The Logical Structure of the Linnaean Hierarchy

ROGER C. BUCK AND DAVID L. HULL

The Linnaean hierarchy is a system of nested classes whose members are individual organisms. It will be the purpose of this paper to set out the logic of this system in a manner which is faithful to taxonomic theory and practice, which avoids certain paradoxes in previous attempts, and which is in terms that a biologist might reasonably be expected to understand.

Introduction

The aim of taxonomy is classification. The product of any classificatory enterprise is a set or structure of classes. The Linnaean hierarchy, as it is now understood, is such a structure of classes. The discipline of logic has considerable bearing on the nature of any such structure. Using the tools which logic affords, it is our aim to provide a formal characterization of the structure of the Linnaean hierarchy. Past treatments of this kind in biology have not proved very influential, either because the logicians who have attacked the problem have not appreciated the biological principles and procedures involved or because their formalizations have proved incomprehensible to most biologists. We have tried to avoid both pitfalls. We have provided a formal characterization of what the structure of the Linnaean hierarchy *is* rather than legislating what it *should* be, and we have done so with a minimum of technical terms and symbols. In fact we use only two symbols foreign to taxonomy: $\in$ and $\subset$.

The most significant previous attack on our problems is to be found in the work of John R. Gregg (Gregg, 1954; see also Woodger, 1937, 1948, 1952). By and large, our analysis of the structure of the taxonomic hierarchy agrees with Gregg's, although in one important respect we depart from him. As a consequence of his formalization, Gregg is forced to disparage the current practice of monotypic classification. Our treatment requires no such limitations. We shall also present our analysis as far as possible in ordinary English. Gregg's exposition employs not just the ideas but also the notation of that branch of logic (or mathematics) called set theory, making his presentation forbiddingly difficult to follow. It is this difficulty, we believe, which has robbed his work of the impact and readership it deserves.

We shall characterize the logical relations among individual organisms, among the classes of these individual organisms called taxa, and among the classes of these classes called categories. Although these logical relations are central to the practice and theory of classification, taxonomists have not always been completely clear as to the differences between them. Examples of such problem areas include the differences between:

a) the relation which an individual organism has to a taxon and the relation a lower taxon has to a higher taxon,

b) the relation a lower taxon has to a higher taxon and the relation any taxon has to its category,

c) the type of properties attributable to individual organisms and the type of properties attributable to taxa,

d) all the relations mentioned in (a–c) and the relation which categories have to each other.

It will be the purpose of this paper to explain these formal distinctions as informally as possible by means of the logic of classes. We therefore turn first to an informal exposition of some logical properties of classes. Our examples are deliberately chosen from taxonomic and non-taxonomic areas. This procedure serves to reinforce our claim that the logic of taxonomy is a special case of the general logic of classes.

Class Membership

The most obvious point about classes is that they have members. They could not lack this property and be classes. It is both true and important that the number of members of some classes may be zero. Even so, these classes do not thereby cease to be classes. Thus, if taxa are classes, they must have members. Taxonomists have been rather uncertain concerning the role that individuals are to play in the taxonomic enterprise. On the one hand, they are the observables. They are what are collected, studied and catalogued. But on the other hand, no reference is ever made to them in a classification. "Classification involves only groups; no entity possible in classification is an individual" (Simpson, 1961). Individuals are identified. Taxa are classified. This terminological distinction reflects the intuitive feeling taxonomists have had for the difference between the relation which an individual has to a taxon and the relation which taxa have to each other. This intuition can be made fully explicit using the logic of classes.[1]

Ordinary taxonomic identification of an organism results in the statement that the organism is a member of a certain class—a taxon. The standard logical notation in which class membership assertions are symbolized is exemplified by

$$\text{Gargantua} \in \text{Primata}.$$

In statements of this sort we shall let lower case letters represent the names of individuals and upper case letters represent the names of taxa. Thus, the general form of a statement which says that an individual is a member of a particular taxon is

$$a \in F.$$

The sign $\in$ is read "is a member of." The name of the member always appears to the left of $\in$, and the name of the class of which it is a member always appears to the right of the sign. This conventional ordering is important because $\in$ is what logicians call an asymmetrical relation. This is to say that wherever "$a \in F$" is true, it cannot be the case that "$F \in a$" also obtains. Indeed, the converse of a true (or even intelligibly false) membership claim will be meaningless. The statement that Primata $\in$ Gargantua does not succeed even in saying something false. It is nonsense.

The notational distinction between upper and lower case letters has additional importance because class names can also appear to the left of $\in$, wherever it is said, for example, that *Homo sapiens* is a species or that Chordata is a phylum. These statements claim that a given taxon is a member of a certain category. Since a taxon is already a class, the category of which it is a member must be a class of classes. In an obvious extension of our symbolism, we shall use the Roman numerals in large type face to represent the names of categories. The general form of a statement assigning a taxon to a category is

$$F \in \text{II}.$$

Thus, the relation which individuals have to taxa is the *same* relation which taxa have to their respective categories—class membership. This formulation can be seen to be appropriate by asking what are the members of the category species? The answer would be a list of the names of spe-

[1] Taxonomists sometimes try to avoid explicating the relation between individuals and taxa by introducing classes termed "populations," which are not named and do not occur in a classification but which are the operational units in the process of classification. Nevertheless, populations are still classes, and they must be classes *of* something. They still must have members. As useful as populations are for other purposes, they do not spare the taxonomist the necessity of explicating the role which individual organisms and their properties play in taxonomy.

cies, e.g., *Homo sapiens, Bos bos, Canis familiaris,* and *Endameba gingivalis.* What are the members of the species *Homo sapiens*? The answer would be a list of the names of individual human beings, e.g., John R. Gregg, Roger Buck, David Hull, and Ludwig von Beethoven.

Perhaps the significance of these class membership levels can be made clearer by some examples, reading from left to right:

Individuals		*Taxa*		*Categories*
Gargantua	∈	Chordata	∈	phylum
Gargantua	∈	Vertebrata	∈	subphylum
David Hull	∈	*Homo sapiens*	∈	species
Flicka	∈	Mammalia	∈	class
Flicka	∈	Chordata	∈	phylum

Besides introducing symbols for the three levels just discussed, Gregg proposes two additional symbols, G and C. G is the class of all taxa and is, therefore, a class of classes. Besides saying, as we did earlier, that Chordata is a phylum, we might also want to say that Chordata is a taxon. This would be symbolized by saying that Chordata is a member of Gregg's G. C is the class of all categories and is, therefore, a class of classes of classes. If we want to say that the genus is a taxonomic category, we are making another membership claim. This would be symbolized by saying that the genus is a member of Gregg's C. Since we have already stipulated that upper case letters are to stand for names of taxa, Gregg's symbolism will not be followed here. Instead, the phrases "the class of all taxa" and "the class of all categories" will themselves be used.

Gregg's four levels of expression related by ∈ are as follows:

N^1—the names of individual organisms (our lower case English letters)
N^2—the names of taxa (our upper case English letters)
N^3—the names of categories (our Roman numerals) and the class of all taxa (Gregg's G)
N^4—the class of all categories (Gregg's C).

Using these level designations, Gregg lays down criteria for well-formed membership statements (∈-statements). A well-formed ∈-statement must be such that the expression on the left hand side is exactly one level lower than that on the right hand side. Any other combination of levels is ill-formed.

An intuitive idea of what it means for a statement to be well-formed or ill-formed can be obtained by reference to truth or falsity. A well-formed membership statement will be either true or false. For example, "Chordata ∈ phylum" and "Chordata ∈ kingdom" are both well-formed, although the first is true and the second false. An ill-formed membership statement will be neither true nor false but nonsensical; for example,

Gargantua	∈	genus
Gargantua	∈	category
Genus	∈	phylum
Mammalia	∈	Chordata

These statements are ill-formed and are strictly without sense. This can be appreciated by translating them into ordinary English. They seem puzzling and vaguely "false." The logical distinctions just made should show why they are puzzling and why they should not be termed literally false.

Failure to distinguish between Gregg's levels N^1 through N^4, like any logical failure, makes no difference providing no misunderstandings ensue. But misunderstandings *have* ensued and *serious* misunderstandings. For example, Ernst Mayr (1942: 103) says:

> Even though Linnaeus recognized the variety as an infraspecific category, in exceptional cases, nevertheless the species was, for him, the basic category.

Here Mayr is saying that *the* variety is a member of the class of infraspecific categories and that *the* species is a member of the class of basic categories. In each case he is saying that a category (N^3) is a member of a class of categories (N^4).

A few lines further he says, however:

> Before we begin to define them we must emphasize that to some extent all taxonomic categories are collective units.

> Even the subspecies is generally composed of a number of slightly distinct populations, some of which connect it with neighboring subspecies.

It is difficult to decide what Mayr means in this quotation because he calls both taxa and categories "categories." He tries to distinguish between them by referring to a taxon as *a* category and a category as *the* category, but English syntax does not permit a consistent use of the articles in this way. To make matters worse, the name of the category most frequently referred to—species—is the same in the singular and the plural. The most likely translation of Mayr's statements into the terminology which he himself has now adopted[2] is that all taxa are collective units, since even subspecies (N^2) are generally composed of a number of slightly distinct populations which occasionally connect one subspecies to another.

In the preceding quotations, Mayr not only ranges over three different logical levels related by class membership with no terminological distinctions to signal these shifts, but also has shifted in the last sentence from the membership relation to an entirely different relation under the cover of the phrase "composed of." Categories are collective units because they have taxa as members. Taxa are collective units both because they have individuals as their members and because they have taxa included in them. Populations are related to subspecies, subspecies to species, species to genera, and so on—but *not* by the membership relation as we shall soon see.

Class Inclusion

The most obvious relations between classes in taxonomy are those generated by the grouping of lower taxa under higher taxa, which account for the use of the word "hierarchy" to characterize the product of the taxonomic enterprise. The major claim of taxonomy to the title of a systematic discipline lies in its stress on such groupings. The standard notations of taxonomy, including the indentation of the names of lower taxa under those of higher taxa and the repetition of generic names as parts of species' names, serve to emphasize such groupings.

What then is the relation between a lower taxon and a higher taxon under which it falls? For example, what is the relation between Vertebrata and Chordata? The answer is clear enough. Such a grouping says that the members of the taxon mentioned on the left hand side are also members of the taxon mentioned on the right hand side. This relation is called by logicians "class inclusion" and is symbolized by $\subset$, such that the statement

$$\text{Vertebrata} \subset \text{Chordata}$$

will mean that all vertebrates are also chordates—that all the members of the former are also members of the latter.[3] Since in all inclusion relations the members of one class are claimed to be members of another class, *only* classes can be related by $\subset$ and *both* classes must be on the same N-level. Thus if we say

$$F \subset G,$$

we are claiming that F and G are both classes on the same N-level, that the members of F are on the same N-level as the members of G, and that this level is one level lower than the level at which F and G are classed. If F and G are taxa, then the members referred to are individual organisms. A striking difference between the inclusion relation and the membership relation thus emerges. $\subset$-related entities *must* be on the same N-level. $\in$-related entities *cannot* be on the same N-level.

In taxonomy, inclusion relations hold only between taxa, the classes at the N^2-level.

[2] Mayr (1965). For further discussion of more recent examples of the taxon-category confusion, see Hull (1966).

[3] The actual relation in taxonomy, however, is stronger than that conveyed by "All F are G," since it entitles us to say not only that all primates are vertebrates but also that if anything *were* a primate, it *would* be a vertebrate. This extra strength is not a feature of all inclusion relations and will be discussed later.

Statements asserting that one taxon is included in another are extremely common in taxonomy. For example, the statements that Mammalia belongs to, falls under or is contained in Vertebrata are all translated uniformly as

$$\text{Mammalia} \subset \text{Vertebrata.}$$

The absence of inclusion relations at other N-levels is readily seen. Individual organisms at level N^1 cannot in the relevant sense include each other, because inclusion relations are defined only for classes. The notion of two individuals having members in common is nonsense. At level N^4 only one class, the class of all categories, is mentioned, and at a minimum two classes at the same level are required for any interesting inclusion relations.

The situation at level N^3 is different. Here the names of categories are listed; e.g., species, genus, family, order, etc. Since these entities are classes and there are many of them, an inclusion relation *could* hold between them. It is a function of the aims and procedures of taxonomy that categories are *not* included in each other. Taxonomic categories are deliberately intended by taxonomists to divide the taxa which are their members into mutually exclusive classes. Even a partial overlap between two categories, for example, a single taxon classified as both a family and a subfamily, would violate the spirit of the taxonomic enterprise and destroy its clarity. Thus, if *partial* overlap is ruled out, then the *complete* overlap entailed by inclusion would surely be ruled out at the category level.[4]

A final point about the inclusion relation and the membership relation is already implicit in what has been said. The inclusion relation is "transitive" and the membership relation is "intransitive." A transitive relation is such that if the first member has it to the second and the second has it to the third, the first member will have it to the third. For example, *taller than* is a transitive relation. If we know that Lincoln was taller than Henry VIII and Henry VIII was taller than Napoleon, then it follows with no further historical enquiry that Lincoln was taller than Napoleon. An intransitive relation is such that if the first member has it to the second and the second has it to the third, the first member *cannot* have it to the third. *Is the father of* is an example of an intransitive relation. If T. H. Huxley was the father of Leonard Huxley and Leonard Huxley was the father of Julian Huxley, then it follows necessarily that T. H. Huxley cannot be the father of Julian Huxley.

Now sometimes (actually not very often) logical principles such as the preceding are set forth as an aid in determining whether certain arguments are valid and whether (supposing their premises true) their conclusions should be accepted. But much more frequently logical principles are enunciated to provide a systematic basis for characterizing the *antecedently known* validity or invalidity of certain arguments. For example, it is a logical principle that when an argument is valid and its premises true, the conclusion of the argument must be true. In everyday life, someone may reason that since both psychopathic deviates and amateur pilots are happy-go-lucky, all amateur pilots are psychopathic deviates. To challenge his argument we may well construct another of the same form which has true premises and a patently false conclusion. An analog to the preceding case would be that since both husbands and wives are married, all husbands are wives. Probably any English speaking person would reject the latter (and hence the former) argument as invalid. And it is unlikely that any person who failed to reject these arguments as invalid could come to understand any logical principles. How-

[4] Since Gregg's G is also an N^3-level class, inclusion relations do exist at level N^3. For example, the class of species $\subset$ the class of taxa. What is ruled out is the class of species and the class of genera having some members in common. Partial overlap in categories must not be confused with overlap of taxa, however. Two categories overlap if a single taxon is a member of two different categories. Taxa overlap if one taxon is included in each of two distinct taxa at the next higher level.

ever, not all arguments are as simple as these, and even if they were, some of those who reject the invalid arguments might be curious as to where and how these arguments had gone wrong and might welcome a systematic explanation of invalidity based on logical principles. The same, we hope, can be said of taxonomists. The remainder of this section will be devoted to just such an explanation with respect to various examples of taxonomic reasoning.

The following taxonomic arguments are set forth both in English and schematically. In the English versions the deliberately ambiguous expression "falls under" is used. In the schematic version the inclusion relation is distinguished from the membership relation. The advantages of this procedure should become readily apparent.

One valid line of reasoning found in taxonomic writings is as follows:

IIIA.	Vertebrata falls under Chordata	$V \subset C$
	Chordata falls under Animalia	$C \subset A$
	Hence, Vertebrata falls under Animalia	$V \subset A$

If all vertebrates are chordates and all chordates are animals, then it follows necessarily from the transitivity of the inclusion relation that all vertebrates are animals. As Aristotle would have said, "Chordata" is serving as a middle term. All classes mentioned are on the same N-level (N^2), and no problems arise. This argument is, of course, an instance of one of the forms of valid argument most familiar in elementary logic courses. Its form is the same as

All Greeks are men
All men are mortal

Hence, all Greeks are mortal.

The membership relation on the other hand is intransitive. The following argument looks superficially like IIIA but is *not* valid.

IIIB.	Gargantua falls under Vertebrata	$g \in V$
	Vertebrata falls under subphylum	$V \in III$
	Hence, Gargantua falls under subphylum	g ? III

This argument has true premises and in the English version a wildly false conclusion. Therefore, it must be invalid. The invalidity flows from the intransitivity of the $\in$-relation. If $g \in V$ and $V \in III$, it is not legitimate to reason that g has either the inclusion or the membership relation to III. "Falls under" in the conclusion cannot be read as inclusion because Gargantua is not a *class* and because Gargantua and subphylum are on different N-levels. "Falls under" in the conclusion cannot be read as membership because Gargantua is *more* than one N-level below subphylum. In either case, the conclusion is ill-formed.

An analog to this invalid argument from ordinary English is

Sitting Bull is an Indian
Indians are disappearing

Hence, Sitting Bull is disappearing.

The second premise means that the class of Indians is a member of the class of disappearing classes; that is, of the class of classes whose membership is becoming less numerous and approaching zero.

Another line of reasoning quite similar to IIIB is given below. It differs from IIIB only in that all of the statements have been moved up one N-level.

IIIC.	Chordata falls under phylum	$C \in II$
	Phylum falls under category	$II \in$ category
	Hence, Chordata falls under category.	C ? category

Although this time both Chordata and category are classes, "falls under" in the conclusion cannot be read as inclusion because these two classes are not on the same N-level. Nor can it be read as membership since the two classes are more than one N-level apart. Again, in either case, the conclusion is ill-formed. An analog to this invalid argument from ordinary English is

The U.S.A. is a member of the U.N.
The U.N. is a supranational body

Hence, the U.S.A. is a supranational body.

In the preceding arguments, the relations in the premises were either both inclusion

(IIIA) or both membership (IIIB and IIIC). There is a second valid line of reasoning found in taxonomic writings which combines the membership relation and the inclusion relation in the premises.

IIID.	Gargantua falls under Vertebrata	$g \in V$
	Vertebrata falls under Chordata	$V \subset C$
	Hence, Gargantua falls under Chordata	$g \in C$

Because of the order of the membership and the inclusion relations in the premises, this argument is valid. If Gargantua is a member of Vertebrata and if all vertebrates are chordates, then Gargantua is also a member of Chordata. Again, this argument is an instance of one of the forms of valid argument most familiar in elementary logic courses. It is of the same form as

Socrates is a man
All men are mortal

Hence, Socrates is mortal.

The following argument looks superficially similar to IIID, but it differs in the *order* in which $\in$ and $\subset$ appear in the premises.

IIIE.	Vertebrata falls under Chordata	$V \subset C$
	Chordata falls under phylum	$C \in \Pi$
	Hence, Vertebrata falls under phylum	$V \ ? \ \Pi$

Here again our argument as presented in English must be invalid since the premises are true and the conclusion is false—but this time not wildly false. "Falls under" in the conclusion cannot be read as inclusion because Vertebrata and phylum are on different N-levels, but it *can* be read as membership. "Vertebrata $\in$ phylum" is well-formed. Our only complaint against it is that it is false. An analog from ordinary English would be

John Birchers are Americans
Americans are numerous

Hence, John Birchers are numerous.

Certainly the argument is fallacious since the truth of its premises does not guarantee the truth of its conclusion. Its conclusion, however, considered separately is a well-formed sentence and could logically be true. It is political convictions, not logical considerations, that falsify the conclusion. Similarly, in argument IIIE, it is taxonomic considerations, not logical considerations, that falsify the conclusion that Vertebrata is a phylum.

With these examples, it seems that the logical character of the membership and inclusion relations has been explained at sufficient length. It has been the purpose of the first half of this paper to explain and argue for a formal characterization of the taxonomic hierarchy whose main points are set forth in the following schematism:

N^1—individual organisms		N^2—taxa		N^3—categories
LBJ	$\in$	Mammalia	$\in$	class
		$\cup$		
LBJ	$\in$	Primata	$\in$	order
		$\cup$		
LBJ	$\in$	Hominidae	$\in$	family
		$\cup$		
LBJ	$\in$	Homo	$\in$	genus
		$\cup$		
LBJ	$\in$	*Homo sapiens*	$\in$	species

The analysis of this structure and various lines of reasoning permissible within it serves to emphasize that taxonomy has a logical as well as a biological significance.

Gregg's Paradox

Until now our presentation, though highly informal, has departed in no significant respect from that of Gregg. But Gregg's logical principles lead him into paradox in any case of monotypic classification. We must turn our attention to the identification, diagnosis and avoidance of what we shall call Gregg's Paradox. In broad outline Gregg employs set-theoretic procedures for specifying the meanings of taxonomic names, which lead him necessarily to a certain criterion of identity for taxa. This criterion states that two taxa are identical if they have all and only the same members, which may be symbolized using $=$ for the sign of identity as follows:

$$(F = G) \text{ if and only if } [(F \subset G) \text{ and } (G \subset F)]$$

By definition, in every case of monotypic classification, two taxa with exactly the same members are classed at different category levels. Thus, it follows that every case of monotypic classification will generate an overlap between two categories which are meant to be mutually exclusive. These results, not surprisingly, lead Gregg to make disparaging remarks about the practice of monotypic classification. For example, he concludes his 1954 monograph as follows:

> It is not hard to see how category overlapping hopelessly complicates the problem of stating anything general about the relations between the levels of a taxonomic system and the categories associated with that system—for not only may a single category have its membership distributed over several levels . . . , but also a single level may have its membership distributed over several categories. . . . For these reasons, and not only these, the whole theory of categorical association needs to be much further clarified.

The example of monotypic classification which Gregg discusses is taken from the classification of G. G. Simpson. In this classification (1945), the cohort Mutica includes only one order Cetacea, which in turn includes two suborders, whales and porpoises. In symbols this situation is expressed as follows:

$$\text{Mutica} \subset \text{Cetacea}$$
$$\text{Cetacea} \subset \text{Mutica.}$$

From the preceding definition of = it follows that

$$\text{Mutica} = \text{Cetacea.}$$

But Mutica is also a member of the category cohort and Cetacea is a member of the category order; i.e.,

$$\text{Mutica} \in \text{cohort}$$
$$\text{Cetacea} \in \text{order.}$$

Thus, it follows from the identity of Mutica and Cetacea that

$$\text{Mutica} \in \text{order}$$
$$\text{Cetacea} \in \text{cohort}$$

and the categories cohort and order are shown to overlap having two members in common.

G. G. Simpson (1961), one taxonomist who has read and understood Gregg's book, reacts to this paradox as follows:

> The examples just given demonstrate that these particular applications of set theory and symbolic logic to classification do not adequately take into account some relationships that are involved in actual classification and that do appear to be perfectly logical. That does not imply that these relationships could not be adequately incorporated in some system such as Gregg's or that it would not be worthwhile to do so.

Simpson is surely right. There are sometimes good reasons for classifying monotypically, and on occasion such classification can indeed be perfectly logical. For example, relatively great phenotypic distance (divergence) between two groups of taxa can be indicated, when one of the taxa has not diversified, only by classing both at a relatively high category level, making the undiversified taxon monotypic. Only considerations of symmetry in the hierarchy mitigate such practices. Simpson is also right in not ruling out the possibility of satisfactorily providing for monotypic classification in a system *such as* Gregg's. This is exactly what we propose to do. But before our proposal can be properly understood, we must pinpoint that aspect of Gregg's procedure which leads to his paradox.

Taxonomic Names: Extensional Definition

The following quotations from Gregg (1954) convey the spirit in which he approaches the definition of taxonomic names:

> It is frequently advantageous to be able to construct a designation for a given set by enumerating its members. In fact this is an excellent method of defining sets that are otherwise difficult to define.
>
> We shall construe taxonomic group names—'Insecta,' 'Primata,' '*Homo sapi-*

> *ens*,' 'Arthropoda,' and the rest—as set names, one and all, i.e., as names of sets of organisms.
>
> Now this decision is not made on the grounds that taxonomic groups 'really are' sets of organisms; it is made solely for the reason that by making it we can bring statements about taxonomic groups within the scope of the systematized and precise idioms of set theory, with all the advantages this carries. Those who wish to put forward other interpretations of taxonomic group names are, of course, free to do so . . . .

According to Gregg, great set-theoretic advantages accrue from treating taxa as sets, but of course it does not follow that they "really are" sets of organisms. Although Gregg does not give any procedure for defining taxonomic names, it is nevertheless clear that he thinks of taxa as sets whose constitution would be given solely by some set-theoretic sophistication on simple enumeration. The names of such sets are defined extensionally rather than intensionally. What this means basically is that a taxon is thought of as simply the collection of its members. No question is raised concerning any *qualifications* for membership. There is no suggestion that some set of properties might be such that any organism possessing them would have to be a member of the taxon, or that some other set of properties of an organism would with equal necessity preclude membership. Gregg makes this decision even though in an earlier article (Gregg, 1950) he made the following statement:

> Simply by way of convention, let us agree to call a statement of the properties which an organism x must have in order to be a member of a particular taxonomic group A, the *group-description*; for example, a list of properties which an organism must have in order to belong to the genus *Daphnia*, is a group-description.

Gregg's decision to treat taxa as sets of organisms and to define them by enumerating their members is unfortunate in two respects. First, it conflicts with all present and past taxonomic theory and practice. Second, and more importantly from his point of view, no enumerative definition of the name of any actual taxon could be given, for the simple reason that we are unable to enumerate all the members of any taxon. Part of what is involved in definition by enumeration can be seen in the following schema:

$$F =_{Df} \text{the class K such that } (a \in K) \text{ and } (b \in K) \text{ and } \ldots (n \in K).$$

Here, as before, the lower case letters are the names of individual organisms, the capital letters are the names of classes of such organisms, and the dots represent our inability to list the names of all the individuals who are members of this class. (Df is read "by definition.") For example,

$$\textit{Homo sapiens} =_{Df} \text{the class K such that (Gregg} \in \text{K) and (Hull} \in \text{K) and (Buck} \in \text{K) and} \ldots \text{and (Beethoven} \in \text{K).}$$

This schema seems initially satisfactory since it shows what an actual enumerative specification of the class would be like. The feeling of satisfaction is illusory, however, in that it rests on the understanding that the dots are to be replaced by additional ∈-statements whose left-hand side will be the names of other entities *of the same kind* as those already listed. But here lies the illusion! Of what kind? No *kind* has been mentioned. Of course, according to taxonomic procedure, the answer to this question is the kind of entity which has properties similar to those of the four individuals already mentioned, but our enumerative definition does not and cannot *say* this. When a class is given by enumeration, membership in that class promises no features or properties of the member. Our enumeration tells us nothing at all of properties common to the mentioned entities. Their only stipulated common property is their joint membership in a class called *Homo sapiens*. And no qualifications for such membership can be stated or even understood if the class is truly given by enumeration.

This is the key point. All there is to an enumeratively given class is a list of its members. The names of its members must be *proper* names, in the sense that no connotations about something's properties are to be thought of as inferable from the names. Fido and Gargantua would have been as suitably mentioned in the list as were Gregg and Buck. When there is no basis for classification, no qualifications for membership, one entity's name is as good or as bad as that of any other in defining a class name. One of the reasons why Gregg says so little about the relation of individual organisms to taxa may easily be that with enumeratively defined taxa names there is virtually nothing to say. Gregg's Paradox flows directly from this situation. Two classes which have the same membership are necessarily identical. They are "really" not two classes but one.

Although the preceding would be sufficient to eliminate enumerative definitions as a candidate for the definition of taxonomic names, there are still other problems remaining. For example, is an extinct taxon to be considered memberless? If so, then all extinct taxa are identical, having the same membership, namely zero.[5] On the other hand, if membership in a taxon is to be cumulative, it might be objected that the definitions of enumeratively defined taxa would change with each new organism born. This objection is tricky. It is true that, say, *Canis familiaris* would *change* if the name of a newly born pup were added to its definition. But here again, we are thinking of this species name being defined via properties. If "*Canis familiaris*" is actually defined only via some list of named members, then there would be no more reason to add a newly born pup to its extension than a newly born centipede or cockroach. Any one of these additions would change the meaning of "*Canis familiaris.*" The enormity of such extensional definition is best brought out by noticing that the latter two changes would be no greater than the former.

[5] The point of our earlier provision for classes having zero members lies partly here. An extinct species must be both the sort of thing that has members (a class), and it must also now have no living members.

Taxa Names: Intensional Definition

How are names of taxa in fact defined in taxonomy? As Gregg mentioned in his earlier article, they are defined intensionally; i.e., via sets of properties which supply the qualifications for membership. *Didus ineptus* and *Tyrannosaurus rex* are not identical taxa, because the qualifications for membership in the two taxa are quite different. Traditional taxonomic theory and practice have treated each taxon as in principle definable by a single set of properties of organisms which qualifies some organisms for membership and excludes all others. Possession of these properties was held to provide both *necessary* and *sufficient conditions* for membership. More precisely, possession of each of the several properties has been held to be necessary and possession of all of them sufficient for membership. The Linnaean hierarchy was constructed on this (unfortunately untenable) view of the definition of taxonomic names. Beginning with the work of Michel Adanson (1707–1788), taxonomists have slowly come to realize that this view is overly simple, whether or not they agree with some of Adanson's other views; e.g., those concerning weighting. Seldom if ever can a precise set of necessary and sufficient property qualifications be provided for any taxon. What has been needed is a method of definition (or specification of meaning) according to which no property is unconditionally necessary and no one particular set of properties is unconditionally sufficient. Some progress has been made in this direction, working with the logic of what are called "clusters" or "empirical classes."[6]

[6] The work in this area has been divided between two schools of taxonomy. Numerical tax-

Although the lessening of stringency with respect to defining properties complicates the explication of the definition of taxonomic names and, hence, the logical structure of the Linnaean hierarchy, one point is clear. It will not involve abandoning the idea that taxonomic names are defined by the properties of organisms. For the sake of simplicity and because it represents the logical picture which underlies the Linnaean hierarchy, we shall operate with the traditional view. Our purpose is to contrast an intensional, property-oriented view of the definition of taxonomic names with Gregg's definition by enumeration. For that purpose no distortion is involved in using the traditional approach.

It is the definitions of the names of taxa in terms of their properties that determine which taxa stand in inclusion relations to each other. In some areas of discourse, inclusion relations must be called accidental. For example, if in a particular elementary school all red-haired boys were fifth graders, then in these circumstances and for that school it just so happens that

Red-haired boys $\subset$ fifth graders.

It seem highly probable, however, that this inclusion relation is not grounded in any features or properties of the classes. We would not want to say that if someone *were* a red-haired boy, he *would* be a fifth grader. In taxonomy the situation is otherwise. The inclusion relation is vital in the logic of taxonomy. Indeed, the idea of a hierarchy rests on the notion of higher and lower taxa ordered by the logical notions of the extension and the intension of class names. The extension of a class name is simply the set of members of the class; the intension is the set of properties which qualifies an entity for membership. Ordinary classes and all taxa are intensionally defined, which means that the intension of the class name determines its extension.

What determines which taxa stand in inclusion relations to each other is the distribution of properties among the organisms being classified, as reflected in the definitions of taxonomic names. To say that Vertebrata is included in Chordata is to say that the properties required for classifying an organism as a vertebrate include (besides others) all the properties required to classify it as a chordate. The lower taxon, having *fewer* individuals as members, requires *more* properties to meet the qualifications for membership. The higher taxon, having *more* individuals as members, requires *fewer* properties to meet its qualifications for membership. This situation is a standard exemplification of a frequently enunciated logical principle, which asserts the inverse variation of extension and intension. Reducing the number of qualifications increases membership; increasing the number of qualifications decreases membership. Vertebrata can be regarded as a class generated by tightening up the membership requirements for Chordata, or Chordata can be regarded as a class generated by relaxing the membership requirements for vertebrate status. Vertebrata is included in Chordata solely on the basis of the membership requirements for the two taxa, without requiring an enumeration of the members to establish the fact.

Of course, the practice of indentation used in the Linnaean hierarchy obviates the need for a repetition of the properties used to define higher taxa in the definitions of the taxa included in them. For example, Mayr (1942) writes:

> On the other hand . . . the description of species in well-known groups should not be burdened with superfluous information. The descripition of a new species of *Drosophila* does not and should not contain a reference to any character

onomists following Adanson treat all properties as of equal weight (Sokal and Sneath, 1963). The phylogeneticists treat properties in the same manner as the numerical taxonomists with the exception that the properties are assigned various phyletic weights according to their presumed phyletic significance. For a formal treatment of the subject of weighting see Kaplan (1946), Kaplan and Schott (1951) and Schutz (1959).

> that is common to all known species of the genus. One of the principal advantages of the binary system is that it permits the omission from the description of all characters common to the genus, family, and higher categories [taxa].

But even though the properties used to define the name of a higher taxon may not be repeated in the definition of the name of one of its included lower taxa, these properties are still necessary for membership in the lower taxon.

The point of the preceding discussion is that the inclusion of one taxon in another is determined by the definitions of the names of these taxa based on the distribution of the properties of the organisms being classified. Thus, *within a single inclusion series*, taxa can be termed higher or lower relative to each other using *just* the definitions of their names. But most taxa do *not* stand in any inclusion relation to each other. They are disjoint. For example, *Canis familiaris*, Hominidae and Apidae stand in no inclusion relations to one another, and yet *Canis familiaris* is lower than Hominidae and Apidae, and the latter two are of equal rank. To make the higher and lower orderings within a single inclusion series determinate and to fix the ranks of disjoint taxa, reference must be made to the definitions of the names of categories. Disjoint taxa are termed higher and lower with respect to one another not because they stand in any inclusion relations but because they are members of different categories.

Until now everything that has been said about the logic of taxonomy could be said with no reference to certain divisive factors in taxonomy. But when we turn our attention to the definitions of category names and ask why some are lower or higher than others, any further development must take into account conflicting opinions as to the purposes of taxonomy. The nature of category definitions has given rise to more confusion than any other aspect of taxonomy. Some problems are understandable, since the idea of categories as classes of classes of individual organisms is far from easy to grasp, but some are not. Taxonomists have unduly complicated their task by failing to distinguish carefully between the types of properties which can be used to define the names of taxa and those which can be used to define the names of categories. Both taxa and categories are defined in terms of properties which their members have. Category names are defined in terms of properties which taxa have, and taxa are defined in terms of properties which individual organisms have. Since taxa and individuals are of different logical types, their properties are of different logical types.

The overwhelming number of properties with which taxonomists are concerned are attributable only to entities at level N^1—to individual organisms. Individual organisms and only individual organisms are fur-bearing, have mammary glands, interbreed, live in water, nest in trees, etc. It makes sense to say that Gargantua has hair. It makes *no* sense to say that the taxon Primates has hair. Instead, one would say that primates have hair. Similarly, having mammary glands may be a defining property of the taxon Mammalia, but the taxon Mammalia does not have mammary glands. Only mammals have mammary glands. Here the logical distinction between attributing properties to classes and attributing them to individuals is reflected in ordinary taxonomic discourse.

It must not be assumed, however, that all properties are appropriate to N^1 entities. For example, a species such as *Didus ineptus* may be extinct. It would make no sense to say of any individual bird, dead or alive, that it was extinct. Other properties which taxa may have but which individuals can neither have nor lack are, for example, being numerous or being distributed throughout North America. Problems arise, however, since the distinction between properties of individuals and properties of classes of individuals is not always recognized in ordinary discourse. The same is true of taxonomic discourse. For example, it is often said that two individuals are overall phenetically similar and that two populations or taxa are overall phenetically similar. Further, it is often said that two individuals interbreed and that two populations interbreed. In both cases, the same word or phrase is being used to refer to properties which are logically quite different.

When it is said that two *individuals* are overall phenetically similar, what is meant is that they share a certain *percentage* of their phenetic properties. When it is said that two *classes* of organisms are overall phenetically similar, what is meant is that a certain *percentage* of the members of each share a certain *percentage* of their phenetic properties. Two dimensions of variation are present. Similarly, when it is sàid that two individuals interbreed what is meant is that two individuals have united sexually with the production of offspring. When it is said that two classes interbeed, what is meant is that a certain *percentage* of their appropriate members have interbred. Interbreeding is a property of individuals. Percentage interbreeding as here explained is a property of classes. The logical confusion in these examples is much like that obvious in the claim that the average American wife has 2.27 babies. Similar confusions in taxonomy are not always so obvious. (For further details, see Hull, 1966.)

Now that the nature of the intensional definition of the names of taxa and categories has been explained, we turn to the solution of Gregg's Paradox. The formal exposition of the solution requires a new set of symbols to stand for properties of organisms. We shall use capital Greek letters for this purpose. Thus Φ might stand for possession of a notochord, Ψ for having an opposable thumb, and Γ for having hair. Taxonomists' definitions of taxonomic names make essential reference to such properties. Thus, if Φ, Ψ and Γ are among the defining properties of the taxon A, it will follow that for any organism x, x is not and could not be a member of A unless it has Φ, Ψ and Γ. And if Φ, Ψ and Γ are sufficient for being an A, it follows that any organism which has Φ, Ψ and Γ must be an A.[7] Thus,

$$x \in A =_{Df} [(x \text{ has } \Phi) \text{ and } (x \text{ has } \Psi) \text{ and } (x \text{ has } \Gamma)].$$

Two taxonomically desirable consequences are immediately apparent. First, the question whether some as yet unclassified organism is a member of A becomes sensible, and there is an empirical procedure for answering it. That procedure is, of course, the traditional one of examining the organism to see if it has properties Φ, Ψ and Γ. This examination will issue in a decision and that decision will be *grounded in* the qualifications for membership in A. Finding out that another individual is a member of A will in no way change the definition of A. This procedure contrasts strikingly with the already noticed case for an enumeratively defined taxon. For such a taxon, the fact that an organism is as yet not classified guarantees that it is not (now) a member of A. While a decision can be made to add it to A's membership, such a decision is based on *no grounds whatever*, and the meaning of A will be changed accordingly.

The second desirable consequence concerns identity of taxa. With A defined as above, it is clear that no higher or lower taxon can be identical with A unless it has exactly the same property qualifications for membership. Even if two taxa happen to have the same members, if they are defined in terms of *different* properties, then they are *different* taxa. They would be extensionally identical but intensionally different. The solution to Gregg's Paradox is now obvious. There are property qualifications for membership in Cetacea which are not also qualifications for membership in Mutica. Thus, while it is true that all members of

[7] These and the following formulations could be expressed entirely in set-theoretical terms by interpreting predication as a membership claim. Then, x has hair would be translated to read

$$x \in \text{the class of hairy organisms.}$$

Mutica are members of Cetacea, it is not true that if any organism *were* a member of Mutica, it would be a member of Cetacea. A marine animal might be found which would be classed in Mutica but not in Cetacea. As long as this is the case, the practice of monotypic classification need not be legislated against since no spurious identification of Cetacea and Mutica is implied and no overlap between the categories cohort and order threatens.

According to our view, a taxon B includes a taxon A if and only if all property requirements for membership in B are also required for membership in A. Using Σ as a property variable, ranging over Φ, Ψ, Γ, etc., a more formal version of our definition is

$$A \subset B =_{Df} (\Sigma)(x)[\{(x \in B) \text{ entails } (x \text{ has} \Sigma)\} \supset \{(x \in A) \text{ entails } (x \text{ has } \Sigma)\}].$$

Taxa A and B are identical only if this inclusion relation holds in both directions: from A to B and from B to A. Since this is not the case with Mutica and Cetacea, these two taxa are not identical and are properly classed by the customary taxonomic notation of indentation, as follows:

Mutica
    Cetacea

It will thus be seen that we can accept Gregg's definition of identity for taxa as joint inclusion and still reject his paradox.[8]

Conclusion

Our aim has been to expound informally a number of important points in the logic of taxonomy, many of which are derivative from Gregg's more formal presentation. Following Gregg, we have argued that clarity about the membership relation, the inclusion relation, and the difference between them is of extreme importance in the logic of taxonomy. Of equal importance is the difference between the types of properties attributable to individuals, taxa and categories. We follow Gregg also in arguing that the relation of organism to taxon and of taxon to category is that of membership, whereas the relation between a taxon and another that falls under it is the very different inclusion relation. Departing from Gregg, we have substituted intensional definitions of names of taxa, of inclusion and of identity for his enumerative extensional approach. Our definitional procedure is surely closer to that actually used in taxonomy. It enables us to avoid his paradox and to dissociate ourselves from his consequent strictures against monotypic classification. We nevertheless believe that our presentation preserves and is fully compatible with many important logical truths of taxonomy which he has set forth.

REFERENCES

EDEN, M. 1955. *Review of*: The language of taxonomy. Amer. Scientist 43:128a–130a.

GREGG, J. R. 1950. Taxonomy, language and reality. Amer. Nat. 84:419–435.

[8] Several attempts have been made to avoid Gregg's Paradox without introducing intensional definitions of the names of taxa. For example, if lower taxa are considered to be related to higher taxa by ∈ instead of ⊂, as Parker-Rhodes (1957) suggests, then monotypic classification causes no problems, but statements such as "primates are vertebrates" become nonsensical. One paradox is replaced by a host of others. Sklar (1964) advocates the introduction of an index-object for each category level. Each species would have as its members numerous individual organisms and the species index-object, each genus numerous individual organisms and both the genus and the species index-objects, and so on. Sklar admits that this method of avoiding Gregg's Paradox is *ad hoc*, but, he contends, so is much of biological classification itself. Van Valen (1964) avoids this paradox by an attempt which has overtones of the intensional method. The irregularities in the taxonomic hierarchy are first removed by the classification of all entities at every possible category level—increasing the number of monotypic taxa and the extent of monotypic classification tremendously. The logical difficulties which arise from such a procedure are then removed by the introduction of conditionally empty sets, such that every taxon has at least two taxa included in it, one of them being conditionally empty. Both of these latter workers, as well as Eden (1955), realize that intensional definitions would solve Gregg's Paradox but fear that such a treatment would be prohibitively intricate. We would like to think that our sense-perserving, intensional reconstruction has shown this fear to be groundless.

1954. The language of taxonomy. Columbia Univ. Press, New York.
HULL, D. L. 1966. Taxa, categories and definition. In preparation.
KAPLAN, A. 1946. Definition and specification of meaning. Jour. of Philos. 43:281–288.
KAPLAN, A. and H. F. SCHOTT. 1951. A calculus for empirical classes. Methodos 3:165–190.
MAYR, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York.
PARKER-RHODES, A. F. 1957. *Review of*: The language of taxonomy. Philos. Rev. 66:124–125.
SCHUTZ, W. C. 1959. Some implications of the logical calculus for empirical classes. Psychometrika 24:69–87.
SIMPSON, G. G. 1945. The principles of classification and a classification of mammals. Bull. Amer. Mus. Nat. Hist. 85:1–350.
1961. Principles of animal taxonomy. Columbia Univ. Press, New York.
SKLAR, A. 1964. On category overlapping in taxonomy. *In* J. R. Gregg and F. T. C. Harris [eds.], Form and strategy in science, p. 395–401. D. Reidel Publ. Co., Dordrecht-Holland.
SOKAL, R. S. and P. H. A. SNEATH. 1963. Principles of numerical taxonomy. W. H. Freeman and Company, San Francisco.
VAN VALEN, L. 1964. An analysis of some taxonomic concepts. *In* J. R. Gregg and F. T. C. Harris [eds.], Form and strategy in science, p. 402–415. D. Reidel Publ. Co., Dordrecht-Holland.
WOODGER, J. H. 1937. The axiomatic method in biology. Cambridge University Press, Cambridge.
1948. Biological principles. Routledge and Kegan Paul, London.
1952. Biology and language. Cambridge University Press, Cambridge.

ROGER C. BUCK is Chairman of the Department of History and Philosophy of Science, Indiana University, Bloomington, Indiana. **DAVID L. HULL** is an Assistant Professor of Philosophy, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin. Professor Hull wishes to express his appreciation for the support he received from the Penrose Fund of the American Philosophical Society and Research Funds of the University of Wisconsin-Milwaukee.

Notice of 1966 Meeting

The annual meeting of the Society of Systematic Zoology will be held in conjunction with the annual national meeting of the American Association for the Advancement of Science, in Washington, D. C., 26–31 December 1966. Headquarters for the AAAS will be the Sheraton-Park Hotel, but some societies meeting with the AAAS will be headquartered at the Shoreham and Washington Hilton hotels. Further details concerning the SSZ meeting will appear in the September issue of *Systematic Zoology.*