THE PRODUCTION OF STRATIGRAPHICAL RANGE-DIAGRAMS BY AUTOMATIC METHODS

by IAN E. PENN

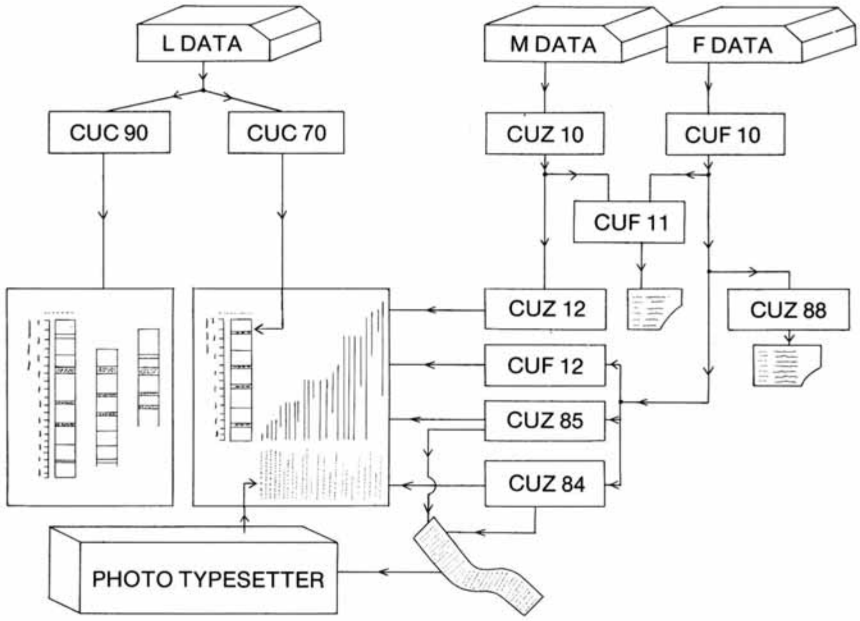
ABSTRACT. A package of computer programs has been developed on the Institute of Geological Sciences' IBM 1130 computer which, after accepting and sorting data submitted on a simple input format, draws stratigraphical range and abundance diagrams to any desired scale. The range-diagrams may be of first occurrence; last occurrence; taxonomic order; any specified order including sorting within sub-groups. An ornamented lithological plot is accurately aligned and a correctly ordered, punctuated and type-set list of fossil names may be produced from companion dictionaries. The resultant diagram is suitable for direct publication. It is estimated that, the programs having been installed, the computer method produces diagrams in one-tenth of the time and at one-quarter of the cost (within the Institute of Geological Sciences) of manual production.

IN his review of the familiar faunal range-diagram, Bursch categorized the commonly used kinds of diagram and quoted Moore's (1948, p. 324) point that the presentation of data, often painstakingly collected, in a single tabular manner may obscure the full significance of the results (Bursch 1950, pp. 479–484). This is probably because 'the *format* of the chart itself *dramatizes whichever viewpoint* is being used' (McLean 1972, p. 1). Though highly desirable, the production of several diagrams from the same data has never been fully practised perhaps because of the time-consuming nature of the work which is in itself not strictly palaeontological. Automatic data-processing methods are an obvious solution, for example McLean (1972, p. 5) devised a semi-automatic system using relatively cheap card-punch and card-sorting equipment.

By mid-1972 the Institute of Geological Sciences Computer Unit had written a package of computer programs in 1130 Fortran for the display of cartographic and lithostratigraphic data on the commonly used IBM 1130 computer configuration. It was then decided to write programs to produce the conventional faunal range-diagrams and link them with part of the larger package which drew ornamented lithological sections. At the same time dictionaries were established containing full fossil names and the appropriate type-setting instructions so that a copy of the species-list for a particular diagram could be automatically punched on paper tape to be fed into a phototypesetter which can produce, if required, a correctly type-set and punctuated faunal list. This list, together with the diagram, results in the automatic production of a faunal range-diagram suitable for publication.

The major features of this palaeontological program package are here outlined (text-fig. 1) while full program listings may be obtained on request. Full details of the more generalized central part of the package are in Farmer and Johnson (*in press*).

[Palaeontology, Vol. 17, Part 3, 1974, pp. 553–563.]



TEXT-FIG. 1. Flow chart of the various program packages. L DATA, M DATA, and F DATA represent the input of lithological mineralogical and fossil data respectively. The various identifiers within rectangles represent the various program decks. CUF 11 and CUZ 88 produce only line-printer output; CUZ 84 and 85 may produce paper tape.

DATA INPUT

Data input (text-fig. 2) is short and simple. Lithological data (L DATA in text-fig. 1) consists of: the thickness of each successive lithological unit followed by the appropriate two-letter code corresponding to the desired lithological ornament punched on successive cards. For example, text-fig. 2A shows the entries on successive cards of a lithological sequence of (in descending order) 25·12 m of argillaceous limestone; 4·23 m of mudstone; 0·12 m of oolitic limestone; 1·56 m of clay. The whole entry occupies no more than the first nine columns of a standard 80-column punch card. At present some twenty-five ornaments are available while a zero entry results in blank space which may be ornamented manually if desired.

The fossil data (F DATA in text-fig. 1) or the mineral data (M DATA) are also presented in stratigraphical succession. Thus the first card of each stratigraphical sample (text-fig. 2B and C) states the number of species or minerals in the sample together with the upper and lower depth limits of the sample; and each subsequent card corresponds to each of the species or minerals in these samples. The species or minerals are not recorded by name but are given a code number (occupying the first four columns of the punch card) which is followed, in the case of species, by a simple

(1–4) abundance code corresponding to 'present', 'fairly common', 'common', and 'very common' (text-fig. 2B). A zero is entered where the occurrence is doubtful. For example, text-fig. 2B shows the entries on successive cards of the first three species (code numbers, 4, 3, and 56) of twelve species determined from a depth range of 17·52–17·92 m, which are estimated as being present, very common and possibly occurring respectively.

The abundance code is the only difference between the fossil and mineral data input formats. This is because the mineral 'abundance' may be in the form of percentages (text-fig. 2C) and so one extra column is required on the data card. For example, text-fig. 2C shows the entries on successive cards of the three minerals (code numbers 1, 4, and 5) of a sample at 11·45 m depth which have been found by analysis to constitute 85%, 2%, and 10% of the rock. Percentage fossil data can, of course, be input on the M DATA format, while the replacement of abundance codes by equivalent, arbitrary percentage values would enable semiquantitative lithological and/or faunal data to be displayed alongside quantitative data. It is therefore possible to combine lithostratigraphical and biostratigraphical data on the same diagram. It will be seen then that the bulk of the data, whether in L DATA or M DATA format, occupies no more than the first seven columns of a standard data card.

Dictionaries giving the full fossil name and corresponding number code are permanently stored by accession order and by taxonomic order on punch cards (text-fig. 2D) as well as on a magnetic disc. It is simple to insert a new card when a name has to be changed or added to the list or to shuffle the deck of cards when the taxonomic order has to be changed.

A third kind of dictionary entry will be seen between the code number and the fossil name (text-fig. 2D). This is a letter code corresponding to convenient groupings, e.g. B = brachiopod, BR = rhynchonellacean brachiopod. This code enables the computer rapidly to break down large lists of species into meaningful sub-groups if desired.

In addition to the code number and full fossil name, type-setting symbols are inserted (text-fig. 2D) so that the output will be in italic or roman type and have upper

A *lithological data input*

```
12345678901234567890
  25.12AL
   4.23MU
   0.12OL
   1.56CL
```

B *fossil data input*

```
12345678901234567890
  12 17.52  17.92
   4 1
   3 4
  56 0
```

C *mineralogical data input*

```
12345678901234567890
   3 11.45  11.45
   1 85
   4  2
   5 10
```

D *dictionary input*

```
12345678901234567890123456789012345678901234567890123456789012345678901234567890
 224  BR£/ @R%HACTORHY NCHIA  OBSO LETA  #@( D% AVIDSON  /NON  #@S%OWERBY)
```

TEXT-FIG. 2A–D. Each row of data represents the left-hand side of one 80-column data card and shows the entries in their correct columns (as indicated by the italicized numbers) to produce the four different kinds of input.

MIDDLE JURASSIC MACROFOSSILS IN ACCESSION ORDER

0 br*Kallirhynchia sp.*
1 br*Kallirhynchia superba* S. S. Buckman
2 br*Rhynchonelloidella sp.*
3 br*Rhynchonelloidella smithi* (Davidson)
4 br*Rhynchonelloidella smithi crassa* Muir-Wood
5 br*Rhynchonelloidella wattonensis* Muir-Wood
6bts*Avonothyris sp. nov.* A
7bts*Tubithyris sp.*
8bts*Wattonithyris sp.*
9bts*Wattonithyris* cf. *fullonica* Muir-Wood
10bts*Wattonithyris* cf. *midfordensis* Muir-Wood
11btl*Ornithella sp.*
12btl*Ornithella bathonica* (Rollier)
13btl*Ornithella bathonica bathiensis* (Rollier)
14btl*Ornithella pupa* Muir-Wood
15btl*Rugitela sp.*
16btl*Rugitela cadomensis* (Deslongchamps)
17 at*Holzbergia sp.*
18 at*Holzbergia schwandorfensis* (Arkell)
19 at*Morrisiceras sp.*
20 at*Morrisiceras comma* (S. S. Buckman)
21 at*Morrisiceras krumbecki* Arkell
22 at*Morrisiceras morrisi* (Oppel)
23 at*Morrisiceras sphaera* S. S. Buckman
24 at*Morrisiceras fornicatum* (S. S. Buckman)
25 at*Morrisiceras skipnum* S. S. Buckman
26 at*Morrisiceras korustes* S. S. Buckman
27 az*Procerites sp.*
28 az*Procerites progracilis* Cox & Arkell
29 an*Procymatoceras baberi* (Morris & Lycett)
30 at*Tulites sp.*
31 at*Tulites subcontractus* (Morris & Lycett)

MIDDLE JURASSIC MACROFOSSILS IN MEMOIR ORDER

PLANTAE

428 h*Solenopora jurassica* Nicholson
427 h*Solenopora sp.*
563 h'*Solenopora*' *sp.*
426 halgae [frags]
430lgnwood [frags]
137lgnlignite [frags]

PORIFERA

423 s'*Peronidella pistilliformis*' Lamouroux
424 s'*Cliona*' *sp.*
425 ssponge [frags]

ANTHOZOA

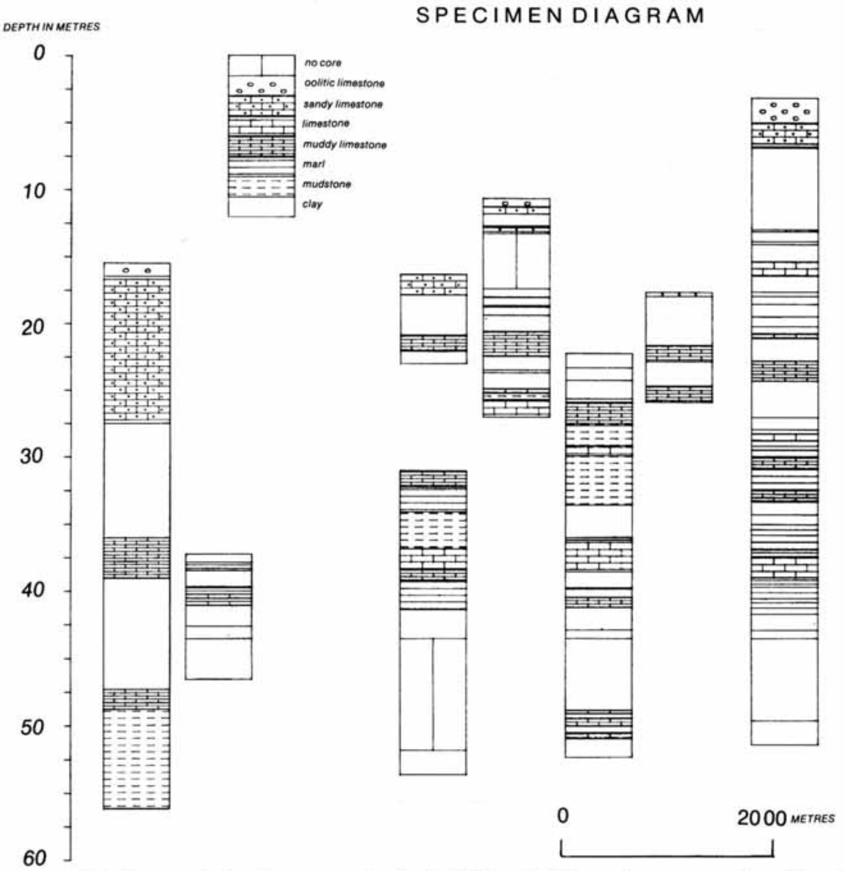
188 sc*Calamophyllia sp.*
176 sc*Chomatoseris hemisphericus* (Milne Edwards & Haime)
177 sc*Chomatoseris orbulites* (Lamouroux)
178 sc*Chomatoseris porpites* (Wm. Smith)
175 sc*Chomatoseris sp.*
181 sc*Cladophyllia babeana* (d'Orbigny)
180 sc*Cladophyllia sp.*
187 sc*Convexastrea sp.*
195 sc*Dimorpharaea defranciana* (Michelin)
199 sc*Ewardsomeandra vermicularis* Milne Edwards & Haime
198 sc*Ewardsomeandra sp.*
174 sc*Montlivaltia sp.*
196 sc*Thamnasteria neptuni* (d'Orbigny)
197 sc*Thamnasteria terquemi* Milne Edwards & Haime
194 sc*Thamnasteria sp.*

TEXT-FIG. 3. Output of fossil dictionary in accession order (to the left) and in taxonomic order (to the right) at 30.9.72. The list has been reduced in scale to the smallest possible size available.

or lower-case letters. The typesetter is fed by tape produced from the dictionaries held on the magnetic disc and produces correctly type-set and punctuated lists (text-fig. 3) which it can reduce (or enlarge if desired) so that a large list can be copied at a convenient size. The input characters for typesetter instruction have been carefully chosen so that in other forms of dictionary output (e.g. the fast lineprinter output from the computer or the tape-typewriter print of the paper-tape contents) they will not be printed. These other forms of dictionary lists are then much easier to read than they would be were the redundant type-setting instructions retained. For example, text-fig. 2D shows the dictionary entries for a species (code number 224) of a rhynchonellacean brachiopod (**BR**). The typesetting instructions for upper and lower-case @ and % precede the appropriate letters and the symbols # and / precede the text to be printed in roman or italic print. The symbol £ instructs the typesetter to delete all the entries to its left if required.

THE PROGRAMS AND THEIR OUTPUT

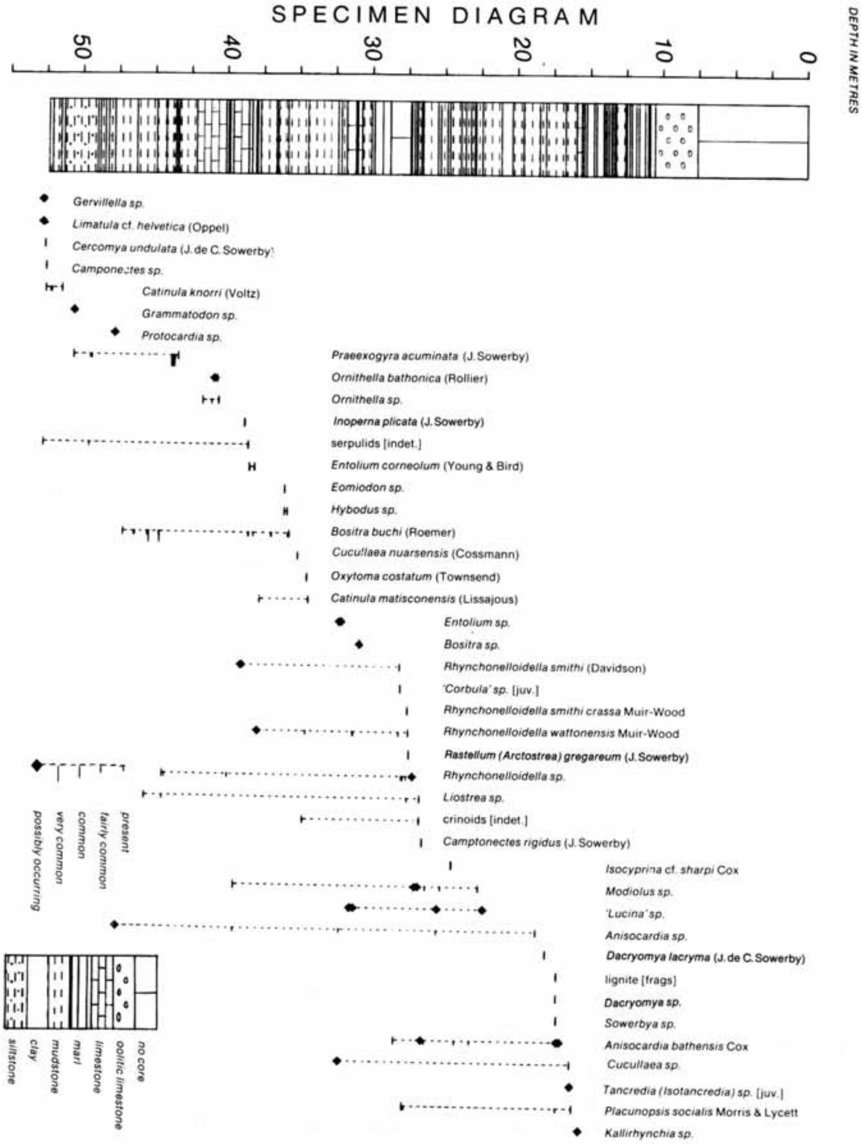
The programs available within the package are labelled (text-fig. 1) by their program identifiers. CUC 90 plots a series of lithological sections side by side in specified horizontal and vertical relative positions (text-fig. 4) thus providing a skeleton correlation diagram. CUC 70 plots individual lithological sections in the correct vertical position relative to plots of stratigraphical range data. Both programs provide



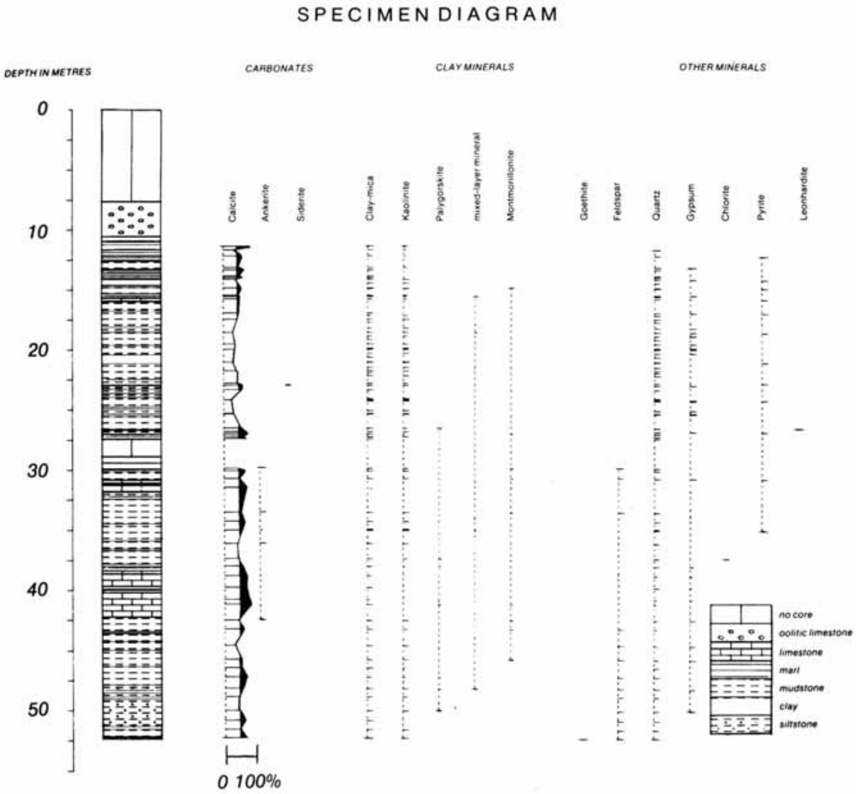
TEXT-FIG. 4. Skeleton correlation diagram as output by CUC 90 in which the sections are correctly positioned stratigraphically and geographically. Labels and tie-lines may be added manually.

a separate list of those horizons which are too thin to ornament at the chosen scale. CUZ 10 and CUF 10 read M DATA and F DATA respectively and store them on a magnetic disc for accession by the remaining programs. Thus CUF 11 prints out a check-list of the M DATA and F DATA stored by CUZ 10 and CUF 10, and CUZ 88 prints out, via the dictionaries, the full taxonomic name of all the different minerals or species stored by CUZ 10 and CUF 10.

CUZ 12 and CUF 12 deduce the stratigraphic range of each mineral or fossil stored by CUZ 10 and CUF 10 and plot range and abundance by first occurrence or last occurrence (text-fig. 5) or by either of those within designated sub-groups (text-fig. 6). It will be seen that the horizontal mark (which is proportional in length to species

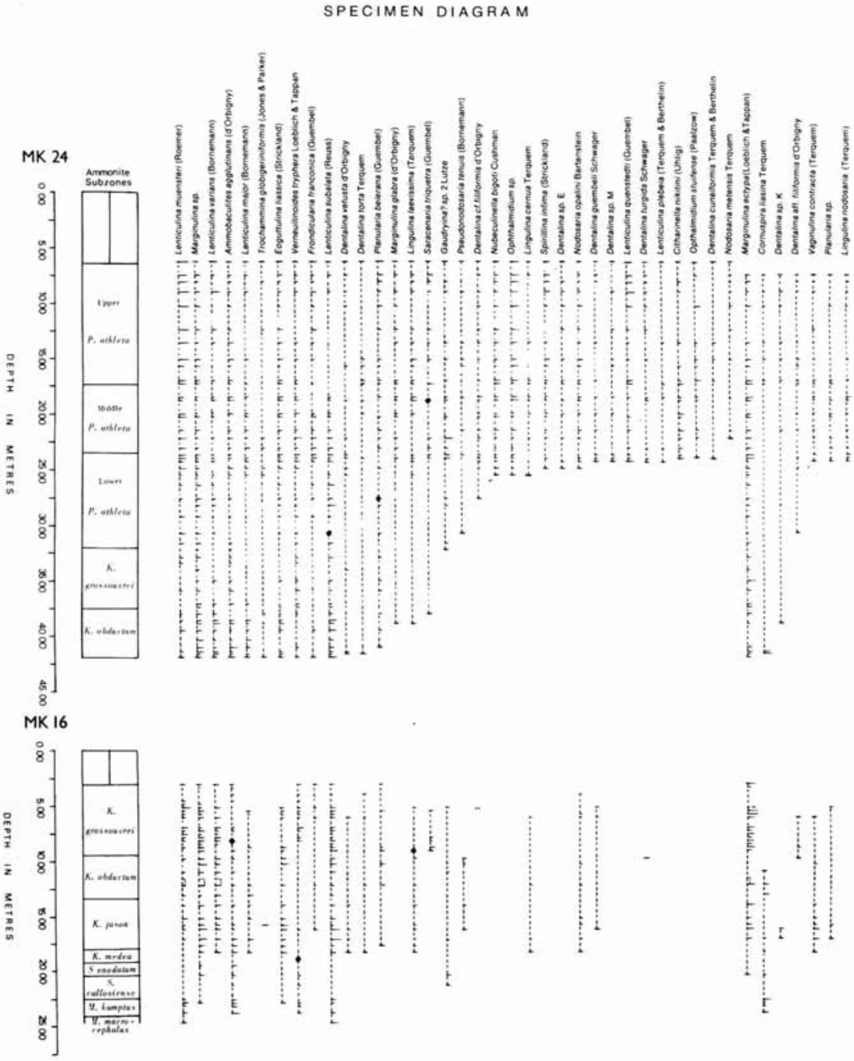


TEXT-FIG. 5. Specimen range-diagram as output by CUZ 85 of last occurrence of Middle Jurassic macrofossils from a borehole near Bath.



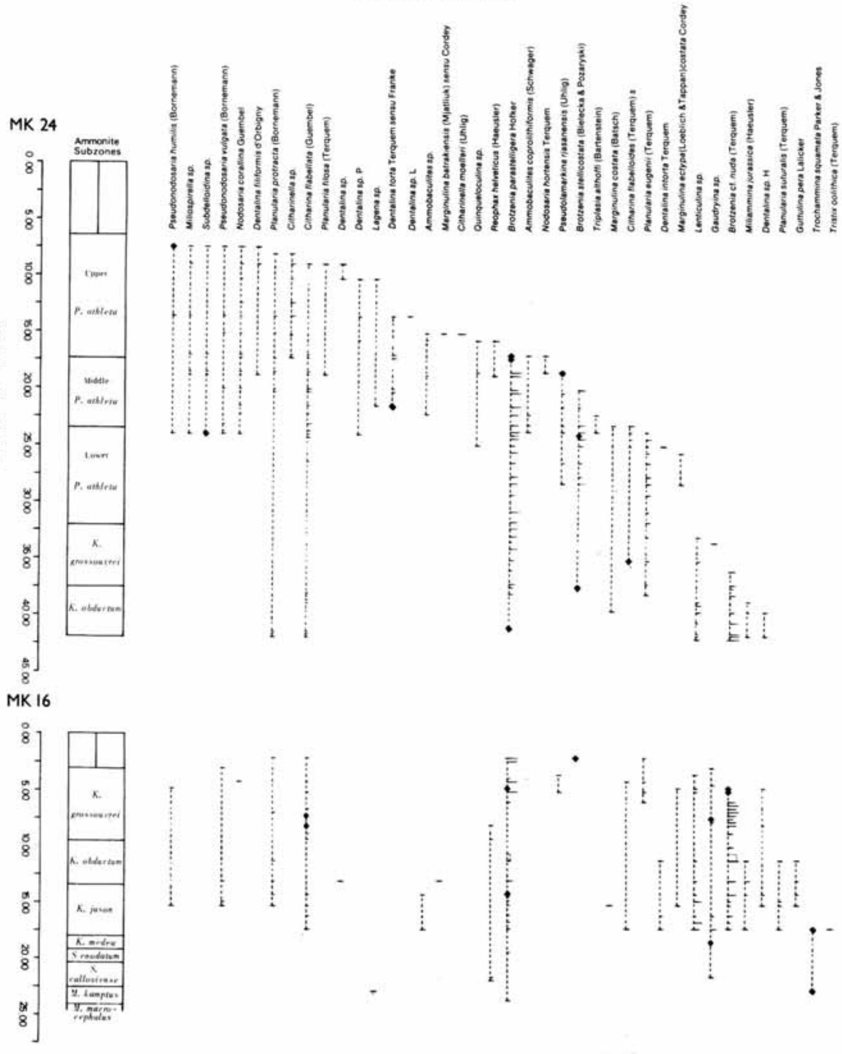
TEXT-FIG. 6. Specimen range-diagram as output by CUZ 12 showing the grouping of minerals. Calcite (shown manually ornamented above the 50% level) was determined quantitatively by wet analysis and this was used as an internal standard for calibrating the peak heights of the X-ray diffractograms. The analysis was not fully comprehensive.

abundance) enables every species occurrence to be localized. Where occurrence is from a range of depth (e.g. when fossils are recorded from 'Bed X') the horizontal mark becomes 'enlarged' to a box whose height corresponds to the range in depth and whose width corresponds to the abundance (e.g. in MK 16 on text-fig. 7). The code number of each fossil or mineral is plotted for identification but use of program CUZ 85 results in the same plots being produced together with a paper-tape list, in the correct order, of the full fossil names obtained from the appropriate dictionary corresponding to the code numbers on the plot. This paper-tape when fed through the phototypesetter produces a correctly typeset list which may then be affixed to the diagram in preparation for publication. CUZ 84 produces a paper-tape in the same way as CUZ 85 but plots the range and abundance data in fixed-order formats, e.g.



TEXT-FIG. 7. Upper Jurassic foraminifera from two boreholes in Central England. The data of borehole MK 24, penetrating younger strata, have been arranged by last occurrence as output by CUZ 85. Data of borehole MK 16, penetrating older strata, have been arranged as output by CUZ 84 in the order discovered in MK 24. Ammonite subzones have been plotted instead of lithology. Species abundance key as in previous diagrams.

SPECIMEN DIAGRAM



according to the current position of species in the taxonomic-order dictionary. In fact, by simply listing the desired sequence of code numbers, a range-diagram of species in any order at all may be drawn (text-fig. 7).

PROGRAM PERFORMANCE

At present the size limits of the data are 100 (species) × 100 (horizons) and the number of species at any one horizon must be limited to 50. In addition, the maximum size of any sub-groups of the data is 20 and the maximum number of sub-groups is 10. The preliminary listings given by CUF 11 and CUZ 88 may be used to inspect data size and to break down larger data into sub-groups of convenient size. Most projects so far handled have been found to be within the total size restrictions of the system. The largest dictionary presently held contains about 600 entries and a limit of 1000 entries is envisaged. Search times for this size of dictionary have proved to be short, so that the number of items is not considered of critical importance.

A comparison of time taken by the computer method against manual methods (Table 1) is based on an example containing some 750 data entries, in this case species occurrences. In Table 1, line 1 comprises the time in entering the data on, for example, standard Fortran data sheets. The rate of card punching (line 2) is fast because of the short input format. Operator time (line 3) is considered negligible since the bulk of time is spent in automatic plotting and paper-tape punching. This figure also includes the time involved in printing the names on the phototypesetter. The estimate of time (line 4) for the manual operation is based on a fixed format plot which is quicker (by an estimated 50%) than an initial plot in the first or last occurrence. The draughtsman's time (line 5) is again a minimum estimate and may be twice as long. Successive diagrams prepared by hand from the same data (lines 8–10) need complete reprocessing of the data whereas by the computer method it is only prepared once. It is here that most time is saved.

It will be seen (Table 1) that the manual time required to produce a diagram is of the order of ten times as long as the time taken by the computer method and increases if several diagrams are prepared from the same data. It has been calculated that this represents about three-quarters saving on costs within the Institute.

The use of the package has been presented from the point of view of its main objective, i.e. the speedy and cheap production of stratigraphical range-diagrams,

TABLE 1. Comparative times of computer and manual methods. For full explanation see text.

	Time in hours	
	Manual	Computer
1. Data preparation for computer		2·00
2. Data punching and checking for computer		1·00
3. Computer sorting and plotting		0·75
4. Total data preparation and plotting	20	3·75
5. Drawing office production	16	
6. Typing of fossil names	4	
7. Total for first diagram	40	3·75
8. Total for first and second diagrams	80	4·50
9. Total for first, second, and third diagrams	120	5·25
10. Total for first, second, third, and fourth diagrams	160	6·00

but it will be readily appreciated that the package is helpful in other respects, e.g. it has focused on errors in existing manual records; and may be used in analysis as well as in description. Thus appropriate sorting of the data cards enables species lists and plots to be generated for formations and groups or biostratigraphic zones within a region, and the ability to plot in any specified order enables diagrams to be prepared that are directly comparable to diagrams prepared from data from other areas. The dictionaries, while giving some personal nomenclatorial stability, may be output via a phototypesetter to produce simple faunal lists correctly typeset and as up to date as possible. In the long term the computer readable data may be used for input to a computerized data bank.

Acknowledgements. The programs for the sorting and display of the stratigraphic range and abundance data were written by Mrs. L. Johnson (Computer Unit, IGS) and that for the display of the lithological logs by Mr. D. G. Farmer (Computer Unit, IGS) who was also responsible for integration of the computer system and the development of ancillary programs. Dr. A. W. Medd and Mrs. B. E. Coleman (Palaeontology Dept., IGS) made text-fig. 7 available, Mr. R. J. Merriman (Petrographical Dept., IGS) provided the data of text-fig. 6, and Dr. T. J. Dhonau (Editorial and Publication Section, IGS) advised on the use of the phototypesetter. The paper is published by permission of the Director of the Institute of Geological Sciences.

REFERENCES

BURSCH, J. G. 1950. The range chart as an aid to formaniferal correlation. *J. Paleont.* **24,** 479–484.

FARMER, D. G. and JOHNSON, L. *In press. Rep. Inst. geol. Sci.*

MCLEAN, J. D. 1972. A philosophy of data treatment in ecology, stratigraphy and paleocology. *Manual of Micropaleontological techniques,* **12,** 1–257. McLean Paleontological Laboratory, Alexandria, Virginia, U.S.A.

MOORE, R. C. 1948. Stratigraphic paleontology. *Bull. geol. Soc. Am.* **59,** 301–326.

IAN E. PENN

Department of Palaeontology
Institute of Geological Sciences
Exhibition Road
London, SW7 5DE

Revised typescript received 23 October 1973