

Correspondence

To the Editor,
The Computer Journal.

Dear Sir,

Cockayne and Hyde, in their article, "Prime Number Coding for Information Retrieval," in *COMPUTER JOURNAL*, Vol. 3, pp. 21–22, describe a system "which so far as we know is novel."

The authors appear to have overlooked the system described by Leibniz in 1679 (C. I. Lewis, *A Survey of Symbolic Logic*, p. 11). The authors may be interested in a theoretical examination of prime-number coding from a strictly mathematical point of view: R. C. Buck, *Studies in Information Storage and Retrieval*. I. On the Use of Gödel Indices in Coding. MRC Technical Summary Report, No. 53, University of Wisconsin, October 1958.

Very truly yours,
Clifford J. Maloney.

Biomathematics Division,
United States Army
Chemical-Corps Biological Laboratories,
Fort Detrick, Maryland, U.S.A.
11 January 1961

To the Editor,
The Computer Journal.

Dear Sir,

"Prime Number Coding for Information Retrieval."

The method of labelling combinations by products of un-repeated primes was used, in the context of what is now called Information Retrieval, by Leibniz in 1679 (see, e.g., Cajori, 1930). It is well known, usually by rediscovery, to workers in this field. However, Mr. Berlin's use of Goldbach Pairs (this *JOURNAL*, Vol. 3, p. 210) is to the best of my knowledge a novel as well as an elegant variation.

The method is better suited to discussion of taxonomy (e.g. Fairthorne, 1947) than to coding an actual set of records for selection. Not only are better and cheaper methods available for this, but the prime number system does not label the set of records. It labels the enormously more numerous set of combinations that may be used irrespective of whether they are used or not, let alone of their relative frequency of use. Moreover, it does so with a vocabulary that becomes increasingly more vacuous.

Coding for selection needs to take into account only the number of existing items in the collection and the way in which they are used. One does not have to recover the descriptions from them, only the items by them.

Information Retrieval and allied activities are beginning to attract people with automatic data-processing devices in this country. Perhaps the following brief comments may save some disappointment and expense? First, "Information Retrieval" is an expression with definite meaning, introduced by Mooers (1950) after considerable consideration, trial, and consultation of dictionaries. It was used by workers in the Royal Aircraft Establishment and thence reintroduced into the U.S.A. The activity deals with retrieval of information, not of marked objects only, and therefore entails the analysis of subject-matter as an essential ingredient. Document retrieval, tape, etc., searching, emission of signals referring to reference numbers, and so on, are not "InfRet" but Data Processing. They are none the worse for that and need no alias.

Second, much has been done and is being done in this field, and there is little difficulty in finding out about most of it. Omission to do so can be embarrassing as well as expensive. The professional journals regularly publish articles, abstracts, reviews and notices; e.g. *Journal of Documentation*, *American Documentation*, *Nachrichten für Dokumentation*, *Revue de la Documentation*. The Office of Science Information, National Science Foundation, provide, for a few cents, a comprehensive survey of projects, other than those covered by security, in their periodically published *Current Research and Development in Scientific Documentation*, *Non-conventional Technical Information Systems in Current Use* and the bi-monthly *Scientific Information Notes*. The first two are essential reading; without them the risk of duplicated effort and expense is high.

Third; for Information Retrieval in a specialist collection of less than some tens of thousands of items, computer-like devices are just not worth while. The problem is already solved without their aid, and in many ways. Their use may be valid as a small-scale model of a large scale-system, though experience to date has shown this to be unlikely. Of course data-processing devices may well have profitable use in small specialist collections for purposes other than Information Retrieval.

Finally, may I immodestly suggest that those seeking an easily accessible survey of mechanization in Information Retrieval should glance at the first issue of this *JOURNAL*?

Yours sincerely,

Royal Aircraft Establishment,
Farnborough, and School of Library
Science, Western Reserve University,
Cleveland, Ohio.
4 February 1961
R. A. Fairthorne.

References

- CAJORI, F. (1930). *A History of Mathematical Notations*, Vol. 2, Section 670, Open Court, Chicago.
- FAIRTHORNE, R. A. (1947). "The Mathematics of Classification," *Proc. British Soc. for Int. Biblio.*, Vol. 9, p. 35.
- FAIRTHORNE, R. A. (1958). "Automatic Retrieval of Recorded Information," *The Computer Journal*, Vol. 1, p. 36.
- MOOERS, CALVIN N. (1950). "The Theory of Digital Handling of Non-numerical Information," *Zator Technical Bulletin*, No. 48, Zator Co., Boston, Mass.
- MOOERS, CALVIN N. (1950). "Information Retrieval Viewed as Temporal Signalling," *Proc. Int. Congress of Math.*, Cambridge, Mass., August–September 1950, American Mathematical Society, Vol. 1, pp. 572–73.