Part A. Problem of effectiveness

1

Introduction

1.1 ***One overall purpose.*** The single aim of this presentation is to make possible the integrated use of every detail of available geologic information taken from rocks in order to achieve better resolution in sequence correlation, in paleoecologic interpretation and in logging the course of evolution. At present much paleontologic and other descriptive work is entirely compartmentalised and is usually embarked upon without closely considering such application of the information.

As indicated on the List of Contents, the first part of the book is devoted to an analysis of the problem, followed by a second part of six chapters proposing a complete revision of paleontologic data-handling in preparation for more profitable computing in semi-numerical languages. The third part, starting at Chapter 9, deals with the information-handling resulting from interpretation of the data; this is designed as a preparation of non-numerical computer manipulation through artificial-intelligence languages. The final part contains summaries and is followed by a glossary. This book is not concerned with computer programs of any kind but solely with the logical arrangements for potential input.

1.2 ***Terminology.*** The introduction of new ideas inevitably generates some new terminology so that fresh concepts may start life without confusion. It is also sometimes necessary to discourage use of certain older terms which may prove to be poorly defined or may obstruct. To reduce difficulty caused by these introductions and changes, a comprehensive glossary of all such terms is included at the end of this book.

1.3 ***The material.*** The discussion is based on fossils, for which geological phenomena the details are usually the most rich and complex in characters, thus making their evolution the more detectable; the methods are intended to be equally applicable to other geological phenomena as interpreted from rock, although that intention is not developed here. The discussion is particularly concerned with microfossils because they have often hitherto been inappropriately and wastefully treated as if they were megafossils. From the descriptive and character recording point of view there is no essential difference between any two kinds of fossil, but in numbers of specimens there is great disparity. With megafossils the problem is almost always to collect enough undamaged specimens to make a useful description with a significant record of variation. With microfossils there is almost always no excuse for failing to collect enough specimens to express variation; but the resulting sampling is so restricted and local in contrast with the total numbers of specimens available, that the method of selection becomes all-important in deciding how representative the selection from a sample may be (Gluzbar 1987).

Further, each set of specimens is by custom normally consigned to one specialist 'expert' who frequently neither investigates nor allows for parallel work in adjacent fields; it is difficult therefore for individual paleontologists to plan for integration of results from the many different types of fossil.

1.4 ***'Economics' of paleontologic exploratory work.*** It is perhaps reasonable to attempt to establish the general pattern of purpose behind the frequently laborious handling of microfossils; it has been argued that many of the specimens now studied could as profitably stay unattended either in or on the ground until they happen to be specifically required. On the other hand, because of the very slowly acquirable expertise that is needed for study, it becomes economically sensible to build up and hold information on fossils (and on most other characters of rocks) at a general state of readiness from which rapid expansion of knowledge to meet new demands becomes feasible. Historically, a balance of 'preparedness' has been built up in the world between institutions and universities, and among amateurs, and this appears to work satisfactorily in the present environment. It seems possible, however, that with greater efficiency of data-handling, and consequently more interchangeability of knowledge among paleontologic 'experts', some economies might be possible. Amateur paleontologists, for example, could be positively encouraged to undertake the widespread production of new observation records to revised higher standards. These could then be used in much more effective exploitation work than is currently feasible.

1.5 ***Obstacles of custom.*** In order to make possible the simultaneous advancement in many languages and in many countries of complex studies of interpretations of fossils and of the applications of the results, some universal customs of procedure are necessary for all to follow; these customs have been elaborately developed for more than a hundred years but the problem is the subsequent difficulty of updating them, once they are in place.

The central difficulty is the current 'cluster' approach to selection and employment of material from seemingly very great variety, and the rough approximation type of result to which it leads. Species and other taxa of fossils are clusters of specimens and in general biozones are clusters of information. Parsimony in erecting taxa, known as 'lumping', is a cautious habit thought useful for assisting the human mind in attempting to encompass the scope of nature.

Secondly, most paleontologists believe very generally in 'biologic evolution' but are also for the most part, because of their individual inability to provide any convincing proof of it, uncertain how to proceed in addressing it. In seeking proof an analytical approach is necessary, and this may be correlated with discrimination and a 'splitting' approach to the erection of taxa.

There is even a philosophical argument for splitting and against lumping, regardless of subject matter. If splitting or lumping were each successfully achieved, there is no case for or against either, but if either process be flawed then: (a) the lumped descriptive data will need to be re-observed to make the necessary discriminations, and (b) the split descriptive data may be lumped in an automatic synonymy. Thus splitting is more efficient if it is not too laborious, which is unlikely when computer assistance is available with handling of the results.

Despite the availability for some time now of rapid data-handling procedures which make 'splitting' a much more feasible approach, custom has had the effect of holding many workers back in the 'cluster and lumping' age. This has maintained a pointless rather feeble continuing exploration instead of the effective development and exploitation which the present state of earth science demands.

1.6 ***Scheme to fit the purpose.*** The Paleontologic Data-Handling proposal made here aims to pay more attention to the making of fuller re-usable records of each individual occurrence of fossils than to the aggregation of these records by 'lumping' into taxa, which are entirely human constructions and so much more liable to be flawed. Correspondingly in stratigraphy more attention goes into recording events than into their 'lumping' into zones. In neither case, however, is there any intention to jettison the taxa or the zones which are naturally the important elements of continuity and compatibility with published work of the past. The purpose is to render taxa and zones more universally useful by making it possible to reach back behind them effectively for the basic data when a new display of such data can lead to a new result.

Some of the detailed suggestions are not new but their co-ordinated presentation is. Some details of the package may appear on their own to be trivial and less urgent for consideration, but all are necessary for consistency. The scheme is addressed in particular to newly qualified paleontologists and stratigraphers who may by using it make much greater use of their inheritance, or who may be led to make better and more far-reaching proposals with this same objective in mind. Some caution may however be necessary with more established scientists who may also happen to instruct, edit or employ, and who may need more persuasion that serious change is needed.

2

Current data-handling for fossils

2.1 ***Outline of current procedures.*** The first serious observer of a new kind of fossil names a species from the best group of specimens that can be assembled; his purpose in so doing is to classify natural observations sufficiently to be able to recognise and memorise occurrence that may prove to be usefully similar, and thus to aid the scientific community. In erecting a species an illustration, description and measurements are given together with locality details, and distinction is drawn between the new and any adjacent species which happen to have been already published. A single nomenclatural holotype is selected, and a statement on variation of specimens is provided if possible.

Following nomenclature rules, the species is attributed to a genus, and whenever possible is fully classified among all other organisms, living and fossil. With most authors there is a tendency to parsimony is erecting new species, so that attribution of the new specimens to an existing published species is always considered first.

In nomenclature the species epithet is normally stable because of the existence of the holotype, but the generic element of the species name is subject to change by any subsequent author with new ideas on classification. The generic element of the name is placed before the specific epithet and so is automatically indexed and spoken ahead of it; consequently retrieval of data is frequently complicated by generic name changes.

Additional specimens from the type locality or from anywhere else in space or time are 'identified' with the species and are usually list-recorded under the same name. These specimens are then usually taken by most workers into the compass of understanding of the species, which thereby becomes a growing cluster of specimens. A range of occurrence is arrived at empirically, and when stabilised is ultimately taken to be a character of the species. Uncertain attributions to the species are prefixed 'cf.', but as discussed later there can be no precision in such statements.

The imprecise nature of the descriptive extent of a species or other taxon is widely accepted as unavoidable, but is the cause of failure to achieve more effective applications of the use of species of fossils. Interpretations of more than one species in phylogenetic studies only lead further away from any easy return to the original observations.

Although constraints on the interpretation of 'species' of fossils have long been understood (see e.g. Newell 1956), very little attempt has been made by the majority of paleontologists to employ this knowledge effectively.

2.2 ***International Codes.*** The International Code of Botanical Nomenclature (ICBN; Voss *et al.* 1983) and the International Code of Zoological Nomenclature (ICZN; Ride *et al.* 1985) are two elegant and very carefully prepared legal documents both maintained with the prime purpose of stabilising names of taxa of living organisms in order to reduce or to avoid confusion. They work well, although they are cumbersome; they are subject to continual adjustment of minutiae, and to long lists of *nomina conservanda* for cases in which full application of the rules is deemed to be unhelpful.

2.3 ***Nomenclature for fossils.*** Perhaps a majority of paleontologic colleagues and many editors have faith in the belief that strict adherence to these two Nomenclature Codes amounts to an essential requirement for orderly progress in both paleontologic and stratigraphic studies. The case, however, for treating geologic material in this way is seldom, if ever, argued.

2.4 ***Taxonomy for fossils.*** Taxonomy, as distinct from nomenclature, is believed by many of the same people to be a matter for free expression of opinion by authors. It does not seem to be widely appreciated that although nomenclature in the form of these

Codes has little influence on extant taxonomy of extant organisms, it has the effect of governing taxonomic procedure for fossils absolutely (see Section 6.2 below). This is mainly because of the unstated but important differences of nature of the data between living and fossil material, which will be discussed in the next chapter.

2.5 ***Control through the Codes.*** The Botanical Code is controlled through independent nomenclature sessions at each International Botanical Congress (currently Sydney 1981 and Berlin 1987). Paleobotanists originally, up to 1959 (Montreal), kept their affairs separate from those of the workers on living plants in a distinct Code 'Appendix for Fossil Plants', which recognised organ-parataxa. However, they have since, on an egalitarian wave of fashion, allowed this useful distinction to lapse; as a consequence paleobotanists are now so heavily outnumbered at meetings concerning the Code that effective thought about separate attention to problems of fossil plants has withered.

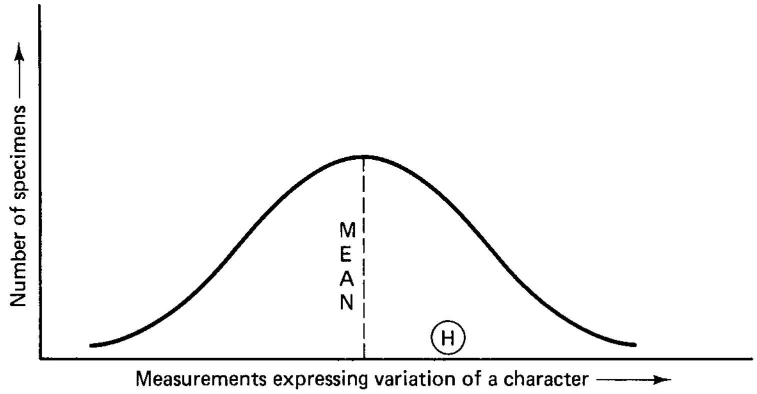
The Zoological Code, because the appropriate International Congresses were discontinued some years ago, has had a committee (with relatively stronger paleontological representation) which has patiently given rulings on problems over the years; because with the principal exceptions of fossil vertebrates and insects, the need there for recognition of organ-parataxa was much less than with fossil plants, treatment of fossils has remained closely similar to that for living organisms.

Thus, from different circumstances and reasoning, unquestioned orthodoxy prevails in both fields and various editorial attempts have been made to enforce it as an apparent simplification of procedure. As will be seen below, such nomenclature control dictates taxonomic style which inevitably and undesirably restricts enterprise in all data-handling for fossils.

2.6 ***Holotype concept for extant organisms.*** Because a species of living organisms is in constant growth and minor change, making all individual organisms transitory, a single preserved nomenclatural holotype is required for reference. This specimen remains as the 'legal' point of reference for the name in resolving disputes and uncertainties of identification, but it does not necessarily reflect the mean of variation of any character of the organism (see e.g. Fig. 2.1). Because the definition (description) of the living species remains verifiable by observation, such usage has proved in the past to be adequate and satisfactory. There is, however, the minor theoretical difficulty that the nomenclatural holotype was in most cases collected several generations ago and in the course of continuing evolution some change in the populations of a species will have occurred since that time; because such change will be very small in scope, although perhaps potentially increasing, it is by custom ignored.

2.7 ***Holotype concept for species of fossils.*** It is attractive to copy this satisfactory holotype procedure for use with fossils because in most cases the fossil specimens themselves are readily preservable and durable, although they are often very far from representative of any complete organism. There being no question of further change to such material, stability appears to reign.

Fig. 2.1. Diagram illustrating normal variation of a species of extant organisms, deemed to occur on a single Holocene time-plane. The nomenclatural holotype (H) is customarily a complete and well-developed specimen selected by eye; it is not related to any detail of character variation, and could in theory contribute to any part of the curve but is usually a little larger or more prominent than the mean.



The Nomenclature Codes provide understandably (but unhelpfully to a paleontologist believing in evolution) that when fossil and living specimens appear under the same name, the holotype and name shall be taken from the living or recently dead (ICBN, Art. 58; ICZN, Art. 20, in a less comprehensive sense). There is also the problem (Fig. 2.2) with fossils representing organisms changing through time, that a single morphotype may represent different parts of the variation of successive 'species' in conformable strata, and so mislead when only small numbers of specimens are available for comparison. Such would even be the case if Philip and Watson (1987) proved to be right in their warning about the normal distribution.

2.8 ***Fossil 'species' as clusters.*** A species of fossils begins as a number of originally described specimens from a designated sample, from which one specimen is selected as a holotype;

Fig. 2.2. Diagram to illustrate how one kind of specimen (morphotype), as indicated here for simplicity by one character measurement (M), can be recognised at successive stratigraphic horizons although in each case it may represent a different part of the variation. This apparent constancy of occurrence could mislead in cases where numbers of specimens collected or studied had been insufficient to establish the true variation at each horizon.