf{h}_3) = 0.333_{(B,C)A}$, because a real number probability cannot be an explanation of a necessarily unique event.

The logic of historical identification, and distinguishing “or” from “and”

The testability of phylogenetic systematics stems from the logic that a statement of historical individuality can be replaced by a conjunction of basic statements—each historical thing constituting a *finite*, spatiotemporally restricted, group of individuals (e.g., Tetrapoda consists of Lissamphibia and Reptilia; each of these in turn consists of . . .). In this sense, a historical individual is a *numerical universal* (Stamos, 1996). *Strictly universal* statements are different from numerical universal statements in that they cannot be replaced by a conjunction of singular statements, because universals are spatiotemporally unrestricted.

Strictly universal statements are falsifiable and are said to constitute laws of nature, because they make statements about natural necessity. While statements of historical individuality are, by comparison, only descriptive of nature, they too are falsifiable, because they are prohibitive. Consider that the basic statement of historical individuality $(A,B)C^9$ prohibits the individual C from the group (A,B) and likewise B from (A,C) and A from (B,C). All members of a set of exhaustively enumerated basic statements for n terminal taxa have equal empirical content (equal improbability), because each statement of that set prohibits (logically) all competing statements.

The relationship between universal statements and their falsifiers can be argued in purely logical terms (e.g., Popper, 1959). Empirically, however, a falsifier of a universal statement does not prove the statement false (e.g., Popper, 1983) because either the universal statement or the falsifier can be erroneous. Thus, statements of historical individuality are empirically falsifiable with synapomorphies, i.e., propositions of shared-derived character states, in which a synapomorphy congruent with a basic statement disconfirms the alternatives (Wiley, 1975). For example, a synapomorphy diagnostic of the group (A,B) counts against (A,C) and (B,C), one of (A,C) counts against (A,B) and (B,C), and one of (B,C) counts against (A,B) and (A,C). Unlike strictly universal statements, numerical universal statements are evaluated in relative terms—by weight of evidence—the

least disconfirmed of the competing basic statements having the highest degree of corroboration (*sensu* Kluge, 1997a; Popper, 1959, 1983; *contra* Rieppel, 2002a).

According to Popper (1959, p. 145), employing the “principle of parsimony . . . restrains us from indulgence in ad hoc hypotheses and auxiliary hypotheses . . . and the degree of falsifiability of a system [is] thus protected [from going] to zero.”

And as Farris (1983, p. 9) pointed out, the “reason for this preference [in phylogenetic systematics] is that each falsifier of any accepted genealogy imposes a requirement for an ad hoc hypothesis [of homoplasy] to dispose of the falsifier.”

In minimizing *ad hoc* hypotheses of homoplasy, a synapomorphy congruent with a basic phylogenetic statement is thus hypothesized to be homologous—shared derived character states are explained in terms of inheritance from a common ancestral condition (Farris, 1983; Hennig, 1966; *contra* Rieppel, 2002b)—no probability is necessary to connect the present to the past (*contra* Sober, 1988). Further, the cause and effect in this deductive scheme of inference is not viciously circular, as some may think (for the nature of circular reasoning see Hull, 1967). Consider that all homologues are synapomorphies, but not all synapomorphies are homologous. Thus, some shared derived character states must be erroneous (Farris, 1983; Kluge, 1999). That propositions of homoplasy can be further critically examined outside the context of a phylogenetic hypothesis and character congruence is an example of reciprocal illumination (Farris, 1983; Hennig, 1966; Kluge, 1997b).

According to this logic of phylogenetic systematics, the existential statements of historical individuality are empirical, not metaphysical. Thus, I believe the paradigm of phylogenetic systematics provides a standard by which to judge the growth of knowledge in other comparative biological sciences. Indeed, I predict that much research that is ecological and evolutionary will be found to be unscientific (Stamos, 1996, p. 185).

In thinking of a phylogenetic hypothesis probabilistically, the objective uncertainty of a single, necessarily unique, event is readily distinguished from a sequence of events of one kind—“or” from “and”— $pS_{(A,B)C}$ or $pS_{(A,C)B}$ or $pS_{(B,C)A} = 1$ from $pS_{(A,B)C}$ and $pS_{(A,C)B}$ and $pS_{(B,C)A} = 1$. Given the notation $p(E) =$ “probability of a historical event,” the rules of the probability calculus sufficient to define these alternative phylogenetic formalisms are $p(E) = 1$, i.e., E must have occurred, or $p(E) = 0$, i.e., it did not, and if and only if E_1 and E_2 and $E_3 \dots$ and E_n are all of the possible n instances of a class that is a subset of E , then $\sum_i p(E_i) = 1$.

The former rule covers the objective uncertainties of inferring a necessarily unique event, like the relative recency of common ancestry being (A,B)C or (A,C)B or (B,C)A, whereas the latter rule must be judged

⁹ That is to say, A and B share a more recent common ancestor than either does with C.

inadmissible in the inference of historical individuals, because it is contradicted by historicity, i.e., history does not repeat itself (Popper, 1957).

That not all systematists may be persuaded by the logic outlined above, and by the distinctions between “or” and “and,” is evident from the title of Felsenstein’s (2002) recent book, *Inferring Phylogenies*. To be sure, there are competing phylogenetic hypotheses (the “or” of historical identification); however, there is only one phylogeny to be inferred, and the conditional (frequency) dogmatists’ approach to probabilifying phylogenies is rendered inadequate. In other words, “inferring phylogeny” deductively is logically consistent with the nature of historical things like species, whereas “inferring phylogenies” statistically is logically inconsistent.

In order to improve our communication of ideas, it is important that all forms of psychologism be eliminated from scientific discussion, like when common ancestry is defined in terms of the conditional (frequency) probability of nucleotide *similarity* (Ghiselin, 1966). I believe the inference of historical individuals is better understood as one of “identification,”¹⁰ or “retrodiction” (objective postdiction), not “intensional definition” or “prediction”—it is not one of objective “estimation.” Historical identification is consistent with what is being inferred from a system that already exists, that of phylogeny, whereas estimation is inapplicable, because it involves predicting future events. The inference of historical individuals may even be understood as one of “discovery,” as in exploring knowledge of some thing, but not in claiming to *know* some thing with certainty (e.g., the intensionally defined taxa, class concepts, of pattern cladism; see Platnick, 1982).

Other consequences: some approaches to phylogenetic inference are guilty

Only certain methods and data are consistent with the identification of relative recency of species common ancestry, i.e., phylogeny. All skeptical treatments of hypotheses involve a concern for evidence, **e**, in relation to hypothesis, **h**, in light of background knowledge, **b**, as summarized by the numerator of Popper’s (1959) degree of corroboration, $C(\mathbf{h}, \mathbf{e}, \mathbf{b}) = p(\mathbf{e}, \mathbf{hb}) - p(\mathbf{e})$ (Kluge, 1997a). A hypothesis of species common ancestry, like $\mathbf{h}_{(A,B)C}$ or $\mathbf{h}_{(A,C)B}$ or $\mathbf{h}_{(B,C)A}$, is falsified objectively (see above), assuming “descent, with modification,” as

background knowledge, according to the degree to which it has survived tests with critical evidence (*contra* Patterson, 1978; see Kluge, 2001). The degree to which corroboration relativizes falsifiability to hypothesis identification is clearly expressed in one of its most familiar denominators, $p(\mathbf{e}, \mathbf{hb}) - p(\mathbf{eh}, \mathbf{b}) + p(\mathbf{e}, \mathbf{b})$. Consider, “if **e** supports **h** (given the background knowledge **b**) then $C(\mathbf{h}, \mathbf{e}, \mathbf{b})$ is positive. If **e** undermines **h** (so that non-**e** supports **h**) then $C(\mathbf{h}, \mathbf{e}, \mathbf{b})$ is negative. If **e** does neither, so that it is *independent* of **h** in the presence of **b**, then $C(\mathbf{h}, \mathbf{e}, \mathbf{b})$ equals zero” (Popper, 1983, p. 241). Thus it is that the least falsified, most parsimonious, hypothesis of species relationships is considered the better supported, e.g., $C(\mathbf{h}_{(A,B)C}, \mathbf{e}, \mathbf{b}) > C(\mathbf{h}_{(A,C)B}, \mathbf{e}, \mathbf{b}) > C(\mathbf{h}_{(B,C)A}, \mathbf{e}, \mathbf{b})$. Such hypotheses, however, deserve no special status beyond serving as “*objects for further tests*” (Popper, 1979, p. 13, italics in the original). Arguably, all possible competing hypotheses can have some objective support.

All that claims to be part of an inference scheme must be judged against the ontological status of what is being inferred (Grant, 2002). Thus it is that all inductive approaches to phylogenetic inference, like maximum likelihood and Bayesian probability, may be judged with suspicion. General classes of methods, like taxonomic congruence and a posteriori character weighting, suffer the same consequence (Kluge, 1997b). Even some of the most commonly used methods in phylogenetic inference are inadequate. While the bootstrap and the jackknife are touted as measuring hypothesis support, they assume a conditional (frequency) interpretation of probability, where no frequency exists, or they do not index the criticalness of the tests made of those kinds of hypotheses.

I conclude that any approach claiming to *estimate* a necessarily unique historical event from a sample of those instances must be charged as guilty of an illogical thesis. As Salmon (1966, p. 132), one of the leading proponents of frequency probability, readily admits at the end of his book on the foundations of that kind of scientific inference, the guilty verdict may apply more generally, because a “well-established method for ascertaining fundamental probabilities” and a “justification of induction” are still wanting. The phylogeneticist who claims either a subjective or an objective probabilistic induction of species history must answer all three of these challenges or remain guilty as charged.

Epilogue

If “nothing in biology makes sense except in the light of evolution” (Dobzhansky, 1973), then surely phylogeny sets conditions for all biological study, because nothing evolutionary escapes the historically contingent

¹⁰ This is the kind of description that follows deductively, in which “descent, with modification,” is a necessary part of the *explanans* (Kluge, 2001). It is not the pattern cladistic, instrumentalist, kind of inductive identification (e.g., Brady, 1985), in which “the explanandum (phenomenon to be explained) is logically separate from and prior to its explanans (explanatory hypothesis or ‘justification’)” (Brower, 2002, p. 222).

nature of the process of species diversification.¹¹ Furthermore, in just this latter condition, there is sufficient reason to believe that those who hold to the idea that “the ultimate goal of the study of the evolution of anything is the discovery of the evolution of nomic patterns” (Bunge, 1998) are certain to fail, because there can be no statistical aggregate on which to base a predictive kind of explanation (e.g., see Losos et al., 1998).¹²

To be sure, as Hull (1988, p.424) explained, “[t]he past ‘constrains’ the future only to the extent that it has structured the present” and that requires the genetic and developmental integration of the ancestral species phenotype. While species history may not be predictive of the processes of species diversification, because of its necessarily unique nature, it may nonetheless be understood as *limiting* the phenotype through developmental constraints and stasis. Still, it is perfectly clear that evolutionary nomicists can do no better than identify particular *trends*, like those in development and the fossil record.

Even the health care industry is beginning to recognize that conditional (frequency) probability can have meaning only in the sense of class concepts, like actuarial issues, in which there is a concern for maximizing success in the repeated instances of some kind of syndrome (like the class concept of battle; see above). It is both meaningless and irresponsible to attribute statistical parameters to the prognosis of the individual who is a member of some syndrome, but whose situation is unique, and so cannot be determined a priori as belonging to one effect category or another (Lynn et al., 1997). The very fact that individual patients are now being genotyped in relation to drug responsiveness is long overdue recognition that there are historical (individually inherited) contingencies that are explanatory of prognosis. In other words, even “prognosis” is moving in the direction of individualization (M. E. Siddall, personal communication). Both the scientific and the clinical sides of human health care are possible without assuming homogeneous classes.

Acknowledgments

This paper was first presented, at the invitation of Walter Fitch, to the members of the Institute for Theoretical Physics, University of California at Santa Bar-

bara, in May 2001. Special thanks go to Taran Grant for forcing me to consider more aspects of uncertainty than I had originally intended. Two anonymous reviewers requested my explication of the logic of the testability of phylogenetic systematics, even though what I have presented adds nothing of significance to Farris’ (1983) classic treatment of that subject. Others, including Marc Allard, Jennifer Ast, Ron Brady, Dick Olmstead, Mark Siddall, Mike Sorenson, and the members of my Seminar on Skepticism, University of Michigan, Winter Term, 2002, also severely criticized the central thesis of this paper—that conditional (frequency) probability is inadequate in the inference of phylogenetic hypotheses. Hopefully, this discussion will continue, with the conditional (frequency) probability dogmatists attempting to prove the innocence of their inference. The initial and final drafts of the manuscript were prepared at the Cladistics Institute, Harbor Springs, Michigan.

References

- Brady, R.H., 1985. On the independence of systematics. *Cladistics* 1, 113–126.
- Brower, A.V.Z., 2002. Cladistics, phylogeny, evidence and explanation—A reply to Lee. *Zool. Scr.* 31, 221–223.
- Bunge, M., 1998. *Philosophy of Science: from Explanation to Justification*, 2. Transaction Publishers, New Brunswick, NJ.
- Carpenter, J.M., 1992. Random cladistics. *Cladistics* 8, 147–153.
- Darwin, C., 1859. *The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. Murray, London. [1964 facsimile of the first ed., Harvard University Press, Cambridge, MA].
- de Quieroz, K., Poe, S., 2001. Philosophy and phylogenetic inference: a comparison of likelihood and parsimony methods in the context of Karl Popper’s writings on corroboration. *Syst. Biol.* 50, 305–321.
- Dobzhansky, T., 1973. Nothing in biology makes sense except in the light of evolution. *Am. Biol. Teacher* 35, 125–129.
- Elsasser, W., 1966. *Atom and Organism*. Princeton University Press, Princeton, NJ.
- Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), *Advances in Cladistics II*. Columbia University Press, New York, pp. 7–36.
- Felsenstein, J., 1987. Comment [on Barry and Hartigan, 1987]. *Stat. Sci.* 2, 208–209.
- Felsenstein, J., 2001. The troubled growth of statistical phylogenetics. *Syst. Biol.* 50, 465–467.
- Felsenstein, J., 2002. *Inferring Phylogenies*. Sinauer, Sunderland, MA.
- Ghiselin, M.T., 1966. On psychology in the logic of taxonomic controversies. *Syst. Zool.* 15, 207–215.
- Grant, T., 2002. Testing methods: the evaluation of discovery operations in evolutionary biology. *Cladistics* 18, 94–111.
- Hacking, I., 1965. *Logic of Statistical Inference*. Cambridge University Press, Cambridge, UK.
- Hacking, I., 2001. *An Introduction to Probability and Inductive Logic*. Cambridge University Press, Cambridge, UK.
- Hennig, W., 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana.
- Howson, C., Urbach, P., 1993. *Scientific Reasoning: the Bayesian Approach*, second ed. Open Court, Chicago.
- Huelsenbeck, J.P., Rannala, B., Masly, J.P., 2000. Accommodating phylogenetic uncertainty in evolutionary studies. *Science* 288, 2349–2350.

¹¹ The following is the complete syllogism from which this part is taken: “Nothing in biology makes sense except in the light of evolution. Nothing in evolution makes sense except in the light of phylogeny. Nothing in phylogeny makes sense except in the light of cladistics. Nothing in cladistics makes sense except in the light of total evidence” (Sytsma and Pires, 2001, p. 726).

¹² Nomic refers to the necessary natural law, as contrasted with the contingent rule.

- Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001a. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294, 2310–2314.
- Huelsenbeck, J.P., Ronquist, F., Hall, B., 2001b. An introduction to Bayesian inference of phylogeny. [<http://morphbank.ebc.uu.se/mrbayes/manual.pdf>].
- Hull, D.L., 1967. Certainty and circularity in evolutionary taxonomy. *Evolution* 21, 174–189.
- Hull, D.L., 1974. *Philosophy of Biological Science*. Prentice-Hall International, Englewood Cliffs, NJ.
- Hull, D.L., 1980. Individuality and selection. *Annu. Rev. Ecol. Syst.* 11, 311–332.
- Hull, D.L., 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. University of Chicago Press, Chicago.
- Kluge, A.G., 1990. Species as historical individuals. *Biol. Philos.* 5, 417–431.
- Kluge, A.G., 1997a. Testability and the refutation and corroboration of cladistic hypotheses. *Cladistics* 13, 81–96.
- Kluge, A.G., 1997b. Sophisticated falsification and research cycles: consequences for differential character weighting in phylogenetic systematics. *Zool. Scr.* 26, 349–360.
- Kluge, A.G., 1998. Total evidence or taxonomic congruence: cladistics or consensus classification. *Cladistics* 14, 151–158.
- Kluge, A.G., 1999. The science of phylogenetic systematics: explanation, prediction, and test. *Cladistics* 15, 429–436.
- Kluge, A.G., 2001. Philosophical conjectures and their refutation. *Syst. Biol.* 50, 322–330.
- Lewontin, R.C., 1966. Is nature probable or capricious? *BioScience* 16, 25–26.
- Lidén, M., 1990. Replicators, hierarchy, and the species problem. *Cladistics* 6, 183–186.
- Losos, J.B., Jackman, T.R., Larson, A., de Queiroz, K., Rodríguez-Schettino, L., 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279, 2115–2118.
- Lynn, J., Harrell, F., Cohn, F., Wagner, D., Connors, A.F., 1997. Prognosis of seriously ill hospitalized patients on the days before death: implications for patient care and public policy. *New Horizons* 5, 56–61.
- Mindell, D.P., 1991. Similarity and congruence as criteria for molecular homology. *Mol. Biol. Evol.* 8, 897–900.
- Murphy, W.J., Eizirki, E., O'Brien, S.J., Madsen, O., Scally, M., Douady, C.J., Teeling, E., Ryder, O.A., Stanhope, M.J., de Jong, W.W., Springer, M.S., 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294, 2348–2351.
- Patterson, C., 1978. Verifiability in systematics. *Syst. Zool.* 27, 218–222.
- Platnick, N.I., 1982. Defining characters and evolutionary groups. *Syst. Zool.* 32, 282–284.
- Popper, K., 1957. *The Poverty of Historicism*. Routledge & Kegan Paul, London.
- Popper, K., 1959. *The Logic of Scientific Discovery*. Harper & Row, New York. [1968 edition].
- Popper, K., 1979. *Objective Knowledge: An Evolutionary Approach*. Oxford University Press, New York.
- Popper, K., 1983. *Realism and the Aim of Science*. Routledge, London.
- Popper, K., 1990. *A World of Propensities*. Thoemmes, Bristol.
- Railton, P., 1978. A deductive-nomological model of probabilistic explanation. *Philos. Sci.* 45, 206–226.
- Rieppel, O., 2002a. Beyond the logic of scientific discovery: a theory of empirical classes. In: *Biodiversity in the Information Age. Sixth International Congress of Systematic and Evolutionary Biology*. University of Patras, Patras, Greece, p. 30 [Abstract].
- Rieppel, O., 2002b. The meaning of homology. In: *Biodiversity in the Information Age. Sixth International Congress of Systematic and Evolutionary Biology*. University of Patras, Patras, Greece, p. 195 [Abstract].
- Salmon, W.C., 1966. *The Foundations of Scientific Inference*. University of Pittsburgh Press, Pittsburgh, Pennsylvania.
- Sarkar, S., 1998. *Genetics and Reductionism*. Cambridge University Press, Cambridge, UK.
- Siddall, M.E., 2001. Philosophy and phylogenetic inference: a comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. *Cladistics* 17, 395–399.
- Siddall, M.E., Kluge, A.G., 1997. Probabilism and phylogenetic inference. *Cladistics* 13, 313–336.
- Saber, E., 1988. *Reconstructing the past: Parsimony, evolution, and inference*. Bradford Book, MIT Press, Cambridge, MA.
- Stamos, D.N., 1996. Popper, falsifiability, and evolutionary biology. *Biol. Philos.* 11, 161–191.
- Stamos, D.N., 2001. Quantum indeterminism and evolutionary biology. *Philos. Sci.* 68, 164–184.
- Sytsma, K.J., Pires, J.C., 2001. Plant systematics in the next 50 years—Re-mapping the new frontier. *Taxon* 50, 713–732.
- Walton, D.N., 1985. Are circular arguments necessarily vicious? *Am. Philos. Q.* 22, 263–274.
- Watkins, J.W.N., 1984. *Science and Scepticism*. Princeton University Press, Princeton, NJ.
- Wenzel, J.W., Carpenter, J.M., 1994. Comparing methods: adaptive traits and tests of adaptation. In: Eggleton, P., Vane-Wright, R.I. (Eds.), *Linnaean Society Symposium Series. Phylogenetics and Ecology*, 17. Academic Press, London, pp. 79–101.
- Wiley, E.O., 1975. Karl R. Popper, systematics, and classification: a reply to Walter Bock and other evolutionary taxonomists. *Syst. Zool.* 24, 233–242.