



Taylor & Francis
Taylor & Francis Group

Society of Systematic Biologists

Rainbow's End: The Quest For an Optimal Taxonomy

Author(s): L. A. S. Johnson

Source: *Systematic Zoology*, Vol. 19, No. 3 (Sep., 1970), pp. 203-239

Published by: [Taylor & Francis, Ltd.](#) for the [Society of Systematic Biologists](#)

Stable URL: <http://www.jstor.org/stable/2412206>

Accessed: 11/03/2014 17:00

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

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



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

<http://www.jstor.org>

RAINBOW'S END: THE QUEST FOR AN OPTIMAL TAXONOMY¹

L. A. S. JOHNSON

Synopsis

The aims and justification of taxonomy are discussed, followed by an examination of the foundations of ordination and classification. The Adansonian or phenetic philosophy is critically examined and it is concluded that its claims of objectivity and precision are ill-founded, since subjective or arbitrary choices and definitions are necessary concerning acceptable or relevant attributes, homologies and correspondences, measures and commensurabilities of attributes, and measures of similarity. Phylogeny, represented topologically as a temporal branching sequence, is held to be the nearest approach to a firm basis of reference in nature for biological classification. The charge that phylogenetic reconstruction involves viciously circular reasoning is discussed and rejected, though some positive feedback is admitted.

The theory and application of numerical taxonomy are discussed in general and the potential value of numerical phyletics is stressed. The possibility of using DNA base matching as a solid foundation is briefly examined. It is shown that, while phenetic classifications are infinitely variable, topological phyletic reconstructions do not themselves supply the kind of taxonomy that is usually demanded, since they do not adequately express significant evolutionary changes in patterns of organisation. No optimal classification can be defined, but improvement is possible up to a point of inherent instability.

While the underlying facts and processes can be scientifically studied as part of systematics, classification itself remains largely a disciplined art, which is not convertible to an exact science by any form of arbitrary quantification. The necessity for compromise and continued synthesis is stressed. The suggested replacement of the Linnaean hierarchy by "numericulture" is mentioned but held to be premature and to have serious disadvantages. Mathematical concepts are discussed wherever relevant to the foundations of the subject.

"He said 'I hunt for haddocks' eyes
Among the heather bright,
And work them into waistcoat-buttons
In the silent night . . ."

LEWIS CARROLL (1871)—
"Through the Looking-Glass, and
What Alice Found There."

The "aged, aged man" would seem to have made two misjudgements: he sought unnecessary and rather unsuitable materials for his purpose, and then looked in

a most unlikely place for them. Is the search by some taxonomists for the one "correct" classification of organisms equally futile and misguided? Before we can answer this question we shall need to look rather closely at some of the foundations of taxonomy.

The history of taxonomy and of taxonomic attitudes and methods has been reviewed in detail often enough. A presidential address gives one licence to wander at large over a field of interest. That is what I propose to do, examining certain relevant aspects or opinions, and offering comments and animadversions upon them, in the hope of discovering or deciding what taxonomy can or should hope to achieve. I shall use the terms "taxonomy" and

¹ Reprinted from PROCEEDINGS OF THE LINNEAN SOCIETY OF NEW SOUTH WALES, VOL. 93, Part 1, pp. 1-45 (Issued 19th November, 1968), through the courtesy of the author and the Society. The paper was originally a Presidential Address, delivered 27 March 1968.

The following article is reprinted from another journal, circulation of which in North America is limited to a few persons and institutions. That the editors of *Systematic Zoology* would break with the tradition that reserves space in the journal only for original works not printed elsewhere hopefully suggests that the article has unusual significance. In fact, the author has done an informed and wide-ranging synthesis and has provided us with a unique and lively view of contemporary taxonomy that should generate considerable interest in many quarters.—Ed.

“systematics” as roughly equivalent, with a tendency towards the principles-*versus*-practice distinction made by some American authors (e.g., Simpson, 1961). Except in passing, I shall mean by taxonomy the classification of organisms and not, as is fashionable today, of automobile components, criminals, Latin texts (Griffith, 1967), or heraldic beasts. In order to achieve reasonable precision, I shall have occasion at times to use some of the elementary language and concepts, but little of the notation, of certain aspects of mathematics, in particular of set theory and higher geometry.

THE STATUS AND JUSTIFICATION OF TAXONOMY

Taxonomists are often heard to complain that other biologists accord them insufficient respect, that universities pay little attention to taxonomy, and so forth—in fact that they are kept in subjection in a scientific peck order in which biology itself ranks none too high, molecular biology apart (for the obvious reason that molecular biologists are asking simpler, and in a sense genuinely more fundamental, questions which can be answered—sometimes—even by physical scientists). I shall now risk the wrath of my colleagues and assert that, although there is indeed some cause for these complaints, not a few non-initiates pay the taxonomic priesthood more respect than we perhaps merit. The historico-nomenclatural system, which is probably much less justified and immutable than systematists are wont to claim, burdens taxonomy with an enormous dead-weight of out-dated and sometimes very bad work, professional and amateur, old and new. In any other non-historical discipline such publications would lapse into the obscurity they deserve.

Although systematics continues to attract more than its fair share of unshakable conservatives and legalists, much of its accumulated dross must eventually be swept away. If it is not, we may find that tra-

ditional systematics has been by-passed, whether we like it or not. I say this as one who has rather punctiliously followed the rules of nomenclature, though with increasing impatience—not at the existence of rules, but at the stultifying consequences of those we have.

Despite these criticisms, much systematic work of the past is in fact not out-dated—just as Pythagoras's theorem is as valid today as ever, or, less absolutely, as Newton's mechanics is all that we need to cope with those many familiar problems where the scale is such that the answers given by quantum mechanics and relativity theory would not differ, at the order of accuracy required, from the Newtonian answers. Also, following a becalmment in the doldrums during the first forty years of this century, systematics was able to recruit more workers possessing sound scientific training and some idea of the complexity and dynamics of the populations from which their specimens are drawn. Taxonomic research improved markedly as a result.

As we all know, this revivescence was associated with the progress of cytogenetic and evolutionary theory, and in its early days it first found full expression in the symposium “The New Systematics” (ed. Huxley, 1940). Much earlier, the Darwinian revolution had imbued many taxonomists (chiefly those whose education was not completed before the establishment of evolutionary thinking!) with the notion that if the phylogeny of a group could be successfully reconstructed, the taxonomic problems concerning it would be solved. Whether or not classification was thereby improved, some life certainly came into the subject for a few decades. Other systematists toiled on, sorting and shuffling, without much theoretical interest at all and, when favoured by nature with clear-cut situations, some among them produced classifications which continue to satisfy most users. There are groups in which nobody cares except the specialists them-

selves. Sometimes the technicality, obscurity, and dullness of the specialists' esoteric works have helped to bring about and preserve this state of blessed isolation.

In general, whether good or bad, theoretically disposed or otherwise, taxonomists themselves, like other scientists and scholars, find their own activities self-justifying because they bring some degree of intellectual satisfaction. This is doubtless the chief motivation for most of us, whatever rationalizations we offer about usefulness when called upon to show why society should support us. Nevertheless, society does support us, niggardly though we claim this support to be. Why? Partly because, from sheer inertia, the dispensers of funds will usually keep a going concern alive, but also, as any worker in a systematic service organization can testify, because there is a genuine call, from science and industry as well as many other activities of man, for the services of the namer and classifier.

THE NEED TO CLASSIFY

None of us can communicate, or even live, unless we sort or classify the phenomena of our experience; further, we need in some way to label the classes or sets to which we assign these phenomena (for an introduction to the mathematical theory of sets see, for example, Room and Mack, 1966, and for an application to taxonomy see Buck and Hull, 1966). Often we know that someone else, the specialist classifier, is more efficient at this task than we are, and so we avail ourselves of his services. Having classified and named (labelled) a phenomenon or group of phenomena, we can ascertain further information about it and can talk about it. We may even, naïvely, consider that we know what it is. How often an inquirer asks, "What is this organism?" and, on being told "*Melania alba*", considers himself somehow wiser than before. Anthropologists know well the supposed power conferred by knowledge of a name.

The role of taxonomy in education is emphasized by Davis and Heywood (1963), who say that it "remains the principal agency through which the student may gain acquaintance with the diversity of organisms, the patterns of variation in the living world and . . . the evolutionary mechanisms which have brought these patterns about. The taxonomic approach is a focal point for students of biology and enables them to fit together into a framework of ideas a mass of otherwise uncoordinated facts".

The classificatory nature of the whole of science is sometimes stressed. Many of my generation will remember from their schooldays a textbook of chemistry by Sherwood Taylor (1939), which begins with the words: "Natural science is the process of systematically arranging and classifying man's knowledge of the world about him." Perhaps this only partially defines the nature of science, but the search for more inclusive and if possible more simply and generally describable classes, of relations as well as things, is an important factor in arriving at the general but precise "laws" (as they are so inappropriately called) which science seeks.

THE QUESTION OF OPTIMALITY

So much for the need for some sort of taxonomy. A chain of questions then arises: Are some classifications better than others? If so, when does the improvement (assuming that we have some means of assessing it) justify changing an existing system—or, how are we to choose between co-existing systems? If improvement is possible, then is there, logically, a *best* classification (practically attainable or not) to which we can try to approximate? Finally, if there is an optimal classification, what is its basis?

As soon as we attempt to answer these questions, or even to define the terms used in them, we find ourselves faced with many more questions. In any field where definitions are required we are led into an infinite

regress; so we must agree to take certain things for granted—to set up starting points, as it were. In the more rigid contexts of logic or mathematics one starts with certain postulates, but even in the expression of these one must perforce use certain terms and syntactic relationships of undefined meaning. For example, in modern mathematics, the concept of “point” is usually left undefined; likewise, logical links such as “it follows that” must be accepted more or less intuitively (Courant and Robbins, 1961).

Much has been written on the philosophical bases of classification, including biological classification. While this literature is academically interesting and has important implications for essentially simple (but not always easy!) domains of inquiry such as mathematics, its application to essentially vague (i.e., very highly complex) domains such as biological taxonomy seems to be of little practical assistance. Most practising systematists continue their studies without worrying about philosophical foundations and it appears that, in their own work, many philosophically minded authors proceed in much the same way as other competent systematists. In contrast, evolutionary theory has noticeably influenced the methods, approach, results, and particularly the interpretations, of those systematic workers who are strongly aware of the genetical and evolutionary background and implications of their work, whether they call themselves “biosystematists” or eschew such labels.

THE PHILOSOPHY OF PHENETICISM

Despite this lack of practical effect, the modern philosophical school of taxonomy, foremost among whose early spokesmen were the botanist J. S. L. Gilmour (e.g., Gilmour, 1940), now at Cambridge, and (somewhat later) the Oxford zoologist A. J. Cain (e.g., Cain and Harrison, 1958; Cain, 1959), has had an increasing appeal over the past thirty years to theoretically minded taxonomists (especially those edu-

cated in certain universities). The background to this way of thought is the “operational” approach of logical positivism, a more far-reaching anti-metaphysical philosophy than empiricism but, like empiricism, of obvious appeal to the scientific mind (Britton, 1958, and references therein and in Gilmour and Walters, 1954, and Carolin, 1967).^{*} It has culminated in the so-called Adansonian, neo-Adansonian, or phenetic credo which, perhaps by historical accident, is now so closely associated with what is labelled numerical taxonomy (Sneath and Sokal, 1962) or taximetrics (Rogers, 1963). The strong influence of the phenetic viewpoint is exemplified in the generally well-balanced and comprehensive modern textbook of angiosperm taxonomy by Davis and Heywood (1963), where its merits are accepted as almost self-evident.

Strictly, *Adansonianism* advocates multi-attribute classification, grouping on the basis of many equally weighted attributes drawn from many parts of the organism,

^{*} To define my own philosophical standpoint, I should perhaps say that, while preferring logical positivism to metaphysical philosophies, I find it entirely reasonable to believe in the material reality of the universe—material, of course, in the sense that, despite the Uncertainty Principle, we can learn a great deal about the behaviour of “matter-energy” at various levels of aggregation and organization, and that the fundamentals (though not always the details) of this behaviour are independent of the existence of ourselves or other observers. As to the place of purpose and particularly of life in this physical world, I can see no justification for any form of supernaturalism, mysticism, transcendentalism, or neo-vitalism, nor can I see anything inconsistent with physics in the fact that, in the sense of information theory but not of physical thermodynamics, organisms increase negentropy while they grow as individuals or populations. Life is the state of possessing mechanisms for self-perpetuation, replication, and (in a sense) occasional increase of the organization of systems of physical components; that is the only way in which it is “different”. Explicitly or not, these views are probably shared by most biologists. Neo-vitalism today seems often to arise, curiously enough, among physicists (e.g., Wigner, reviewed by Pais, 1967).

and refusing to attach greater weight *a priori* to certain "essential" characters. Michel Adanson's approach (Lawrence, 1963), in the context of the eighteenth-century Enlightenment, represented a deliberate break from the blend of debased Platono-Aristotelian philosophy and theology which constituted mediaeval Scholasticism. Although it was abandoned by progressive philosophers, Scholasticism continued to obfuscate some scientists' thought long after the Middle Ages and its outward forms, at least, were preserved in the concepts of "genera", "species", "characters", and "differentiae", as well as in the doctrine of divine special creation, accepted by Linnaeus and other orthodox taxonomists of his day (see also Hull, 1965). In practice, many good systematists were partial Adansonians long before the term came into its present vogue. Phylogeny, as we understand it, was a concept unknown to Adanson, though he was no divine-creationist, but modern *pheneticism* not only purports to be stringently Adansonian but also emphatically rejects the use of phylogenetic considerations in reaching taxonomic conclusions (that is to say, in arriving at particular classifications). However, its proponents are mostly careful to state that they do not reject subsequent or independent phylogenetic interpretation, or "speculation" as many prefer to put it. Most pheneticists also reject taxonomic use of the "biological species concept" developed, with variations, in numerous publications by Dobzhansky, Huxley, Mayr, Stebbins, and others, and strongly advocated in rather purist terms by Åskell Löve (e.g., Dobzhansky, 1951; Huxley, 1942; Mayr, 1939, 1957*a*, 1957*b*, 1963; Stebbins, 1950; Löve, 1964).

Since many evolutionary taxonomists and biosystematists have thought that pheneticism unjustifiably neglects the important gains in understanding contributed by their approach to biology, a spate of discussions has ensued, some of them strongly polemical, in which various issues have been

debated and frequently confused. Traditional taxonomists have joined in with more pragmatic arguments and from fear of displacement or interference by technical innovations which, perhaps often rightly, they regard as unnecessary. On the whole the pheneticists have spoken more loudly and with the confidence of revolutionaries who sense that the *Zeitgeist* is on their side. In the course of a decade the numerical pheneticists have come to hold the centre of the stage. Such successes are at times due to salesmanship rather than scientific merit, and we shall do well to look critically at the foundations of phenetics. Time and performance, of course, will provide the acid tests.

I shall defer most of the discussion of numerical techniques in taxonomy until we have dealt with the more general subject of systematizing the objects of our experience.

THE NATURE OF ORDINATION

As several authors (e.g., Williams, 1967) have recently reminded us, an orderly arrangement of objects, or more usually of the symbols representing them, need not be a classification; we may settle for an *ordination*, that is, we may assign relative positions to objects according to their states with respect to a set of their "attributes", a separate dimension being necessary for each attribute. This procedure establishes an "attribute-space" (usually abstract), in which the objects are represented by points or sometimes regions. Obviously, to be meaningful in the ordination, each of these attributes must be represented by at least two states in the set of objects under consideration; in other words, they are "pluri-state attributes".* An example of the simplest non-trivial case would be the ordering in one dimension of, say, men according to their measurements

* I shall use the prefixes "pluri-" and "multi-" to signify respectively "more than one" and "more than two" (or "more than three" for dimensionality).

(states) of the attribute *height*, or of events according to *sequence in time* (unless relativistic considerations are negligible the latter case requires specification of a frame of reference). Two or more objects may occupy the same position in the attribute-space. They are then indistinguishable with respect to the attributes concerned, but we may still regard them as distinct entities by reference to other attributes.

Ordination *does* involve some implicit classification, as we shall see. Consider a set of objects each possessing what we may call “elementary attributes”. For the moment we may define the latter as any properties we can and wish to specify, for example six-leggedness, a length of four millimetres, or orange colour. Before we can assign positions to the objects in any ordination we must establish working homologies between them. This implies setting up one-to-one (pluri-unique) correspondences, over the object set, between some of the elementary attributes of the objects. The result may be more or less reasonable according to the circumstances. It is reasonable to establish a correspondence between the surface areas of a moss leaf, a cycad leaf, and a lycopod leaf if we are interested in photosynthetic capacity—it is scarcely reasonable if our interest is in phylogeny, since most botanists regard these leaves as evolutionarily non-homologous. The establishment of a correspondence ranging over the object set, whether reasonable or not, is tantamount to a classification of the attributes: we have assigned a certain one of the elementary attributes of each object to a class. From the property by which we define this class we derive a pluri-state attribute applicable to all members of the object set; the elementary attributes then become the individual “states”.

The property “measurable length” is here regarded as an abstraction from the set of actual measurable lengths; the very definition of this set implies a classification of these elementary attributes on the basis of

a property common to them all (the property of being a *length*). We could look at this relation in other ways: for instance, we might regard the possession of a particular length as dependent on having length at all; again, the particular lengths are attributes of the individual objects while length is a property of all the objects.† It is unfortunate that the term “attribute” (or “character”) is used in the literature at these different levels. In what follows it will be necessary to use it in both senses but either the context or some qualifying term should make the particular application clear.

It has been assumed in the preceding that for a single object the “states” should be mutually exclusive—ordinarily a leaf cannot have more than one measurement of length (at any instant). When the actual observations are made upon *parts* of the objects under ordination (or classification) the state may be expressed by some suitable statistic, for example, the mean leaf-length or the largest observed leaf-length (the object here being a plant, a population, or a leaf over an extended time). Although this procedure will simplify the ordination, it will result in loss of information. Alternatively, the objects (being composite with respect to the attribute or attributes concerned) may be represented in the attribute-space not as single points but as regions, which may overlap. For complete specification we then require an evaluable density function. For any one attribute (dimension) this

† When appropriate, the zero state must be included in the range of the pluri-state attribute. Here length, even if sometimes zero, is considered to be a property of all the objects if a comparison on the basis of length is meaningful. Thus, for a snake, it is reasonable to treat the length (zero) of the external limbs as a comparable attribute to the non-zero lengths of such organs in other reptiles; on the other hand for an amphioxus, say, the concept of a measure of length, zero or not, of limbs is inapplicable. In the latter case, if comparing with other chordates, the two-state attribute “presence or absence of limbs” might be appropriate.

may or may not be easily expressed. If the distribution of variation is Normal, we need to specify only (estimates of) the mean and the standard deviation (or the variance); if it is non-Normal but follows some other regular pattern it may also be expressed in terms of the distribution function and a small number of parameters. Sometimes a logarithmic or other non-linear scaling transformation may yield such a simple distribution. However, frequency (= density) distributions of variation may be quite irregular, or plurimodal, or discontinuous (e.g., flowers 4-merous 21%, 5-merous 79%). Moreover, since more than one attribute is usually involved, the *covariation* within the composite object also becomes important. Thus, the pluri-dimensional density functions may be very complex indeed, and the extent to which they should be simplified for composite objects (e.g., species or higher taxa) will be a matter of judgment and practicability in each case.

Serious difficulties are introduced into the dimensional representation if some (pluri-state) attributes depend for their expression on the existence of certain states of others, and are therefore not uniformly relevant over the whole set of objects. These circumstances result in the attribute-space being nonhomogeneous as to dimensionality (see also Reynolds, 1965). In numerical taxonomy this is a practical problem which has usually been somewhat unsatisfactorily avoided by various shifts and devices such as redefinition of attributes, or by weighting methods (Kendrick, 1965) which most numerical pheneticists frown upon (see Long, 1966, for a facile dismissal and misunderstanding of the serious dependence problem raised by Kendrick).

Still other problems arise when "states" which it seems reasonable to group into a single "attribute" do not admit of an unequivocal measure by which they may be arranged serially, for example, variants of a chemical constituent which differ in re-

placement-groups. A representation which preserved the symmetry of such cases, and in which they could be combined with serial (or two-state) cases would call for a space wherein some attributes were expressible in terms of linear co-ordinates and some in terms of higher-order symmetrical (e.g., triangular, tetrahedral, etc.) co-ordinates. This of course implies dimensions within dimensions and enters realms of complexity which, though perhaps beyond practical handling and certainly introducing anisotropic properties into the spaces concerned, should not be glibly passed over, since they *are* inherent in the general problem of ordination. Anticipating matters to be discussed later, in numerical *classification*, which is not necessarily subject to the same practical restrictions as to the spaces implied in the models used, Lance and Williams (1967c) have devised an information-statistic computer strategy to deal with mixed data including non-exclusive, non-serial,* multi-state attributes. Still more recently, Wallace and Boulton (1968) have developed another mixed-data strategy.

Although ordination implies assignment of relative position, it is not essential for it to be metric; no particular fixed measure is obligatory. The relation expressed by saying that, in a one-dimensional linear (in the sense of non-closed) space, B lies *between* A and C does not require either that we know just how far B is distant from A and C, or that it has any fixed position, or indeed that a measure of distance has any meaning at all. Here three distinct points will define an ordering, without any specification of direction in the space. If a direction is assigned, two points A and B are sufficient to define an ordination. Further, ordination remains mathematically meaningful, though difficult to handle, if (in each or any of the dimensions) the arrangement is cyclic

* "Non-serial" seems preferable in this context to "disordered", as used by Lance and Williams.

(then, however, we should require a minimum of four points in the non-directional non-metric case, three in the directional non-metric) or in some other way not simply linear. Moreover, no concept of a continuum is necessarily required; there may be only discrete positions. Thus, the abstract space in which we represent an ordination may differ from our everyday concept of physical space in one or more of five ways:

- (1) it may be finite,
- (2) it may be more than three-dimensional,
- (3) it may be non-Euclidean,
- (4) it may be non-metric,*
- (5) it may be discontinuous.

It must, however, possess certain topological properties. In loose terms: however we may deform our representation by variation of scales in any direction, the ordination must not be altered. In any but the simplest cases it is no easy matter to elucidate and explain the topological and metrical properties of the infinite variety of definable spaces; *this is no reason to assume*, as is often done, *that simple spatial models are particularly appropriate to taxonomic ordination*. (For an introduction to some abstract spaces see, e.g., Sawyer, 1955, and Courant and Robbins, 1961; there are many more advanced texts, for instance that of Kelley cited by Williams and Dale, 1965).

Non-mathematicians are conditioned to believe that there is something "natural" about Euclidean space and the Euclidean metric. The latter means, roughly, that if we set up rectangular, similarly-scaled, Cartesian co-ordinates in an n -dimensional space then the distance between any two

points P and Q is given by the Pythagorean function

$$d(P,Q) = \left[\sum_{i=1}^n (x_{iP} - x_{iQ})^2 \right]^{\frac{1}{2}},$$

where x_{iP} , x_{iQ} are the co-ordinates of P , Q respectively in the i th dimension. This is an extension to n dimensions of the familiar theorem of Pythagoras that the length of the hypotenuse of a right triangle is the square root of the sum of the squared lengths of the two orthogonal sides. (Oblique co-ordinate axes may also be set up in a Euclidean space, and are used with the Mahalanobis "generalized distance" measure (Sokal, 1965; Menitskii, 1966) though Kendall (1957) and Williams and Dale (1965) have challenged the validity of the procedure in respect of this particular statistic.) However, from what we may know of cartography and navigation most of us will admit that something rather different is involved in the geometry of the surface of a sphere. We customarily think of a sphere as an object in three-dimensional linear (Euclidean) space. Mathematically, however, its surface geometry can be considered in complete isolation as a closed, curved, two-dimensional space which we can, if we wish, embed in a three-dimensional Euclidean space. We shall probably also have heard that representation of space-time in the General Theory of Relativity calls for a particular four-dimensional case of a strange collection called Riemannian spaces, which constitute a more general class of curved spaces. These notions of naturalness arise in the ordering of our spatial experience of the physical world, but we are accustomed also to graphical representation of non-spatial quantities in co-ordinate systems in what appear to be Euclidean spaces. Actually, such representations often do not imply, or allow, any concept of inter-point distance involving more than one co-ordinate, and there is then in fact no implication of a Euclidean (Pythagorean) metric. As an extension from such

* Conditions exist in which there is not a complete absence of metrical properties; semi-metric, quasi-metric, and disjoint metric spaces all arise in numerical taxonomy (Williams and Dale, 1965).

representations, the concept of abstract Euclidean spaces of more than three dimensions has become familiar, as has the correspondence (isomorphism) of algebra and geometry, first clarified by Descartes, which freed geometry of the need for pictorial representability.

Euclidean metrics are implicit in most of the models used in classical statistics, including such probabilistic techniques of multivariate statistics as principal component analysis, factor analysis, and canonical analysis, which depend on the method of least squares (e.g., Kendall, 1950, 1957; Seal, 1964). Because of these familiarities and the convenient pre-existence of statistical techniques, some authors on taximetrics (e.g., Sokal, 1961, 1965; Sokal and Sneath, 1963; Goodall, 1964; Jancey, 1966) have shown a strong preference for Euclidean models. Sokal (1961, p. 73) and Boyce (1964) speak (nonsensically in the context of comparison of attribute sets) of the Pythagorean measure as "true distance" when contrasting it with the "mean character distance" (M.C.D.), used by Cain and Harrison (1958). The M.C.D. in fact defines what may be called a lattice metric—the distance function here is

$$d(P, Q) = \sum_{i=1}^n |x_{iP} - x_{iQ}|$$

(ignoring the M.C.D.'s scaling factor $\frac{1}{n}$)

which means that the shortest distance between two points is, in general, "around the corners", as for a rook's moves in chess ($n = 2$) or an ant travelling along the bars of a children's "jungle-gym" ($n = 3$). Such a lattice-metric space can be embedded in a Euclidean space; it differs from the latter in that, given a basis of rational numbers, no concept of algebraic irrational numbers is needed to express distances within it, nor are there any smooth curves or transcendental irrationals like π . Unlike those of a Euclidean space, lattice-space distances are not invariant under rotation of axes, indeed in general

(with a rational-number basis) the "fixed" points themselves will "disappear" under such rotation. These details are not relevant to practical taximetrics but are mentioned to illustrate that some of the mathematical properties which we take for granted in the Euclidean metric can change drastically with a very simple change of metric. Lance and Williams (1967c) refer to this lattice metric as the "Manhattan metric" and point out that it is the first-order case of a general class (Minkowski metrics), the Euclidean metric being the second-order case.

The infinite-order case is of interest since it is the "chess-king's metric" in which

$$d(P, Q) = |x_{iP} - x_{iQ}|_{\max},$$

that is, the distance is simply the value of the greatest single co-ordinate difference between the points.

There is no *a priori* reason why such a metric, or indeed many others, should yield a less "realistic" measure of "distance" (or its complement, similarity) between sets of attribute-states. Nevertheless, even the mathematically sophisticated Williams and Dale (1965) favoured Euclidean systems, on grounds of statistical convenience and "pictorial" representability (which fails in any case due to the distortions and indeed the semi-metric property of projective mappings of higher-dimensional spaces on to their lower-dimensional subspaces, i.e., originally distinct points may not be distinct in the mapping). Jancey (1965) suggests that the intuitive concept of similarity in terms of real spatial relationships renders the Euclidean metric worthy of retention. If this merely reinforces preconceptions, its value seems doubtful and, as Macnaughton-Smith (1965) says, "this 'visualizable' quality is by no means a necessity, and one would wish to use the most appropriate function regardless of whether it was visualizable". More recent papers on taximetrics (e.g., Lance and Williams, 1967c) show less attachment to Euclidean representations.

For particular purposes, ordination procedures may be more efficient than classifications in that they need not involve us in loss of information, nor do they set up so many "artificial" distinctions or segmentations of gradual transitions. Their usefulness will naturally depend on the appropriateness to our purposes of the attributes chosen, and on the measure adopted for assigning position in each dimension. Incommensurability in the various dimensions is a problem only if we are interested in such concepts as similarity over a multiplicity of attributes or "distance" between objects (points or regions) in the attribute-space. Perhaps, if our mental apparatus were differently organized, multidimensional ordination would meet most of our requirements for the organization, retrieval and comparison of information. This is a question for those studying the design of logical and quasi-mental machines.

THE NATURE OF CLASSIFICATION

As it happens, our minds cannot cope with multidimensional systems and we need to classify, even though we lose information and introduce distortions and artificialities in the process.

In dealing with the chain of questions posed earlier, it will be useful to refer frequently to one of the most definitive statements of the "philosophical" attitude, a contribution by Gilmour and Walters (1964) entitled "Philosophy and Classification". I should make it clear that, although I shall offer some severe criticisms of the purely phenetic approach, a good deal of what Gilmour and Walters have to say is not in dispute and that they are aware of of some of the unresolvable aspects of the problems raised.

It has been repeatedly stressed, especially by Gilmour and his followers, that the values of classifications should be assessed according to the range of their purposes. We need not here discuss classifications of narrowly-defined purpose, although many of the problems associated with more gen-

eral classifications may arise, *in parvo*, in the special cases also.

Biological taxonomy is expected to produce classifications of broad utility; Gilmour and Walters have termed these *general-purpose classifications*. More essentialistic taxonomists have aimed to produce so-called "natural" classifications which, it is claimed, would in part meet the requirements of a general-purpose classification but would also, as it were, reflect some more fundamental truth about nature. As Gilmour and Walters point out, there are philosophical objections to the term "natural" (and its opposite, "artificial") in this usage. They say: "the view that there are such 'natural kinds', differing in some 'fundamental' way from other, 'artificial' methods of classifying the same objects, is very difficult to sustain; a more useful way of looking at the situation is that these so-called '[natural] kinds' are classes showing [a] high degree of correlation of attributes, differing only in degree from other classes with a less high correlation". (The second part of this statement raises difficulties, discussion of which is deferred.) Later, they outline eight principles which they suggest should be applied to biological classification. Since these represent, in summary, an influential attitude, I shall quote them in full as a framework for critical comment.

"(1) The term 'classification' is used by philosophers to describe the act, conscious or unconscious, of grouping objects into classes because of certain attributes they have in common."

Comment: We may have difficulty in defining "objects" and "attributes" but the definition will serve. Here, and in what follows, these authors use "attributes" to include what I have called "elementary attributes" rather than "pluri-state attributes". I shall follow their usage while it remains relevant.

"(2) Classification, used in this sense, is man's basic method of dealing with

the multiplicity of individual objects in the world around him.”

Comment: Whilst the statement is true enough, its restriction to “man” is unnecessary and symptomatic of an anthropocentric viewpoint which pervades philosophy. Other animals classify, so do some machines; certainly any highly intelligent being would do so (though, as I have said, some might be less compulsive classifiers than ourselves). The process of classification does not depend on the existence of man. It does, I suppose, depend on the existence of an *agent* of some kind—“*classifico ergo sum!*”—but we may be led into unprofitable by-ways if we pursue this further.

“(3) Since classification is a product of man’s need to deal with his environment, the actual classifications that he makes are determined by his desires and purposes in relation to that environment.”

“(4) The suitability of any particular classification can only be judged in relation to the purpose for which that classification was made.”

Comment: This is not necessarily so; a classification made for one purpose may suit another very well. It should be judged in relation to the purpose for which it is *required*.

“(5) Two types of classification can be distinguished, with every gradation between them: ‘general-purpose’ classifications and ‘special-purpose’ classifications.”

“(6) General-purpose classifications consist of classes containing objects with a large number of attributes in common, thus making them useful for a wide range of purposes; special-purpose classifications consist of classes containing objects with only a few attributes in common, and hence serve a more limited range of purposes.”

Comment: In a “common-sense” way, we doubtless all think that we can see what is meant here. However, the concept of “number of attributes in common”, or more sophisticated measures of similarity, lies at the heart of the phenetic (Adansonian) approach, and of the mathematical methods which have been developed on the basis of this approach. Therefore we must examine this statement more critically. There are several hidden variables here, and unless we can somehow define, measure, and control them, the statement simply will not do as a basis for the precise, “objective” approach to taxonomy which many pheneticists state to be their aim.

First, we must consider the domain from which the attributes are to be chosen. Every “object”, a term which may be extended to cover any physical or conceptual object of discourse or thought, has an infinitude of attributes. That is to say, the object itself does not set any bound beyond which we can say “no further attributes exist”; this conclusion is not affected by the practical limitations of our thought. (We can ignore such playthings of paradox-fanciers as “the concept which has no attributes”.) A little reflection will show, for instance, that any object bears various relationships, tenuous as they may be, to every other physical object or collection of objects in the universe, at every point on the world-line of every particle in space-time. There are likewise relationships to the past and future states of the object itself, and indeed to innumerable abstract concepts. Perhaps this sounds extreme, but it is an inescapable conclusion from any general concept of an attribute, and there is no *a priori* reason to stop at any particular point in our search for further attributes. It may be objected that relations to other objects are not *intrinsic* attributes of the object under consideration. Reflection will show that *all* describable characteristics are relations to other objects or concepts; we simply cannot speak meaningfully about the properties of a thing in itself, Plato or

Whitehead notwithstanding. Likewise, it is of no practical help to adopt a holistic standpoint and claim that everything is part of one great integrated whole and that the individual objects and their attributes are mental abstractions from this whole—this can only lead one to say “everything is as it is” and to resign from the game. *Second*, the considerations just stated will show also that the attributes themselves are infinitely divisible. Our concept of “elementary attributes” implies no atomicity. *Third*, there can exist no absolute measure of similarity (i.e., matching correspondence) between non-identical sets of attributes which are infinite, unbounded, and unconstrained.

Infinity is no simple subject, but we are now trapped into considering some consequences of invoking it. We find that measures of matching can in fact exist between certain infinite sets. Let us consider, for instance, the set of all positive integers A: {1, 2, 3, 4, . . .} and the set of odd positive integers B: {1, 3, 5, 7, . . .}, and regard the elements of these sets as our selected elementary attributes of two “objects” which, being here equivalent to the sets described, we may also label A and B. Let us then first take these elements in the order given and, by an act of classification, group them in pairs (1,1) (2,3), as “states” of a set G of pluri-state attributes which we shall call {I, II, III, IV, . . .}. We could define the “possible” states of the members of G in various ways; one way would be: the “states” of I comprise the presence of 1 or 2 or 3 or in the first position, likewise for II in the second position, and so on. This would permit the comparison of a whole class of sets similar to A and B. Each of the sets A, B, G is infinite but, in the language of the Cantorian theory of transfinite numbers, denumerable (Courant and Robbins, 1961; Dantzig, 1962).

If we compare all the attributes, or any ordered sample of them, and write 1 for a match, 0 for a non-match, we have the correspondence:

G: I II III IV
 A: 1 2 3 4
 B: 1 3 5 7

giving the matching sequence: 1 0 0 0, that is, a simple matching coefficient (Sokal and Sneath, 1963), if we start from 1, of

$$S_{AB} = \frac{\text{matches}}{\text{matches} + \text{non-matches}} = \frac{1}{n},$$

where *n* is the number of attributes being measured. Obviously, as *n* tends to infinity, *S*_{AB} tends to zero. If we start elsewhere we have *S*_{AB} = 0 in any case. This is not very helpful but, with what follows, it illustrates the point that it is not only the attributes used, but also the way in which they are grouped and arranged, which affects the results.

If we now look at these sets in another way (one of an infinity of ways) and take as our set H of two-state attributes {i, ii, iii, iv, . . .} the presence or absence (anywhere in A or B) of the numbers 1, 2, 3, 4, (this is in effect establishing a new set B' from the union of B with the set {absence of 2, absence of 4,}), we have the correspondence:

H: i ii iii iv
 A: 1 2 3 4
 B': 1 - 3 -

Then, by taking the whole or any ordered sample, we have, scoring for matches as before, 1 0 1 0 1 if we start with an odd number, or 0 1 0 1 0 if we start with an even number, giving *S*_{AB'} = ½ for any ordered sample of an even number of attributes from H and *S*_{AB'} → ½ (as the sample size increases) for any ordered sample of an odd number of attributes from H. We can indeed say that *S*_{AB'} = ½ for the comparison (made in this particular way) of the two infinite sets of elementary attributes. Moreover, we can say that statistically the most probable estimate of *S*_{AB'} from any random

sample from H is also $\frac{1}{2}$,* and can in fact evaluate the probability of obtaining this or any other result for a sample of any given size. We have, in fact, for our defined sets and procedure, a parameter $S_{AB'}$ and a probability distribution for sampling error.

The infinite sets A, B, and B', just discussed, are in fact (if considered as linearly ordered sets) bounded in one direction and not in the other, but it would not affect the argument if we took the sets of all integers and of all odd integers (positive, zero, and negative), which are not bounded in either direction. What is more, we may calculate the matches between pluri-dimensionally ordered sets, such as the two-dimensional arrays:

$$C: \begin{Bmatrix} 1 & 2 & 1 & 3 & 1 & 4 & \dots \\ 2 & 1 & 3 & 1 & 4 & 1 & \dots \\ 3 & 4 & 3 & 5 & 3 & 6 & \dots \\ 4 & 3 & 5 & 3 & 6 & 3 & \dots \\ 5 & 6 & 5 & 7 & 5 & 8 & \dots \\ 6 & 5 & 7 & 5 & 8 & 5 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{Bmatrix} \text{ and } D: \begin{Bmatrix} 2 & 2 & 3 & 3 & 4 & 4 & \dots \\ 1 & 2 & 1 & 3 & 1 & 4 & \dots \\ 4 & 4 & 5 & 5 & 6 & 6 & \dots \\ 3 & 4 & 3 & 5 & 3 & 6 & \dots \\ 6 & 6 & 7 & 7 & 8 & 8 & \dots \\ 5 & 6 & 5 & 7 & 5 & 8 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{Bmatrix}$$

giving the “obvious” matching coefficient of $S_{OD} = \frac{1}{4}$, though by considering other attributes derived from these sets (e.g., presence or absence of a prime number in a given position) other similarities would be obtained. Such comparable arrays may be bounded or unbounded in any direction, provided that they are conformable or can be rendered conformable. If there is no “natural” bound, one must of course choose appropriate or arbitrary starting points for a matching procedure.

Thus, on such a model, having set up a particular correspondence and using any particular measure of similarity, there *does* exist an actual parametric value of similarity over these infinite sets which we may hope to estimate in various ways. Why, then, do I say that in the general case of

* Strictly, for odd-numbered samples the *two* most probable values are those two of the possible values nearest on either side of $\frac{1}{2}$, and these converge on $\frac{1}{2}$ as the sample size increases.

all attributes, no such parameter exists—in other words, that no meaning can be attached to the concept of such a parametric value? The reason is that our examples possessed underlying (“built-in”, if we like) regularities or patterns which pervaded each of two conformable sets and that, furthermore, we could find a simple and pervading relationship between these two patterns themselves, even though the sets were infinite. We had, in fact, a set-up case with several constraints. In the all-attributes case, no such pervading regularity can exist, since there are no constraints to prevent us bringing in any number of other infinite sets of attributes (elementary or pluri-state) to swamp into infinitesimal proportions any regularity existing among some of the attributes. As a further complication, describable attributes may be dependent upon the existence of other attributes, for example blood pigments cannot be compared as between a bloodless animal and one with blood. This is one aspect of the “no-comparison” problem familiar in the taximetric literature, and I have alluded to it in the discussion of ordination.

It is abundantly clear that, because of such fundamental difficulties, any recourse to an all-attributes concept or to “true” or overall similarity leads us into a hopeless morass. We can, of course, though not without problems, arrive at numerical answers purporting to indicate degrees of similarity (or “distance”) if we restrict ourselves to finite sets of attributes or to samples from infinite sets of certain restricted types, as we have seen. This involves subjective decisions, as to:

- (i) the set of objects considered usefully comparable;
- (ii) the domain of attributes which we consider relevant to our interest in the objects;
- (iii) the “fineness” with which we analyse the features into elementary attributes (\equiv states);

- (iv) the establishment of equivalences or homologies between parts of the objects under comparison; and the consequent grouping of the elementary attributes into two- or multi-state sets, thus specifying what we usually term "the attributes" or "the characters" ("multi-state" here includes "continuously-varying");
- (v) the method and intensity of sampling of the objects;
- (vi) the method and intensity of sampling of the acceptable sets of "relevant" attributes;
- (vii) the quantitative or qualitative measures to be used in expressing the "states" of each attribute (involving an often arbitrary assignment of working commensurability)*;
- (viii) the measure of similarity (or distance) to be adopted.

Only some of these are susceptible to a more or less quantitative approach; in those cases statistical or other mathematical principles and techniques will sometimes assist in the judgment but subjective, though not always arbitrary or uninformed, decisions as to appropriateness or usefulness are needed all along the line. I shall mention later some useful bases for such judgments but, as an initial restriction, we may reasonably confine our attention to *attributes which show a degree of stability or regularity in the individual objects over the time-range in which we are interested*. Such a restriction must itself be arbitrary;

* Williams and Dale (1965) have recognized "the highly autocratic nature" of the convention that attributes are dimensionless ("dimension" is used here as in dimension theory of physics, not in a spatial sense) and that quantities representing different attributes are jointly available for arithmetical manipulation. "Standardization" to unit variance, or in some other way, may add some reasonableness to the commensurability assumption but does not remove the difficulty; the particular scaling thus derived depends in any case on the constitution of the chosen sample of objects.

furthermore, since it does not in itself introduce sufficient constraint to prevent the swamping of overall regularities by additional attributes, *there is still no parametric value of similarity*.

The concept of a "general" classification has been criticized also by Edwards and Cavalli-Sforza (1964), while some authors (e.g. Olson, 1964) have explicitly but unjustifiably assumed that there exists a finite set of "meaningful" characters.

To resume the enumeration of Gilmour and Walters' principles:

- (7) "General-purpose classifications can be made only when the objects concerned are influenced by a powerful factor, which causes a number of their attributes to be highly correlated in their occurrence; in the absence of such a factor, only special-purpose classifications can be made."

Comment: It will be best to avoid discussion of causality; we have perhaps already ventured too far into the philosophical chamber of horrors! In discussing the previous principle I stressed the importance of pervading regularity or pattern (terms which are free of the causative or the particular mathematical connotations of "factor"). We need to know not only something of the patterns but also the range of their pervasiveness and the extent of their relevance to our particular concerns. High correlation (or any other measure of covariation) of occurrence of attributes (i.e., of "states"), which Gilmour and Walters consider to be the *sine qua non* for a general classification, is meaningful only over a restricted domain, as we have seen, and is therefore dependent on our interest and purpose. If we attempt a "general" classification, for example, of the object set:

{Mao Tse-tung, a peanut, the Sphinx, an electron, a litre of alcohol, the star Achernar},

we shall find it very unsatisfying *because we cannot define the field of our interest*

or readily choose attribute sets showing regularities. Such sets *could* be found among the infinity of attributes, but would appear to us to be absurdly chosen.

The question is whether one can choose such a domain pervaded by a regularity which, firstly, is perceivable by us, and secondly, seems meaningful and useful to us. For a set of miscellaneous everyday objects there are various ways of selecting a set of *many* pluri-state attributes which would show a high correlation of their states. For instance, the selection could be such that a classification derived from a randomly chosen subset of the original attribute set would have a high probability of grouping the objects according to their geographical origin, which would then be the "factor" or regularity determining the high correlation. Our immediate response is, "But this is a special-purpose classification, and the attributes were specially selected." This is true, but when we select (and we *do* select) the acceptable attributes to be drawn upon in making a "general" taxonomic classification we also unconsciously choose a particular set of attributes specially reflecting a "factor" which is strongly linked to our (selected) interests.

In biological taxonomy, many regularities are discernible among the attributes which we customarily consider useful for classification, of interest to other branches of biology, or of practical importance. The reasons for these qualities and for the regularities are themselves closely though complexly related, and this is the nub of the matter. Evolutionary processes, selection and adaptation, population structure and dynamics, and the genetic mechanism itself in its grosser aspects as well as in its physico-chemical basis and its organization for the storage, replication, transfer, and implementation of instructive information—all these, intricately interwoven as they are, are closely associated with those perceivable (phenetic) characteristics of organisms which concern us. How this

bears on our original questions concerning optimal classification will be discussed after dealing with a few remaining points.

"(8) A distinction can be made between typological and definitional methods of making a classification. In the former, which is exemplified in the semantic development of the words used in everyday language, no one or more attributes are necessarily possessed by *all* the objects in a particular class, but rather, these objects show a 'family resemblance' to an imaginary 'type-representative' of the class. The definitional method, on the other hand, involves a conscious laying-down of certain attributes that an object *must* possess in order to belong to a particular class. Each method is appropriate to a particular type of purpose."

Comment: Provided that we define the attributes and attribute-sets, this is a valid methodological distinction. The use of "typological" here may be misleading, however, since it suggests the essentialistic concept of an archetype or *Bauplan* which has been effectively criticized, for example, by Simpson (1961) and Hull (1965). It will suffice to quote with approval Simpson's dicta that "Typological theory is linked with philosophical idealism [i.e. essentialism. L.J.] which on pragmatic grounds (if no other!) must be excluded from modern science" and that "such metaphysical beliefs . . . have no heuristic value". This is not to deny the usefulness of a sort of "working" Type concept derived as an ordering of our experience and used for the purpose of orientation of further experience (Carolin, 1967). Such a concept must be subject to any necessary revision in the light of new data or of reconsideration, and need have neither essentialistic nor phylogenetic implications. (This has no present connection whatever with the system of nomenclatural "types" which are now merely formal reference points for the names of taxa.) The first of

the two methods distinguished by Gilmour and Walters may be redefined to exclude any notion of a "type-representative" and then becomes in effect the *polythetic* method (Sokal and Sneath, 1963; Williams and Dale, 1965) used either intuitively or numerically in most biological systematics today. The second is similar to the *monothetic* method employed by some "old-fashioned" systematists in placing their taxa, and used generally in keys and some other post-classification procedures for discrimination and assignment, as well as in ecological, non-biological and *ad hoc* classifications. Monothetic classifications are by nature arbitrary or suitable only for special purposes, and will not henceforth concern us.

Apart from these "Principles", several other statements of Gilmour and Walters are pertinent as expressions of the phenetic school's viewpoint. They will serve as convenient pegs (not, I hope, as a gallows!) upon which to hang some observations. Gilmour and Walters refer disapprovingly to "the impression that taxonomic work has an aim of its own, apart from the aims of biological science as a whole, and that, if this aim could be fully accomplished, it would result in a single, perfect, ideal classification of living things", and proceed to uphold "the view that neither affinity nor phylogenetic relationship can be regarded as valid aims for taxonomy and that they have been adopted because of a lack of appreciation by biologists of one of the basic principles of classification enunciated by philosophers—namely the principle that it should serve some extraneous purpose."

We may grant that the quest for a single, perfect classification is beset with difficulties, some of which are inherently insoluble. But what do the words "should" and "extraneous" mean in this passage? We have seen that classifications can only be judged in relation to some set of requirements (not necessarily an *aim* of the maker of the classification) but the expression of phylogeny is as valid as any

other requirement, even if it cannot be perfectly fulfilled. Biological classifications are useful precisely because of circumstances inextricable from those involved in evolution in general and the phylogeny of the groups concerned in particular. Hence one cannot speak of "extraneous" purpose as if this must exclude phylogenetic expression.

The authors proceed: "This lack of appreciation has left, so to speak, a vacuum which was filled, in pre-Darwinian times by the semi-theological concept of affinity, and later by the *equally vague* [My emphasis. L.J.] concept of phylogenetic relationship."

I am at a loss to understand how even a philosopher, much less two biologists, could make this astonishing statement! Phylogenetic (cladistic) relationship, though not metrically expressible in any unequivocal way, is (potentially) topologically representable with little ambiguity, at least above the level of the coenospecies. We may not *know* the details of phylogeny but (unless we reject biological evolution) we must accept that they *exist uniquely* in space and time, and therefore form a concrete basis for concepts of phylogenetic relationship, however defined. In contrast, the notion of "affinity" is subject to unlimited variation and any claim for a firm basis for it must be metaphysical. The province of metaphysics appears to be to pose and seek to answer questions which are either practically or inherently unanswerable. If the former become practically answerable then that part of the subject moves into the realm of natural science. The latter have no place in science—or perhaps in any useful inquiry—except as a warning.

Gilmour and Walters then suggest that the purpose of taxonomy should be to make a "broad map of the diversity of living things which, by taking account of as wide a range of attributes as possible, will serve the needs of as many as possible of those concerned with animals and plants"; they regard phylogeny as the "factor" making

such a general-purpose classification possible.

This procedure provides no definable basis for agreement or for testing whether a classification should be changed or not, except that these authors would "regard stability of nomenclature as a very important factor in deciding whether or not to alter the rank of a taxon". Now, under the present system formal changes merely of *rank* (as distinct from transfer from one group to another) are not called for by changes of *position* on the topological map of inferred phylogeny (some palaeontological cases apart) and in fact usually arise from re-evaluation of the importance of phenetic differences, so that it is hard to see how the phenetic philosophy itself affects this situation! From their next remark that "where new knowledge renders an existing classification clearly absurd, changes must be made, but we suggest that such cases should be comparatively few", it is clear that, just like phyleticists, Gilmour and Walters cannot stomach classifications which are manifestly inconsistent with the inferred phylogeny or at least with the inferred genetic constitution. In a later paper, Walters (1965) expresses an even more extreme pragmatism; one may share his impatience with nomenclatural formalities and quibbles, but these are not a product of phyleticism.

The virtue of arbitrarily imposed stability of any unsatisfactory system is not apparent, though frivolous changes should be avoided. Systematics as a humble servant supplying changeless tags for gardeners, agriculturists, foresters, or physiologists would contribute little of interest to biology. Caution, criticism, and common sense play useful roles in science but conservatism for its own sake is inimical to it. If taxonomy were to have only the uninspiring goals set up for it by Gilmour and Walters it might as well be abandoned to any rude mechanicals who cared to gather voluminous data for a computer. Understanding of nature would not be its purpose

and any value it might have for biological theory would be incidental.

THE VICIOUS CIRCLE

Another philosophical defence of phenetics rests on the assertion that to allow phylogenetic conclusions (speculations, hypotheses) to influence taxonomic conclusions involves us in circular reasoning and is therefore inadmissible. This view is forcefully expressed by Sokal and Sneath (1963) and accepted by Williams and Dale (1965). Many practising systematists (e.g., Carolin, 1967) now defend it in theory, but nevertheless appear to be influenced by phylogenetic thinking in their actual work.

It is true that there are some recursive loops in the phyletic taxonomist's reasoning. This applies equally to morphological interpretation and description, in particular to the problem of homology. Attempts to develop "logical", non-phylogenetic definitions of homology (e.g., Mason, 1957; Carolin, 1967) are doomed to failure, being open to the same criticisms as phenetic taxonomy.* Indeed homology needs no separate treatment since comparative morphology implies taxonomy and *vice versa*. The need for very careful interpretation of morphological homology on a basis of evolutionary likelihood is illustrated by the

* Key (1967) attempts to define *operational homology* thus: "Feature a_i of organism A is said to be *homologous* with feature b_i of organism B if comparison of a_i and b_i with each other, rather than with any third feature, is a necessary condition for minimising the overall difference between A and B". Despite its spuriously precise formulation, such a definition is itself thoroughly non-operational except in finite cases, for the reasons already given. When discussing set-correspondences, I have used the term "working homologies" simply for such correspondences of features (of objects of any kind) as we may take, possibly on very complex and subtle grounds, to be a reasonable basis for further steps in the comparative procedure. The "operational homology" concept discussed at length by Sokal and Sneath (1963) boils down to much the same thing; it is only vaguely operational, in the sense of having some empirical consequences which can be checked (Hull, 1967).

highly modified and cryptic structures of many animals and plants, for instance the "bulrush" or "cat-tail" genus *Typha* (Briggs and Johnson, 1968) as considered in relation to other groups of monocotyledons. No phenetic comparison could be of much value here unless the relevant homologies were for the most part correctly worked out beforehand. The comparative survey by Hamann (1961) of the families of the "Farinosae", though most comprehensive and thoughtful, includes a numerical phenetic study of which we cannot confidently accept the results, because of serious doubt as to the homology of many of the features compared. Here, as often in broad surveys, the data have been drawn partly from various published descriptions, which may be very misleading. Throckmorton (1965) points out that phenetic parallelism may be equivalent to genotypic homology, that is, similar conditions which appear to arise independently may be due to separate "assembling" of elements of a common pre-existing genotype.

We have seen that there is a hidden circularity in any definition of a "general classification" and a hidden infinite regress in every aspect of phenetics. The Infinite Regress and the Vicious Circle or ultimate tautology are the two most inescapable hard facts of philosophy, mathematics, and science—at bottom we must come upon one or the other. They have been encountered by every critical child who has asked, "If God [or whatever agent or process one wishes to substitute] made the Universe, then who made God?" In any case, what is before or after, backwards or forwards, on a cosmic (or ultra-cosmic) scale? Such ancient questions are ultimately unanswerable, though they may be pushed back somewhat along the regress, but they may help us to dispense with some unnecessary non-answers to non-questions.

The partial circularity inherent in the phylogenetic approach is not reducible to a *petitio principii* as is the simple vicious circle of logic. It is not just saying tautolog-

ically that conclusion B follows from premise A when we have already used B in formulating A (as, for instance, "a classification produced by a good taxonomist is good", having defined a good taxonomist as "one who always produces good classifications"). Nor is it arguing in an epistemological circle (Hull, 1967), that is, purporting to show that B is true because of A when we can only know that A is true if we know that B is true (as, for instance, "a newly discovered mammal from Africa is a placental mammal because all African mammals are placentals").

Rather, we are accepting a degree of positive feedback in what is not a chain but an anastomosing plexus of reasoning and evidence. The premises and reasoning used in arriving at a representation of a (hopefully) deduced phylogeny are highly complex, as are the conclusions (see, for example, the attempted reconstructions of the phylogeny of Dipsacaceae by Ehren-dorfer, 1964a, 1964b, and of Proteaceae by Johnson and Briggs, 1963; also Thorne, 1963). There is no hope of making them fully explicit, but to deny them validity on that ground would be to deny validity to all but the simplest reasoned conclusions in science or life. Many pheneticists take the view that, in the absence of a "time-machine" (Carolin, 1967), palaeontological data provide the only information of direct phylogenetic significance that can be validly fed into a taxonomic system. But most of us do, after all, believe that the processes of evolution are pretty well understood. Most phylogenetically inclined taxonomists and general evolutionists consider that the existing corpus of evidence justifies the use of evolutionary principles to make sense of the diversity of Recent as well as fossil organisms. This involves a great deal of argument by analogy and extrapolation, but so does most of science. How many of us *know* from direct evidence that we consist of neutrons, protons, and electrons, or even that we have chromosomes? But we believe it because we have convincing

evidence of consistency in the physical world. In the course of describing and analysing the phylogenist's, and in fact the common scientific, method of successive approximation, Hull (1965, 1967) presents powerful philosophical arguments for a similar rejection of the charge that phyletics is based on viciously circular reasoning, in either the logical or epistemological sense.

It is a commonplace that, outside mathematics, most scientific hypotheses and theories are inductively derived, though we aim to reduce the number of such hypotheses to a minimum and to use them as postulates for a deductive superstructure. They are testable by predictive value, self-consistency, and consistency with extraneous evidence. Operationism requires a hypothesis to be potentially falsifiable, that is, subject to tests which *could show* it to be false. If we exclude phylogenetic interpretation, phenetic "general" classifications of organisms are scarcely hypotheses at all*; they are subject merely to practical evaluation by their predictivity within the field of our interest. Phyletic classifications (which of course use phenetic techniques as *part* of their method of derivation) may certainly be equally effectively judged on predictivity; they also embody biological hypotheses (though the classification as such may not fully *express* them), which are subject to many checks of self-consistency and extrinsic consistency—though such checks are not necessarily easily expressed in quantitative terms. It has been claimed (e.g., by Birch and Ehrlich, 1967*a*) that these embodied hypotheses, and indeed synthetic evolutionary theory generally, are not falsifiable. This is due to the unsophisticated assumption that a

useful hypothesis must be a sufficiently simple and definable proposition to be falsifiable in the same way as the statements: "all swans are white" (known to be false) and "the equation $x^n + y^n = z^n$ has no solutions where x, y, z , and n are integers and $n > 2$ " (the celebrated "Last Theorem" of Fermat, which is not a deductively proved theorem at all but has not been shown to be false). Phylogenetic hypotheses are susceptible of demonstration that they are very likely to be untrue, and this is indeed the only disprovability we can hope for in a good deal of science. We can achieve a considerable and reasonably convincing understanding of nature in such ways, and can apply a great deal of criticism to our evidence, arguments, and conclusions. Mackerras (1964), in a cogent defence of the phylogenetic method, says that "to make a preliminary arrangement on general resemblance (as is often done), test its components for phylogenetic concordance or discordance, and then base a classification on the results of those tests, is not circular reasoning in any sense of the words."

Birch and Ehrlich (1967*b*) say, "We *do not* need to consider the unknown history of the organisms . . . in order to classify them . . . Phylogenetic history . . . is not pertinent to most of the uses of classification. Astronomers do not have to know the histories of sidereal bodies, nor physicists of atomic particles in order to do valid scientific work with them." These authors claim precision for phenetic methodology and are opposed to "mixtures" of phenetics and phyletics in classificatory procedures. An outrageous stab on their part is to accuse the evolutionist Ernst Mayr of Platonism; to interpret Mayr's insistence on the "reality" (i.e., reasonable definability) and importance of the species in evolutionary theory as a form of Platonic essentialism is a gross distortion indeed. Birch and Ehrlich's quoted statement is partly true; we certainly *can* classify without reference to phylogeny (Linnaeus did

*This is admitted by Williams and Dale (1965), though we may accept their suggestion that non-probabilistic numerical classifications in other fields have value as hypothesis-generating systems. While they allow that phylogenetic hypotheses may be generated, these authors tend to dismiss them on the ground of their alleged untestability.

it!) and, for the reasons already given, such classifications may work quite well. This is in itself *no* reason at all to reject consideration of phylogeny, which is one important aspect of the evolutionary process that underlies the workability of our classifications. Astronomers *do*, of course, consider the histories of stars, etc., and may modify their classifications accordingly. The case of [sub]atomic particles is simpler and not really comparable; nevertheless elucidation of underlying regularities and the possible transformations of particles is highly pertinent to our mental systematization of them. Improved classification in these cases does in fact involve positive feedback (i.e., a "mixed" procedure).

For the reconstruction of phylogenies we can use evidence of many kinds which is not just derived from the patterns of phenetic diversity we observe in the organisms under study. Cytogenetic data, for example, have a special relevance beyond mere matching of attributes of the chromosomes or of the genetic systems. The contribution of comparative biochemistry, especially where its adaptational significance is evident or where biosynthetic pathways indicate probable directions of change (e.g., Scora, 1967), frequently goes beyond mere phenetic comparison. Quite often there is a fossil record of considerable value. In particular, probable directions in the evolution of organ-systems can be deduced from ontogenetic, ecological, adaptational, and historical evidence and comparisons. Examples are legion and can be found in works on evolutionary theory and in many publications presenting critically argued reconstructions, as in the case of the Proteaceae already cited. A single example will suffice: floral adaptations for pollination by long-tongued insects or by birds cannot have occurred in geological periods before such animals existed. Any phenetic classification which grouped organisms in a manner inconsistent with such a fact (taking all other relevant information into account) would not be

acceptable as even roughly consistent with phylogeny, and there is no reason why we should be asked to accept it simply because of claims of repeatability, objectivity, precision, or stability. Repeatability is not dependent on the phenetic approach and the claimed objectivity and precision are superficial. Stability has been mentioned already and later will be discussed further.

In summary, such philosophical considerations as we have reviewed, far from invalidating the phyletic approach to taxonomy, may be turned against the pheneticists themselves.

TAXIMETRICS*. I. NUMERICAL PHENETICS

Evaluation and methods

We have discussed at length the foundations of phenetics. Only brief mention can be made of the techniques of numerical taxonomy and their applicability. These are being used, often on rather trivial problems, by increasing numbers of taxonomists who often seem to accept very uncritically the philosophical foundations and the mathematical models which they explicitly or implicitly embody. The worth of the results seems usually to be measured against what the taxonomist has done, or would have done, without the techniques—an amusing test, to say the least. Mathematical methods of ordination and classification seem to have thrown very little new

* The forms "taxometrics" and "taxonometrics" are also in current use. "Taximetrics" is etymologically the best-formed since the first root is the Greek *τάξις*, *τάξεως* (or Ionic *τάξις*)—"an arranging", of which the combining form in Greek compounds was *ταξι-* (Liddell and Scott, 1864). The word "taxonomy" is badly formed (cf. the French "taxinomie") but is now firmly established by usage. One may hope that "taximetrics", as used by Rogers (1963), will prevail, or perhaps it is not yet too late to substitute the simpler "taximetry", with the ending on the model of "geometry", "trigonometry", "biometry", "anthropometry", etc. Further development of the first of these examples could lead to the delightful consequence of numerical taxonomists' being called "taximeters".

light on the taxonomic problems of "difficult" groups, despite their usefulness in ecology or in non-biological fields where there is no obviously prime source of regularity comparable with phylogeny and the genetic constitution of organisms.

It is true that Watson *et al.* (1966, 1967) have produced a slightly better classification of the Epacridaceae and a considerably better one of the Ericaceae than those previously existing (not necessarily "accepted" as they put it)—but the latter were almost a century old and based on less complete data and, in the case of Drude's less satisfactory classification, on an analytic or monothetic approach. From the data, it seems to me that many good modern taxonomists should have been able to effect equal improvement by adequate study and phylogenetic consideration. Indeed, by careful study and reasoning, Watson had already cleared up most of the problems, though the computer analysis suggested some placings for doubtful genera. In the more difficult case of the Basidiomycetes, reported by Kendrick and Weresub (1966), "Adansonian" computer analyses gave thoroughly unsatisfactory results as compared with a reasoned phylogenetic approach. It would be difficult to explain these away by criticism of the mathematical models or techniques employed; the source of trouble seems to have been evolutionary convergence. In a group of *Solanum* species and hybrids, Heiser *et al.* (1965) found that the results of a phenetic numerical analysis were less satisfactory than those of a "subjective" study when checked with the considerable biosystematic information.

A recent *coup d'œil* over the grass family, using a probabilistic mixed-data method (Clifford and Goodall, 1967), seems only to confirm the well-established features of modern non-numerical classifications and to indicate that the "difficult" cases actually are so. Its sampling basis of one species per tribe would certainly satisfy neither traditional taxonomists nor

phyleticists. The authors claim it as an "advantage" that "such a limited sampling eliminates all the intra-taxon variability", which hardly engenders confidence in their appreciation of the significance of nature's complexity. In their study of difficult species-complexes in *Cassia*, Irwin and Rogers (1967), using the graph-theory model of Wirth *et al.* (1966), seem to be reasonably well satisfied with the results. The taxa are too closely related to show clear phylogenetic patterns, but little biosystematic information is available. The taximetric clustering procedure seems at least to have promoted the authors' confidence.

A notable exponent of numerical methods recently remarked to me that, given suitable data and an appropriate program, "the computer can produce a better classification than a poor taxonomist". But then, so can a good taxonomist—and, being in possession of much extraneous information and reasoning power which are not in the storage unit or the programmed strategy of the computer, he can often do so with much less tedious recording of data. Might it not be more economical to employ a few more good taxonomists? This ought to be possible out of a world population of 3×10^9 .

The literature provides ample illustrations of considerably divergent classifications of the same material, *from the same data*, produced by various strategies of numerical taxonomy (Boyce, 1964; Olson, 1964; Katz and Torres, 1965; Minkoff, 1965; Rohlf and Sokal, 1965; Sheals, 1965; Kendrick and Weresub, 1966; Lance and Williams, 1966a; Sokal and Michener, 1967; 't Mannetje, 1967). Varying the attribute sampling gives somewhat different results again; so does changed scoring, whether it implies altered attribute definitions or grouping, changes in measure, or all of these. It is true that different methods sometimes yield reasonably concordant results for the major divisions of a group but in these circumstances the taxonomist

is seldom in doubt in any case. If, after a careful investigation, a taxonomist remains in doubt as to the classification of particular groups, it is usually due to intrinsic complexities, past or present, in evolutionary situations. Lack of well-defined clustering and especially of nested groupings is inherent in such situations; consequently the different characteristics of the various taximetric strategies will result in lack of consistency. These instabilities arise from the fundamental difficulties discussed earlier; they cannot be regarded simply as functions of a signal : noise ratio (Lange *et al.*, 1965) (though this may vary with the method), since the signal itself is to some extent what we define it to be.

The theoreticians of numerical taxonomy have enjoyed themselves immensely over the past decade (though not without developing several schools with scant respect for each other!). The mushrooming literature is quite fascinating and new developments tumble after each other. Anyone who is prepared to learn quite a deal of matrix algebra, some classical mathematical statistics, some advanced geometry, a little set theory, perhaps a little information theory and graph theory, and some computer technique, and who has access to a good computer and enjoys mathematics (as he must if he gets this far!) will probably find the development of new taximetric methods much more rewarding, more up-to-date, more "general", and hence more prestigious than merely classifying plants or animals or working out their phylogenies. Unlike the taxonomic questions themselves, the methodological questions which the taximetric mathematician sets himself are mostly so definite and so answerable—yet hunting for the answers is difficult enough to be interesting:

Does a certain distance function define a metric, semi-metric, quasi-metric, or non-metric space? Is it monotonic (varying throughout in the same direction) as object-groups are successively fused? What are the properties of the Canberra metric as a

variant of the Manhattan metric? What are the relative advantages of divisive and agglomerative strategies? (One can do a little classifying of the strategies themselves—without using the computer, so that there is some intellectual satisfaction in it.) What are the properties and advantages of hierarchical (nested) systems as opposed to overlapping, clustering (clumping) systems? To what extent are particular agglomerative sorting strategies space-distorting? Can dimension-reducing ordination techniques such as principal component analysis and factor analysis be used efficiently as a basis for extracting a classification from an ordination? What are the virtues of rotation of axes in a factor analysis to yield "simple structure"? What do the "factors" of factor analysis mean? (Not so easy to answer.) Are certain matrices always positive semi-definite (symmetric with all eigenvalues non-negative)? Are certain functions algebraically tractable? Do annoying singularities occur? Are probabilistic models really applicable or desirable? Can predictivity be usefully measured? What are the virtues of combining R- and Q-analysis, and what are the appropriate techniques? (As now agreed both by Sokal and Sneath (1963) and Williams and Dale (1965), these terms refer to matrices: the elements of a Q-matrix are measures of comparison of objects (individuals) while those of an R-matrix are measures of association of attributes.) How effective are heuristic "hill-climbing" strategies (Rubin, 1967) which search for optimal structure by systematic trial? How large a data matrix can particular computers handle for particular programs and how much computer time does an analysis take? Are programs easily modifiable? What methods can handle mixed data, and how validly? Should attribute scores be standardized to unit variance, or in some other way? Are double-negative matches to be included? What is negative in any case? (Suppression of a character may well

be the derived condition within our frame of reference; to assign greater importance to “presence” than “absence” is a decidedly subjective judgment.) Is the information-statistic the best basis for agglomerative strategies?

Some of these questions are discussed in an enlightening paper by Williams and Dale (1965). Although these authors make the usual, partially invalid, pheneticist assumptions, their mathematical discussion is most helpful and pertinent. However, reference to various subsequent papers is necessary (e.g., Bonner, 1964, 1965; Macnaughton-Smith, 1965; Hall, 1965, 1967*a*, 1967*b*; Davidson and Dunn, 1966; Goodall, 1966; Gower, 1966; Jancey, 1966; Menitskiĭ, 1966; Wirth *et al.*, 1966; Lance and Williams, 1966*a*, 1966*b*, 1967*a*, 1967*b*, 1967*c*; Davidson, 1967; Estabrook, 1967; Orloci, 1967; Rubin, 1967; Crovello, 1968; Wallace and Boulton, 1968), and indeed any attempt to be comprehensive in this field is obsolescent before it reaches the printer. The well-known text of Sokal and Sneath (1963), which was welcomed by non-initiates as the Bible of the subject, though still important, has already taken on the archaic flavour of an Old Testament, both as to foundations and as to the range and evaluation of mathematical methods and models.

Are phenetic numerical methods, then, of value in practical systematics? I think they can be, especially now that computers can process high-order matrices (approaching 200 “objects”, or even more) and that mixed-data programs have been developed to deal with two-state, multi-state, and infinite-state (continuously-varying) quantities in the same matrix (Lance and Williams, 1967*c*; Wallace and Boulton, 1968). Our ordinary intuitive processes begin to lose efficiency with problems of this magnitude, especially when a comparable number of attributes is used. The computer’s elucidation of “structure” in the data may be useful to us in *suggesting* an appropriate classification, even though

we are aware of the numerous subjective and arbitrary decisions implied in the choice of data and of mathematical models.

I do not believe, however, that we should accept any such classification as the last word or as indicating that answers exist to the chain of questions posed early in this address. Ordinarily one can attempt to evaluate such a classification by evolutionary considerations and can modify it accordingly if necessary—there is no need to stand in awe of its “objectivity”. If this is not feasible, one can accept it as possibly better than one could do by intuitive methods. *Having accepted a finite set of relevant attributes* (not necessarily all used in the analysis), one may judge classifications as to *predictivity within that set* by means of the probabilistic utility function of Goodall, who, in an interesting paper (Goodall, 1966), outlines some of the constraints which are necessary before any probabilistic technique can be validly used.

Warburton (1967) suggests that the purpose of classification should be to maximize “the probability that statements known to be true of two organisms are true of all members of the smallest taxon to which they both belong. It should not be impossible to develop tests for this property to objectively decide which of several rival classifications is best.” Any such test, of course, would depend on definition of a finite set of “statements”. If we agree to that, Warburton’s criterion could provide a useful pragmatic test within a frame of reference, but because of this arbitrary aspect it answers no general question about *the best classification*.

No good reason exists why any particular horizontal cuts across phenograms* should be accepted as meaningful “phenon levels” (Sneath and Sokal, 1962; Sokal and Rohlf,

* Phenograms are phenetic dendrograms purporting to show similarity by a tree-like diagram of nested subsets of the objects—not merely topologically but scaled according to the particular measure used.

1962; Sokal and Sneath, 1963) upon which to erect our formal taxonomy. All claims for phenetic "standards" of rank collapse on analysis, including those reiterated in the rather brash paper in which Sokal and Sneath (1966) set out their recipes for a great leap forward to "efficiency in taxonomy".

Although a number of earlier authors had proposed numerical methods for taxonomic classification, it is interesting that one of the influential streams in the modern period is associated with the work of P. H. A. Sneath (for references see Sokal and Sneath, 1963) on the classification of bacteria, a group in which phylogenetic interpretation has been unsatisfactory and equivocal. Another vigorous stream, associated with W. T. Williams and his collaborators, began in the field of ecology and has close associations with workers in such areas as criminology (Macnaughton-Smith, 1965), industry, and business, in all of which classification has been and must surely remain an *ad hoc* arbitrary matter (however "objective" we may cause it to appear by making the subjective decisions *before* beginning the calculations).

Generally experienced systematists, young or old, who employ taximetric methods usually seem to make excuses for them: "Well, the computer analysis didn't do any better than (or as well as) I could, but it wasn't a bad job and perhaps it gave me an odd idea here and there" would sum these up. Despite substantial response arising from genuine interest as well as from bandwaggoning, the showing after ten years' hard selling is not at all impressive so far as improved practical classification or biological understanding are concerned. The onus must surely be increasingly on the numerical pheneticists to give reasons, other than its meretricious glitter, why we should buy their product except with considerable reserve. Fashion, spurious objectivity, and competition for financial grants (Rollins, 1965) are not very scientific reasons. Neither is fear of being included by

such determinedly iconoclastic zealots as Paul Ehrlich (1965)[†] in the class: {"members of the old school" who "would still like to see a pinch of phylogenetic speculation mixed into their basic data (presumably for sentimental reasons)" and who "will continue to promote this confusion for some time to come"}—and thus excluded from the class: {"those who wish to look forward"}, regarded by Ehrlich (objectively, no doubt) as non-overlapping with the former class. There is a familiar ring to this: "He who disagrees with me is a reactionary."

Our only certain scientific pay-off from phenetic taximetric methods is that, having gathered a lot of data, we shall therefore be less likely to overlook features of evolutionary or practical significance. Often the game may not be worth the candle.

Relation to genetic basis

Sneath and Sokal (1962) introduced the concept of the "matches asymptote", hypothesizing (Sokal and Sneath, 1963) that "the similarity between two operational taxonomic units is some parametric proportion of character matches which we are estimating with a sample of characters" and that "as the number of characters sampled increases, the value of the similarity coefficient becomes more stable". The meaninglessness of a parametric value in the all-attributes case has been demonstrated but it may still seem reasonable to consider that, given a suitable measure of similarity, such a parameter exists for the "matches in the nucleotide sequence of the DNA of the genotype". If that is so, then, over the attributes ordinarily regarded as important, the confidence band for an estimate of this parameter will narrow as the sample size increases. In that there is a certain regularity, related to the genetic

[†] More recently (Ehrlich and Ehrlich, 1967), this author has realised that numerical phenetics is inherently indeterminate in its results—and appears to have adopted a nihilistic attitude to taxonomic improvement in general.

information, over the attributes we are likely to consider, some such convergence is indeed to be expected. The existence of a numerically definite value will of course depend on the acceptance of a *particular* finite attribute set defined by enumeration of its elements, but it is doubtless correct to say, in an imprecise way, that the larger the sample of attributes the better our comparisons will tend to be in reflecting the similarities of the genotypes. However, even if we knew the entire nucleotide sequences over a set of organisms we should still have to make many decisions on matching procedure. (We can hardly speak of comparing "genes"—the gene is no longer a useful operational unit at the level of molecular genetics.) We certainly could not set up pluri-unique correspondences of DNA base sequences over the whole genotype of organisms which were at all diverse or differed in karyotype (see Ehrlich, 1964, and Reynolds, 1965, for possible further difficulties). Taking the step up to the comparison of proteins will not remove this difficulty, though such comparisons as can be made in this field will provide information at least as useful as that from other attributes.

Related to, but not identical with, the "matches-asymptote" hypothesis is the "non-specificity hypothesis" (Sneath and Sokal, 1962; Sokal and Sneath, 1963) which assumes that "there are no distinct large *classes* of genes affecting exclusively one class of characters such as morphological, physiological or ethological, or affecting special regions of the organism". This suffers from the same fundamental difficulties of definition and testing as other phenetic concepts but, so far as the phenetic tests may be accepted as valid, it appears often not to hold very well (e.g., Rohlf, 1965; Thornton and Wong, 1967). Indeed, the lack of correspondence of groupings derived from different sets of data (say, from internal anatomy and external morphology, or from larval and adult stages) is a problem in phyletic as well as in phe-

netic taxonomy, but at least the phyletic approach to its resolution should be more intelligent and subtle than crude lumping of the data or results. An alternative, of course, is to take the view of Ehrlich (1964, 1965) that "phylogenetic speculation is fun, but seems to have little scientific purpose" and hence to regard all classifications as special, so that "one wishing to make predictions about the distribution and ecology of larval mosquitoes would presumably do better to work with a taxonomy based only on characters of the larvae".

Phylogeny cannot be perfectly elucidated, and it is inherently complex and reticulate at those levels of grouping at which interbreeding among groups still occurs. Nevertheless its unique existence provides the nearest thing to a solid base for general taxonomy.

TAXIMETRICS. II. NUMERICAL PHYLETICS

Conditioned as they have become by reiterated assertions to believe that operationist philosophy demands that taxonomy must be purely phenetic, only a few of the theoreticians of taximetrics have shown interest in developing mathematical models and corresponding numerical techniques for the elucidation of phylogeny. One can set up plausible, though clearly over-simplified, models of phylogenetic processes, just as one can for genetic systems. Some at least of these models are reasonably tractable mathematically, as are the rather idealized models of population genetics which have often proved fruitful in suggesting hypotheses and in testing for consistency.

The "advancement index" of Sporne (1948, 1954, 1956, 1960) was an early attempt to assess evolutionary advancement. Sporne's approach was different from later methods discussed here, but is of limited applicability, for the reasons given by Davis and Heywood (1963, p. 39). The concept of a statistically measured overall "advancement" is both too indefinitely based to have much statistical validity and

too generalized for effective reconstruction of phylogenies.

A simple phylogenetic consistency test, based on the postulate of unidirectional change in individual characters, was developed by Wilson (1965). This field of study has been termed *cladistics* by Camin and Sokal (1965) who, having formulated a set of assumptions regarding evolutionary sequences, developed a computer strategy based on the principle of parsimony. Its aim was to construct *cladograms* (phyletic dendrograms) representing an evolutionary minimum-path branching pattern for the set of OTU's under consideration. The method was checked against the palaeontologically well-documented phylogeny of horses and also for a group of imaginary organisms ("Caminalcules") which had been independently generated using the phylogenetic principles enunciated. This important pioneering work has naturally appealed to those phyleticists who do not wish to turn their backs altogether on numerical methods, but has been virtually ignored by many pheneticists. Camin and Sokal themselves continue to prefer a phenetic basis for classification "until an operational system combining cladistics and phenetics can be established"; since any such system *must* be arbitrary, from the nature of the phenetic component, I cannot see how this aim can be achieved with the "objectivity" which Camin and Sokal desire.

Cavalli-Sforza and Edwards (1967) have discussed, in considerably greater depth and detail, some mathematical models and estimation procedures for phylogenetic analysis. Their intention, only partially fulfilled, was to use maximum-likelihood methods to estimate the form and proportions of "the most probable tree uniting the presently living populations". Their "branching random walk" model assumes that evolutionary changes leading to divergences in genetic constitution are, at the level concerned, essentially stochastic, summarized as "random genetic drift and vari-

able selection". One could disagree with these assumptions, and it is clear from the authors' very honest discussion that many arbitrary decisions have to be made as to the nature of the spaces and metrics implied by the models and by the methods of estimation. Tests of the method in cases where there is good fossil and other evidence for a particular phylogeny should be interesting. Cavalli-Sforza and Edwards dismiss Camin and Sokal's paper with the observation that the latter authors' assumption "that evolution *proceeds* according to some minimum principle" cannot justify the use of a "method of minimum evolution", though Cavalli-Sforza and Edwards do not deny the possible usefulness of such methods.

Working with amino acid sequences of cytochrome *c* from an assortment of vertebrates, two insects, and three fungi, Fitch and Margoliash (1967) have used a numerical method of phylogenetic analysis on the basis of "mutation distance", that is, the minimal number of nucleotides which must be altered in order to convert the coding for one cytochrome to that for another. This is an excellent paper, but I cannot resist quoting the metaphorical gem that if "one wishes to test a tree which differs only in the order in which the chicken, duck, and penguin are joined, the only legs in need of recalculation are those five descending to these birds from the avian apex". The mental picture which this conjures up is slightly less surrealistic if we note that Fitch and Margoliash's trees grow upside-down and have legs instead of branches!

Silvestri and Hill (1964) stress the value of the *patristic* approach (i.e., genetic comparison) in microbiological systematics, since there are few reliable data for cladistic hypotheses in that field. However, Silvestri (1964) also points out that phenetic differentiation may be outstripped by DNA evolution because of the degeneracy of the genetic code.

Further development of numerical phylogenetics is certain and seems to have considerable promise, but inasmuch as the methods *are* numerical they impose their inbuilt metric properties on a situation for which topological "reality" is more or less inherent, but "distance", if defined on essentially phenetic grounds, is dependent on our point of view. "Mutation distance" will usually be determinable only in respect of a few loci, and therefore inadequately sampled, but it does perhaps represent the nearest thing to a "natural" *metric* basis for biological taxonomy.

QUALITY AND QUANTITY IN MATHEMATICS

The urge to quantify is upon all biology and the social sciences. Undoubtedly it has already brought considerable benefits, but we should not forget that science is *not* the process of measuring natural phenomena; that is merely a technique of science. Very largely the aim of science is to discern qualitative differences, that is, to simplify and reduce the number of quantities which it is necessary to specify. For instance, in particle physics today the algebraic theory of *groups* (this term has here no connection with "groups", i.e., classes or sets, in classification) has been successfully applied to the ordering and prediction of quantum phenomena, as indeed it was applied earlier to symmetry relations in such fields as crystallography. The essence of the algebraic concept of a group is not quantity but *structure*: a particular set of relations between elements, the latter usually being *operations* in the particular class of relations which satisfy the group concept. This principle is expressed with greater generality in the concept of *isomorphism: the matching of sets of relations, not quantities*. It is these qualitative aspects with which much of modern algebra is concerned (e.g., Maxwell, 1965; Hollingsworth, 1967). The same is true of much advanced geometry, especially the non-metric geometry of position (mostly abstract and expressed in a nota-

tion allied to that of set theory): topology, or *analysis situs* as it was once called.

We should not be bewitched by number, in particular by the continuum (for basic concepts and some enlightening philosophical discussion, see Dantzig, 1962). We shall probably never be able to express or apprehend complex situations in their precise quantitative detail. Over the centuries, beginning with the natural numbers 1, 2, 3,, we have had to extend our concept of quantity to embrace zero, directed (positive and negative) numbers, rational fractions, algebraic and transcendental irrationals, complex numbers, and various kinds of hypercomplex numbers, such as quaternions, vectors, and matrices. To cope with *change* and with infinite processes, the differential and integral calculus was necessary. Further, to deal precisely with matters as uncomplicated (as compared, say, with those of biology) as the specification of simple events and relations in space-time, mathematicians have combined vectors, matrices, and calculus into the formidable subject of tensor analysis—and this in cases where there are no difficulties of multiple incommensurability. *We have no hope of extending this kind of precise quantitative mathematics* to describe the biological situations encountered in taxonomy. Mathematics is hard, paradoxically, because its subject matter is simple—and the subject has thus been able to progress a long way.

Statistics, largely based on probability theory, has been developed to cope with circumstances, in physics as well as in less simple fields, where we are concerned with large numbers of events which display a degree of disorder, that is, lack of individual specifiability (except where individual observation is feasible). Statistics requires certain assumptions, as we have seen, which are often not justified in systems of incommensurables. May we not be well advised to avoid putting all our money on Quantity? Qualitative assessment (which, as we have seen, does *not* exclude

mathematical concepts) may be a surer winner in heavy going. To be sure, we can often extract qualitative structure, or qualitative generalizations, from quantitative data and, when we can, we should doubtless do so.

CLADISTIC SPACES

We can also very often perceive qualitative structure, and subject our interpretation to empirical tests, by qualitative means. Quantitative expression of a situation, or a quantitative check, is of no greater value unless the assumptions underlying it are justified. We now know that often they are not. In the words of Hull (1965), an author whom all the devotees of numerical objectivity (*and* all phyleticists!) should read, "talk of taxonomic space, like talk of gene pools, is strictly metaphorical. . . . Taxonomic space . . . is an amorphous continuum with no intrinsic metric; there is nothing about taxonomic space to indicate how long to make the unit of measurement or at what point to begin measuring once the unit is chosen". To this we may add that the dimensionality of taxonomic space, even in the sense of phyletic space, is arbitrary and the space is not isotropic (similar in properties in all directions).^{*} There is a time dimension which is metrizable for our purposes by the ordinary physical criteria. The remaining dimensions will depend on our model or mental construct of the space. Two main phyletic cases arise: (I) We can say that there are n dimensions (including time) where n is the number of recognized "ends" (individuals, species, OTU's or whatever our smallest unit of taxonomic division is taken to be). This will be a pure "cladon-space", and any cladistic tree within it can be mapped without loss of information into a two-

dimensional space to form a conventional dendrogram, *provided* that we are not concerned with any measure of distance or indeed with any ordering except on the time co-ordinate. In this case it is *only the order of branching* which counts. (II) We can admit $p + 1$ dimensions (including time) where p is the number of attributes taken into account. This class of phyletic-phenetic space, which is merely an attribute-space with a time dimension added, is the basis of Cavalli-Sforza and Edwards's primary model; it has some metric properties but need not be fully metric in the sense of Williams and Dale (1965). A tree in it cannot be mapped into a two-dimensional space without considerable loss of information. Two- or three-dimensional representations such as those advocated by Sporne (1956), among many others, though useful as illustrations, cannot hope to be very accurate in representing either phenetic or evolutionary distance, according to whatever measure we adopt.

Whether or not we agree on some arbitrary working commensurability over the dimensions representing attributes (or some such measure as Fitch and Margoliash's "mutation distance"), in both space classes (I) and (II) time is incommensurable with the other dimensions. We cannot legitimately combine time and the other forms of "distance" into any general distance measure. This is in fact the strength of the time-oriented standpoint; the time-sequence is potentially determinable and often practically inferable. Thus, the only unequivocal dendrograms are topological cladograms which are *qualitative* in nature, except that if we do have palaeontological evidence of actual times of branchings these cladograms become unequivocally metric as to the time axis only.

CLADISTICS AND CLASSIFICATION

Assuming that we have arrived by reasonable inference at a cladistic tree, does this then at last lead us to an optimal *classification*? It certainly does not lead us to

^{*} Edwards and Cavalli-Sforza (1964, Cavalli-Sforza and Edwards, 1967) specifically *define* an isotropic "evolutionary space-time" but this isotropicity is achieved by a transformation which, as they say, "is a reflection of the genetic assumptions which are being made, and these will be peculiar to each case".

an *ordination* of the objects at any time level (more strictly, in any hyperplane normal to the time axis), unless we accept a particular attribute-space as in case (II) just described. But we may consider the possibilities of effective classification without ordination.

Often, by reference to the time sequence of branching, we may refer the objects to a nested (hierarchical) system of sets. However, if we are working wholly or partly below the level of effective genetic isolation, the tree is a banyan, anastomosing in its lower levels at least. The sets will then intersect in a complex manner and in this case there is certainly *no* unique hierarchy. This is one reason why the "biological" species concepts are important in taxonomic theory and practice, however difficult to define they may be in some groups of organisms (see Simpson, 1961, for an elaboration of this matter, and Hull, 1964, 1965, for an illuminating philosophical discussion). In these reticulate cases, which are usually below and about the "species" level in the conventional system, but may include somewhat higher levels if allopolyploidy has occurred, *only subjective or arbitrary classification is possible—there is no one "right answer"*, though some classifications are certainly worse than others. Numerical methods cannot alter this, though they may aid in evaluation of data and obviate the possibility of grossly subjective (that is, "unreasonable") judgements.

For many years I have worked with the large genus (or group of genera) *Eucalyptus*, which presents a maze of anastomosing situations of this kind, though not including allopolyploidy. It has long been clear to me that no single "best" classification, phenetic or "biological", is possible in such genera. It would be a wasted effort to aim at one, and it is a fond hope indeed that some numerical analysis or other is going to "solve" the classification problem. Nevertheless, it is not wasted effort to attempt to understand the group or to describe as well as we can the situations which exist

within it, in terms of phenetic variation, ecological relations, breeding behaviour, and so forth. As Harlan Lewis (1957) has said, "it may be easier to determine why the group is difficult than to decide the most appropriate taxonomic disposition".

This is where the biosystematic way of thought has the advantage over the formalism of the pure pheneticist. Biosystematics, even when not highly experimental, is attempting to understand and describe an aspect of nature in dynamic terms. This need by no means always, or even usually, be pushed to the level of a detailed analysis of population dynamics; we can infer a great deal from many other cases where detailed studies have been made. We need to see the forest as well as the trees, *and* some of the trees as well as the forest, but not always at the same time or place.

Formal taxonomy at the specific level cannot always aim at perfection, though there are in fact many perfectly straightforward cases where no dispute at all is likely to arise as to the correct grouping—nature often does present us with comforting discreteness at a particular time level, hence the not inconsiderable success of traditional systematics. In the difficult cases, we can frequently improve the formal classification up to a certain point by removing the grosser inconsistencies between formal treatment and our biological understanding, although possibly replacing them with lesser inconsistencies elsewhere. When we reach the stage where there is no gain in consistency (this is itself not unequivocally measurable; see Hull, 1964), we may as well describe the situation, and stop. Whatever theorists may suppose, practising systematists are well aware that there are vast fields in which much more effective work can be done before reaching the point where *inherent* instability becomes limiting. Instabilities arising out of ignorance or sheer wrong-headedness are, or were, common enough, but we need not judge any subject by its incompetents. The improvement may sometimes be made with

the aid of numerical methods, but is often quite feasible without them. Whether it is always worthwhile, by any method, is another question. Ehrlich (1964, 1965) thinks not; he may be right, but this is a highly subjective judgment, like many pronouncements of the objectivists. Certainly some more consideration of priorities in taxonomy, as suggested by Ehrlich, should increase the efficiency of the subject as an aid to understanding. This is quite apart from economic priorities, which are already well recognized.

Returning to a level above coenospecies and above allopolyploid formation, where the cladograms are free of loops, the problems are again discussed by Simpson (1961) and by Hull (1964, 1965). They are at two levels: phylogenetic interpretation and formal treatment. We shall here assume that our interpretation is as good as we can get it—how, then, do we group, and how do we rank?

Since lineages are continuous in time (up to the point of extinction), classification *along* the time axis must be largely arbitrary, although the rate of evolutionary change is by no means constant (see, e.g., Simpson, 1953) and some authors (e.g., Takhtajan, 1953) have described evolutionary change as chain-like rather than line-like. I cannot deal here with the difficulties peculiar to the classification of allochronous forms. They are of much less practical importance than those concerning contemporary organisms and may be safely left to the palaeontologists. The text-book of Simpson (1961) and the references given therein cover the subject adequately. Phyletic numerical analysis may doubtless sometimes prove helpful in this field for the elucidation of phylogeny and the quantitative study of anagenetic change, but any purely phenetic analysis across different time levels is inappropriate for general classification. Tuomikoski (1967) gives a useful discussion of palaeontological classification over all time levels but appears

not to perceive the virtual impossibility of a satisfactory solution.

Amongst organisms at any one time level, it may at first seem obvious that, if we have deduced the topological structure of a divaricating tree and expressed it as a cladogram, we then have a hierarchy laid out before us. The nested groupings would then present no problem though assignment of rank would seem to have no non-arbitrary objective basis.

Unfortunately the situation is not so simple. Hull (1964) gives an admirable discussion of the difficulty of fitting the traditional Linnaean hierarchy (indeed any hierarchy) to the phylogenetic background, and the reasons why Simpson's criterion of consistency can be only imperfectly applied. There are two main sources of difficulty:

(1) The definiteness of a topological cladogram depends not only on our knowledge, which may be quite inadequate to indicate the *exact* sequence of branching, but also on the taxonomic *unit* of the cladogram. Any questions of monophyly will depend upon the unit adopted and here we encounter some circularity in the definition of categories, if monophyly is taken as a criterion (as it is by Simpson, 1961, and many other phyletic taxonomists). A stem of a cladogram may be simple if we *define* it as being of generic cross-section, as it were, at all points, but if we dissect it into species strands we may find quite a network.

(2) Providing our representation of time and populations is sufficiently coarse-grained, we may be able to say that species diverged at such-and-such a level but it is not so easy with higher categories. This is fundamentally because we insist on recognizing these categories by the "level of organization" or "broad adaptive patterns" of the organisms we refer to them, as well as by inferred monophyly (e.g., Simpson, 1953, 1961). This familiar double function of the taxonomic categories was neatly expressed by Huxley (1958) when he coined

the terms *grade* and *clade*, now widely adopted. The corresponding terms (due to Rensch, 1960) for processes of differentiation are *anagenesis* (evolutionary “advance”) and *cladogenesis* (evolutionary separation), with the addition of *stasigenesis* (evolutionary stasis).

To illustrate this double requirement, the familiar example will serve: The reptiles, mammals, and birds represent three grades, but cladistically the last two are simply branches (clades, ignoring some parallelism in the case of the mammals) of equal status to other branches still treated as reptiles. In set-theoretic terms, the set: {all vertebrates whose ancestors attained the reptilian level of organization} includes two (now) pretty clearly defined non-intersecting *proper subsets*: {mammals} and {birds}, which we do *not* now call “reptiles”, together with the residue (the complement of the union of {mammals} and {birds}) which we *do* call reptiles but are jointly defined (within the original set) only by the fact that they are not mammals or birds.

Ehrlich (1964) has said that “a system based on phylogenetic relationships will not necessarily represent degree of phenetic differentiation”. Restricting “phenetic differentiation” to what we regard as adaptationally and practically important, this is quite true.

Now, in such cases the traditional hierarchy is *not conformable* with the cladistic hierarchy, and the situation recurs throughout the living world. It may be argued that the classification actually used is essentially phenetic. So, in a sense, it is, but the choice of attributes used and the actual assignment of individual taxa take evolutionary (but not only cladistic) considerations very much into account.

Evolution, indeed, is *not* summed up merely in the topology of a phyletic tree. Phenetic differences of many kinds (underlain by the genetic information and the systems for its exchange, suppression, modification, and so forth) are the very stuff

on which selection works. As I have said earlier, because we are part of the world in which selection operates and because what is a practical matter for selection is often a practical matter for us, we are also concerned with such phenetic characteristics, and a purely cladistic classification will not meet our needs. Nor will a patristic system.

Thus we are presented with a dilemma: If we choose to classify purely phenetically (this *need* not be hierarchic—overlapping clusters or simply ordination are systematizations, though psychologically unsatisfying to most of us) then we have no firm basis and are ignoring phylogeny, which *does* interest us. If we choose to classify purely cladistically (assuming that we have enough information) then we fail to display phenetic aspects of nature which are highly important in the evolutionary process, to the organisms themselves, and to ourselves.

The customary attempted solution of the evolutionary taxonomist is a compromise. Its general rationale is given by Simpson (1961). It has been followed, inevitably with varying degrees of imperfection, by many taxonomists, sometimes avowedly (e.g., Johnson, 1959, pp. 76–77), sometimes implicitly, and sometimes even when accompanied by disavowal.

Such a procedure is obviously not very logical and its defence may present many problems. It is offensive to those who dislike compromise and demand a clear codification of all procedures (e.g., Sokal and Camin, 1965). But it serves our purpose very well, and we should think very hard before we abandon it. For an analogy we may look to language. Natural languages are complex, not very logical systems, full of redundancies and overlaps. Ordinary language will not do for mathematics or even for the discussion of semantics itself, so we invent special symbolic languages and meta-languages. Nevertheless, most of us have not yet found it necessary, feasible, or desirable to discard common language

even in scientific communication. With some modifications and supplementation, it serves us better than any substitute yet thought of, *because it suits our psychology* (which is far from simply logical and linear) and because it links us with the knowledge of the past.

If, for similar reasons, we retain the broad framework of Linnaean hierarchical taxonomy (though we may see fit to do away with some of its more uselessly archaic features), we must face the fact that *no optimal taxonomy exists*. We could discard Linnaean taxonomy for one of the innumerable phenetic systems (there is no optimality there, as we have seen), or we could use a purely cladistic system which could then (in theory) be optimized, down to a certain level, as to the nesting of sets but not as to ranking. As Hull (1964) has shown, this latter would result in a "prohibitively complex and asymmetrical classification", which would not meet our requirements as we see them at present.

Recognizing and accepting the procedures which man has followed in intuitive classification, Davidson (1967) suggests a radical departure from the formal models employed by other taximetrists, whether phenetic or phyletic. He recommends that mathematics should be used to construct a cybernetic model congruent with the "classification program used by man since time immemorial". His efforts will be worth watching.

One of the logicians' objections to the usual formal hierarchy has been the recognition of taxa which may include only one subtaxon of next lower rank, for example, unispecific ("monotypic") genera. This objection arises from a particular set-theoretic treatment under which the existence of such categories leads to what has been called Gregg's Paradox. This paradox, like most others, is a consequence of the logical system adopted. Under the more reasonable method of definition of taxon-names (in the sense of logic) used in a

logical analysis of the Linnaean hierarchy by Buck and Hull (1966), the paradox vanishes. We can thus rest easy with a long-established feature of systematics which is certainly meaningful to most biologists.*

The continual "improvement" of classifications (on whatever basis) leads at present to some instability in names, which certain biologists and others appear to find most vexatious (this is quite apart from formal changes due to newly discovered nomenclatural priorities, and so forth). Systematists themselves often do not mind these changes, at least in groups with which they are familiar, since they feel that the new nomenclature expresses the "cleaning-up" which has been achieved. Nevertheless, the instabilities do cause confusion. Various authors have suggested that there should be a fixed nomenclature, perhaps using single-word names for species, separate from classification. Some (e.g., Michener, 1963; Sokal and Camin, 1965) suggest the use of code numbers. Hull (1966) proposes the adoption of Michener's fixed numbers supplemented by a system of "phylogenetic numericulture" with adjustable "positional numbers" to indicate current views on phylogenetic arrangement. Perhaps some such system will eventually come, but at present it seems premature. The mnemonic value of words, and of binomials in particular, remains, despite challenges, and a great deal of information in the literature would become difficult of access if the present system were discarded. Similar views are expressed by Randal and Scott (1967), who point out the usefulness of natural-language input to data-processing machines, and also the greater detectability of errors by the eye when words are used. Parkes (1967) gives a similar qualified defence of traditional nomenclature, but we retain it chiefly, I think, *faute de mieux*.

* See Addendum.

CONCLUSION

So, neither phyletics nor phenetics will lead us to the optimal classification, the crock of gold, because it is not there. None the less, the elucidation of phylogeny can still proceed, and our unperfectable classifications can still be improved by reducing inconsistency until uncertainty or instability renders further change unprofitable. Above all, taxonomists can set themselves the task, as more and more of them do, of describing the situations they perceive, giving reasons for their judgments, and suggesting lines of inquiry. The investigation of phenetic and "biological" characteristics of organisms and of their evolution is the scientific part of taxonomy. Classification itself remains an art, but a disciplined and respectable art whose aim is to serve science, not to express the personality of its exponents.

The classificatory component of taxonomy cannot itself be made into a science by ill-founded philosophy or essentially arbitrary numerical procedures and taxonomists would be well advised to treat "philosophical" pheneticism as unjustified in its claims and unproductive in practice. They will surely find numerical techniques useful, particularly on the big problems, and especially if the taximetrist can free themselves of their pheneticist dogmas. Taxonomists may remember, however, that many other techniques and tools, new and old, are at their disposal, including the subtlety of the human intellect and its power of perceiving *Gestalt* and of bringing information and theoretical reasoning of all kinds to bear on a problem.

If systematics is made into a sterile exercise, a purely pragmatic service, or a playground for technicians, I would advise intelligent young biologists to steer well clear of it. On the other hand it can be, as I have found it, a stimulating and rewarding (though *inherently inexact*) branch of science, *with its dash of art, like most other stimulating branches of science, including mathematics itself*. To conclude

with the words of that wise and wide-ranging taxonomist, Lincoln Constance (1964), taxonomy in its widest sense, which goes far beyond the formulation of classifications, remains "an unending synthesis".

Acknowledgements

My colleague Dr. Barbara Briggs, by her helpful criticism, has prevented this journey through the mists of theory and philosophy from being even more disordered and discursive than it is; I am most grateful to her. My wife has patiently typed illegible and tortuous manuscript, and borne with long discussions; I am equally grateful to her. I wish to thank also (and to ask tolerance for the cut-and-thrust of debate from) my friends Dr. W. T. (Bill) Williams and Dr. Paul Ehrlich whose stimulating and provocative, but also constructive, words and works have partly brought about the present assessment of trends in taxonomy, with much of which I naturally do not expect them to agree!

References

- ADANSON, M. (see LAWRENCE, 1963).
 BIRCH, L. C., and EHRLICH, P. R., 1967a.—Evolutionary history and population biology. *Nature*, 214: 349-352.
 ———, and ———, 1967b.—Evolutionary history and taxonomy. *Syst. Zool.*, 16: 282-285.
 BONNER, R. E., 1964.—On some clustering techniques. *IBM J. Res. Development*, 8: 22-32.
 ———, 1965.—Cluster analysis. IBM Technical Report 17-168: 1-22.
 BOYCE, A. J., 1964.—The value of some methods of numerical taxonomy, with references to hominoid classification. In Heywood, V. H., and McNeill, J., "Phenetic and Phylogenetic Classification", Systematics Association Pub. No. 6: 47-65. London.
 BRIGGS, BARBARA G., and JOHNSON, L. A. S., 1968.—The status and relationships of the Australasian species of *Typha*. *Contrib. N.S.W. nat. Herb.*, 4: 57-69.
 BRITTON, K. W., 1958.—Logical positivism. In Encyclopaedia Britannica, 1958 British ed., 14: 332-334. (Encyclopaedia Britannica, Ltd., Chicago, London, Toronto).
 BUCK, R. C., and HULL, D. L., 1966.—The logical structure of the Linnaean hierarchy. *Syst. Zool.*, 15: 97-111.
 CAIN, A. J., 1959.—Deductive and inductive

- methods in post-Linnaean taxonomy. *Proc. Linn. Soc. London*, 170: 187-217.
- , and HARRISON, G. A., 1958.—An analysis of the taxonomist's judgment of affinity. *Proc. zool. Soc. London*, 135: 1-31.
- CAMIN, J. H., and SOKAL, R. R., 1965.—A method for deducing branching sequences in phylogeny. *Evolution*, 19: 311-326.
- CAROLIN, R., 1967.—The concept of the inflorescence in the Order Campanulales. (Presidential address.) *Proc. Linn. Soc. N.S.W.*, 92: 7-26.
- CAVALLI-SFORZA, L. L., and EDWARDS, A. W. F., 1967.—Phylogenetic analysis: models and estimation procedures. *Evolution*, 21: 550-570.
- CLIFFORD, H. T., and GOODALL, D. W., 1967.—A numerical contribution to the classification of the Poaceae. *Aust. J. Bot.*, 15: 499-519.
- CONSTANCE, L., 1964.—Systematic botany—an unending synthesis. *Taxon*, 13: 257-273.
- COURANT, R., and ROBBINS, H., 1961.—“What is Mathematics?”, 11th printing. (Oxford University Press, London, New York and Toronto.)
- CROVELLO, T. J., 1968.—Key communal cluster analysis as a taxonomic tool. *Taxon*, 17: 241-258.
- DANTZIG, T., 1962.—“Number. The Language of Science.” 4th ed. (George Allen and Unwin, London.)
- DAVIDSON, R. A., 1967.—A cybernetic approach to classification: Preliminaries. *Taxon*, 16: 1-7.
- DAVIDSON, R. A., and DUNN, ROSALIE A., 1966.—A new biometric approach to systematic problems. *BioScience*, 16: 528-536.
- DAVIS, P. H., and HEYWOOD, V. H., 1963.—“Principles of Angiosperm Taxonomy.” (Oliver and Boyd, Edinburgh and London.)
- DOBZHANSKY, TH., 1951.—“Genetics and the Origin of Species.” 3rd ed. (Columbia University Press, New York.)
- EDWARDS, A. W. F., and CAVALLI-SFORZA, L. L., 1964.—Reconstruction of evolutionary trees. In Heywood, V. H., and McNeill, J., “Phenetic and Phylogenetic Classification”. Systematics Association Pub. No. 6: 67-76. London.
- EHRENDORFER, F., 1964a.—Evolution and karyotype differentiation in a family of flowering plants: Dipsacaceae. In “Genetics Today”, *Proc. XI int. Congr. Genetics*: 399-407.
- , 1964b.—Cytologie, Taxonomie und Evolution bei Samenpflanzen. In Turrill, W. B. (ed.) “Vistas in Botany. IV. Recent Researches in Plant Taxonomy”. 99-186. (Pergamon Press, Oxford etc.)
- EHRlich, P., 1964.—Some axioms of taxonomy. *Syst. Zool.*, 13: 109-123.
- , 1965.—Numerical taxonomy. *Ann. Rep. Proc. Papua and New Guinea sci. Soc.*, 17: 10-14.
- EHRlich, P. R., and EHRlich, ANNE H., 1967.—The phenetic relationships of the butterflies. I. Adult taxonomy and the nonspecificity hypothesis. *Syst. Zool.*, 16: 301-317.
- ESTABROOK, G. F., 1967.—An information theory model for character analysis. *Taxon*, 16: 86-97.
- FITCH, W. M., and MARGOLASH, E., 1967.—Construction of phylogenetic trees. *Science*, 155: 279-284.
- GILMOUR, J. S. L., 1940.—Taxonomy and philosophy. In Huxley, J. (ed.) “The New Systematics.” (Oxford University Press, London.)
- , and WALTERS, S. M., 1964.—Philosophy and classification. In Turrill, W. B. (ed.) “Vistas in Botany. IV. Recent Researches in Plant Taxonomy.” 1-22. (Pergamon Press, Oxford, etc.)
- GOODALL, D. W., 1964.—A probabilistic similarity index. *Nature*, 203: 1098.
- , 1966.—Classification, probability and utility. *Nature*, 211: 53-54.
- GOWER, J. C., 1966.—Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53: 325-338.
- GRIFFITH, J. G., 1967.—Numerical taxonomy and textual criticism of a classical Latin author. *Nature*, 215: 326.
- HALL, A. V., 1965.—The peculiarity index, a new function for use in numerical taxonomy. *Nature*, 206: 952.
- , 1967a.—Methods for demonstrating resemblance in taxonomy and ecology. *Nature*, 214: 830-831.
- , 1967b.—Studies in recently developed group-forming procedures in taxonomy and ecology. *J. S. Afr. Bot.*, 33: 185-196.
- HAMANN, U., 1961.—Merkmalsbestand und Verwandtschaftsbeziehungen der Farinosae. Ein Beitrag zum System der Monokotyledonen. *Willdenowia*, 2: 639-768.
- HEISER, C. B., SORIA, J., and BURTON, D. L., 1965.—A numerical taxonomic study of *Solanum* species and hybrids. *Amer. Naturalist*, 99: 471-488.
- HOLLINGSWORTH, C. A., 1967.—“Vectors, Matrices, and Group Theory for Scientists and Engineers.” (McGraw-Hill, New York, etc.)
- HULL, D. L., 1964.—Consistency and monophyly. *Syst. Zool.*, 13: 1-11.
- , 1965.—The effect of essentialism on taxonomy. Two thousand years of stasis. *Brit. J. Phil. Sci.*, 15: 314-326, and 16: 1-18.
- , 1966.—Phylogenetic numerclature. *Syst. Zool.*, 15: 14-17.
- , 1967.—Certainty and circularity in evolutionary taxonomy. *Evolution*, 21: 174-189.
- HUXLEY, J. (ed.) 1940.—“The New Systematics.” (Oxford University Press, London.)
- , 1942.—“Evolution. The Modern Synthesis.” (George Allen and Unwin, London.)
- , 1958.—Evolutionary processes and tax-

- onomy with special reference to grades. *Uppsala Univ. Arsskr.*, 1958: 21-38.
- IRWIN, H. S., and ROGERS, D. J., 1967.—Monographic studies in *Cassia* (Leguminosae-Caesalpinioideae). II. A taximetric study of section *Apoucouita*. *Mem. New York bot. Gard.*, 16: 71-118.
- JANCEY, R. C., 1965.—Numerical methods in taxonomy. *PROC. LINN. SOC. N.S.W.*, 90: 335-340.
- , 1966.—Multidimensional group analysis. *Aust. J. Bot.*, 14: 127-130.
- JOHNSON, L. A. S., 1959.—The families of cycads and the Zamiaceae of Australia. *PROC. LINN. SOC. N.S.W.*, 84: 64-117.
- , and BRIGGS, BARBARA G., 1963.—Evolution in the Proteaceae. *Aust. J. Bot.*, 11: 21-61.
- KATZ, M. W., and TORRES, A. M., 1965.—Numerical analyses of caespitose zinnias. *Brittonia*, 17: 335-349.
- KENDALL, M. G., 1950.—Factor analysis as a statistical technique. *J. roy. stat. Soc. London, Ser. B.*, 12: 60-73.
- , 1957.—“A Course in Multivariate Analysis.” (Charles Griffin and Company, London.)
- KENDRICK, W. B., 1965.—Complexity and dependence in computer taxonomy. *Taxon*, 14: 141-154.
- , and WERESUB, LUELLA K., 1966.—Attempting neo-Adansonian computer taxonomy at the ordinal level in the Basidiomycetes. *Syst. Zool.*, 15: 307-329.
- KEY, K. H. L., 1967.—Operational homology. *Syst. Zool.*, 16: 275-276.
- LANCE, G. N., and WILLIAMS, W. T., 1966a.—Computer analysis for hierarchical polythetic classification (“similarity analyses”). *Computer J.*, 9: 60-64.
- , and ———, 1966b.—A generalized sorting strategy for computer classifications. *Nature*, 212: 218.
- , and ———, 1967a.—A general theory of classificatory sorting strategies. I. Hierarchical systems. *Computer J.*, 9: 373-380.
- , and ———, 1967b.—A general theory of classificatory sorting strategies. II. Clustering systems. *Computer J.*, 10: 271-277.
- , and ———, 1967c.—Mixed-data classificatory programs. I. Agglomerative systems. *Aust. Computer J.* (Nov., 1967): 1-6.
- LANCE, R. T., STENHOUSE, S. N., and OFFLER, CHRISTINA E., 1965.—Experimental appraisal of certain procedures for the classification of data. *Aust. J. Biol. Sci.*, 18: 1189-1205.
- LAWRENCE, G. H. M. (ed.), 1963.—“Adanson. The Bicentennial of Michel Adanson's ‘Familles des plantes.’” Pt. I. (Hunt Botanical Library, Pittsburgh.)
- LEWIS, H., 1957.—Genetics and cytology in relation to taxonomy. *Taxon*, 6: 42-46.
- LIDDELL, H. G., and SCOTT, R., 1864.—“A Greek-English Lexicon.” 5th ed. (Clarendon Press, Oxford.)
- LONG, C. A., 1966.—Dependence in taxonomy. *Taxon*, 15: 49-51.
- LÖVE, Å., 1964.—The biological species concept and its evolutionary structure. *Taxon*, 13: 33-45.
- MACKERRAS, I. M., 1964.—The classification of animals. *PROC. LINN. SOC. N.S.W.*, 88: 324-335.
- MACNAUGHTON-SMITH, P., 1965.—Some statistical and other numerical techniques for classifying individuals. H.M.S.O. Home Office Research Unit Rep. No. 6.
- MANNETJE, L. 't, 1967.—A comparison of eight numerical procedures applied to the classification of some African *Trifolium* taxa based on *Rhizobium* affinities. *Aust. J. Bot.*, 15: 521-528.
- MASON, H. L., 1957.—The concept of the flower and the theory of homology. *Madroño*, 14: 81-95.
- MAXWELL, E. A., 1965.—“Algebraic Structure and Matrices.” (Cambridge University Press, Cambridge.)
- MAYR, E., 1939.—“Systematics and the Origin of Species.” 1st ed. (Columbia University Press, New York.)
- , 1957a.—Species concepts and definition. In MAYR, E. (ed.), “The Species Problem.” (A.A.A.S., Washington.)
- , 1957b.—Difficulties and importance of the biological species. In MAYR, E. (ed.), “The Species Problem.” (A.A.A.S., Washington.)
- , 1963.—“Animal Species and Evolution.” (Harvard University Press, Cambridge, Mass.)
- MENITSKIĬ, I. V. L. [MENITZKY, J. L.], 1966.—Ispol'zovanie kolichestvennykh otsenok skhodstva v sistematike. [The use of the quantitative estimates of resemblance in taxonomy]. *Bot. Zhurn.*, 51: 354-371.
- MICHENER, C. D., 1963.—Some future developments in taxonomy. *Syst. Zool.*, 12: 151-172.
- MINCKOFF, E. C., 1965.—The effects on classification of slight alterations in numerical technique. *Syst. Zool.*, 14: 196-213.
- OLSON, E. C., 1964.—Morphological integration and the meaning of characters in classification systems. In Heywood, V. H., and McNeill, J., “Phenetic and Phylogenetic Classification.” Systematics Association Pub. No. 6: 123-156. London.
- ORLOCI, L., 1967.—Data centering: a review and evaluation with reference to component analysis. *Syst. Zool.*, 16: 208-212.
- PAIS, A., 1967.—Review of: Wigner, E. P., 1967. “Symmetries and Reflections.” *Science*, 157: 911-912.
- PARKES, K. C., 1967.—A qualified defense of traditional nomenclature. *Syst. Zool.*, 16: 268-273.
- RANDAL, J. M., and SCOTT, G. H., 1967.—Linnaean nomenclature: an aid to data processing. *Syst. Zool.*, 16: 278-281.

- RENSCH, B., 1960.—"Evolution above the Species Level". (Columbia University Press, New York).
- REYNOLDS, K. A., 1965.—Numerical taxonomy and comparative elaborateness, with a speculation on unused genes. *Nature*, 206: 166-168.
- ROGERS, D. J., 1963.—Taximetrics—new name, old concept. *Brittonia*, 15: 285-290.
- ROHLF, F. J., 1965.—A randomization test of the nonspecificity hypothesis in numerical taxonomy. *Taxon*, 14: 262-267.
- , and SOKAL, R. R., 1965.—Coefficients of correlation and distance in numerical taxonomy. *Univ. Kansas Sci. Bull.*, 45: 3-27.
- ROLLINS, R. C., 1965.—On the bases of biological classification. *Taxon*, 14: 1-6.
- ROOM, T. G., and MACK, J. M., 1966.—"The Sorting Process." (Sydney University Press, Sydney).
- RUBIN, J., 1967.—Optimal classification into groups: an approach for solving the taxonomy problem. *J. theoret. Biol.*, 15: 103-144.
- SAWYER, W. W., 1955.—"Prelude to Mathematics." (Penguin Books, Harmondsworth, Baltimore, and Ringwood.)
- SCORA, R. W., 1967.—Divergence in *Monarda* (Labiatae). *Taxon*, 16: 499-505.
- SEAL, H. L., 1964.—"Multivariate Statistical Analysis for Biologists." (Methuen, London.)
- SHEALS, J. G., 1965.—An application of computer techniques to Acarine taxonomy: a preliminary examination with species of the *Hypoaspis-Androlaelaps* complex, Acarina. *Proc. Linn. Soc. London*, 176: 11-21.
- SILVESTRI, L. G., 1964.—[Discussion] In Heywood, V. H., and McNeill, J., "Phenetic and Phylogenetic Classification." Systematics Association Pub. No. 6: 160. London.
- , and HILL, L. R., 1964.—Some problems of the taxometric approach. *Ibid.* 87-103.
- SIMPSON, G. G., 1953.—"The Major Features of Evolution." (Columbia University Press, New York.)
- , 1961.—"Principles of Animal Taxonomy." (Columbia University Press, New York.)
- SNEATH, P. H. A., 1957a.—Some thoughts on bacterial classification. *J. gen. Microbiol.*, 17: 184-200.
- , 1957b.—The application of computers to taxonomy. *J. gen. Microbiol.*, 17: 201-226.
- , 1958.—Some aspects of Adansonian classification and of the taxonomic theory of correlated features. *Ann. Microbiol. Enzimol.*, 8: 261-268.
- , and SOKAL, R. R., 1962.—Numerical taxonomy. *Nature*, 193: 855-860.
- SOKAL, R. R., 1961.—Distance as a measure of taxonomic similarity. *Syst. Zool.*, 10: 70-79.
- , 1965.—Statistical methods in systematics. *Biol. Rev.*, 40: 337-391.
- , and CAMIN, J. H., 1965.—The two taxonomies: areas of agreement and conflict. *Syst. Zool.*, 14: 176-195.
- , and MICHENER, C. D., 1967.—The effects of different numerical techniques on the phenetic classification of bees of the *Hoplitis* complex (Megachilidae). *Proc. Linn. Soc. Lond.*, 178: 59-74.
- , and ROHLF, F. J., 1962.—The comparison of dendrograms by objective methods. *Taxon*, 11: 33-40.
- , and SNEATH, P. H. A., 1963.—"Principles of Numerical Taxonomy." (Freeman and Company, San Francisco and London.)
- , and ———, 1966.—Efficiency in taxonomy. *Taxon*, 15: 1-21.
- SPORNE, K. R., 1948.—Correlation and classification in dicotyledons. *Proc. Linn. Soc. London*, 1960: 40-47.
- , 1954.—Statistics and the evolution of dicotyledons. *Evolution*, 8: 55-64.
- , 1956.—The phylogenetic classification of the angiosperms. *Biol. Rev.*, 31: 1-29.
- , 1960.—Correlation of biological characters. *Proc. Linn. Soc. London*, 171: 83-88.
- STEBBINS, G. L., 1950.—"Variation and Evolution in Plants." (Oxford University Press, London.)
- TAKHTAJAN, A. L., 1953.—Phylogenetic principles of the system of higher plants. *Bot. Rev.*, 19: 1-45.
- TAYLOR, F. SHERWOOD, 1939.—"Inorganic and Theoretical Chemistry." 5th ed. (William Heinemann, London.)
- THORNE, R. F., 1963.—Some problems and guiding principles of angiosperm phylogeny. *Amer. Nat.*, 97: 287-305.
- THORNTON, I. W. B., and WONG SIU KAI, 1967.—A numerical taxonomic analysis of the Peripsocidae of the Oriental Region and the Pacific Basin. *Syst. Zool.*, 16: 217-240.
- THROCKMORTON, L. H., 1965.—Similarity versus relationship in *Drosophila*. *Syst. Zool.*, 14: 221-335.
- TUOMIKOSKI, R., 1967.—Notes on some principles of phylogenetic systematics. *Ann. ent. fenn.*, 33: 137-147.
- WALLACE, C. S., and BOULTON, D. M., 1968.—An information measure for classification. *Computer J.*, 11: 185-194.
- WALTERS, S. M., 1965.—"Improvement" versus stability in botanical classification. *Taxon*, 14: 6-12.
- WARBURTON, R. E., 1967.—The purposes of classification. *Syst. Zool.*, 16: 241-245.
- WATSON, L., WILLIAMS, W. T., and LANCE, G. N., 1966.—Angiosperm taxonomy: a comparative study of some novel numerical techniques. *J. Linn. Soc. London (Bot.)*, 59: 491-501.
- , ———, and ———, 1967.—A mixed-data numerical approach to angiosperm tax-

- onomy: the classification of Ericales. *Proc. Linn. Soc. London*, 178: 25-35.
- WIGNER, E. P. (see PAIS, 1967).
- WILLIAMS, W. T., 1967.—The computer botanist. (Presidential address to Section M, A.N.Z. A.A.S.). *Aust. J. Sci.*, 29: 262-266.
- , and DALE, M. B., 1965.—Fundamental problems in numerical taxonomy. In PRESTON, R. D.—“Advances in Botanical Research”, 2: 35-68. (Academic Press, London and New York.)
- WILSON, E. O., 1965.—A consistency test for phylogenies based on contemporaneous species. *Syst. Zool.*, 14: 214-220.
- WIRTH, M., ESTABROOK, G. F., and ROGERS, D. J., 1966.—A graph theory model for systematic biology, with an example for the Oncidiinae (Orchidaceae). *Syst. Zool.*, 15: 59-69.

ADDENDUM

Since this paper was written in early 1968, it has inevitably become out of date as a review. There is little point in citing the numerous recent publications on methods or applications of taxometrics. The former deal chiefly with mathematical or technical refinements which do not affect my general thesis, while the latter mostly do not seem to me to have demonstrated any outstanding effectiveness in improving taxonomy, and often indeed the authors have a rather naively hopeful approach.

Readers of *Systematic Zoology* and *Taxon* will be able to keep generally abreast of developments, if they wish, but several publications bearing on the general philosophy of taxonomy deserve special mention. The cladistic position is presented in detail by Hennig (1966) and discussions of this important work, and the arguments against a purely cladistic approach, have been given in several papers in *Systematic Zoology* and elsewhere. The important paper on “The Phylogenetic Fallacy” by Colless (1967) and others by the same author have also been discussed in this journal. Colless’s case is well argued but he takes little account of the inapplicability of simple logic to the complexity of systematic decision-making. It is manifestly *not* true that supposedly theory-free phenetics boils down to the same thing as

phyletics, which draws on the whole body of evolutionary theory, even though the latter depends of course on phenetic observations. David Hull (1968) deals devastatingly with naive operationism and upholds the place of theory and the usefulness of the concept of evolutionary homology. Ghiselin (1969) has stressed several of the points also made in the present paper, in particular that it is nonsense to speak of “overall similarity”.

This reprinting has provided the opportunity of correcting a few minor errors in the text. To the remarks on “Gregg’s Paradox” on page 234, I would now add that, although Gregg (1968) criticized Buck and Hull’s formulation on formal grounds, quite satisfactory formalizations are in fact possible and Gregg’s restricted set theory is in fact inapplicable (see also Buck and Hull, 1969; Hull and Snyder, 1969). This pseudo-problem is completely avoided by the more useful and more readily generalized model of Jardine and Sibson (1968, see also Jardine, 1969).

ADDITIONAL REFERENCES

- BUCK, R. C., AND D. L. HULL. 1969. Reply to Gregg. *Syst. Zool.*, 18:354-357.
- COLLESS, D. H. 1967. The phylogenetic fallacy. *Syst. Zool.*, 16:289-295.
- GHISELIN, M. T. 1969. The principles and concepts of systematic biology. In “Systematic Biology,” 45-55. National Academy of Sciences, Washington.
- GREGG, J. R. 1968. Buck and Hull: a critical rejoinder. *Syst. Zool.*, 17:342-344.
- HENNIG, W. 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana, Chicago, London.
- HULL, D. L. 1968. The operational imperative: sense and nonsense in operationism. *Syst. Zool.*, 17:438-457.
- HULL, D. L., AND D. P. SNYDER. 1969. Contemporary logic and evolutionary taxonomy: a reply to Gregg. *Syst. Zool.*, 18:347-354.
- JARDINE, N., AND R. SIBSON. 1968. A model for taxonomy. *Math. Biosc.*, 2:465-482.
- JARDINE, N. 1969. A logical basis for biological classification. *Syst. Zool.*, 18:37-52.

*National Herbarium of New South Wales,
Royal Botanic Gardens, Sydney, 2000, Australia.*