A BINARY-OCTAL CODE FOR ANALYZING HEXAMETERS

FRANK P. JONES
Tufts University

Recent studies of the Greek hexameter have been directed almost exclusively to the "inner metric."[1] Little has been done with the scansion pattern itself (or "outer metric") since the turn of the century. It is generally taken for granted that everything that needs to be said about scansion patterns was said at that time. The most thorough and the most frequently cited of the earlier studies are those of La Roche, who tabulated occurrences of the various metrical schemes in Homer, Hesiod, and Vergil.[2] La Roche's statistics, however, are neither complete nor easily available. They are based on editions which are no longer in use; and they are not free from error. For the first thousand lines of the *Iliad*, fifty-five of La Roche's scansions do not fit the readings of the Oxford text. Forty-two of these differences are due to a difference in the treatment of diphthongs, thirteen to errors in scansion.[3] I believe that a new analysis of the outer metric is called for. To

[1] The terms "inner" and "outer" metric were popularized by E. G. O'Neill, Jr., "The Importance of Final Syllables in Greek Verse," *TAPA* 70 (1939) 256–94; "Word-Accents and Final Syllables in Latin Verse," *TAPA* 71 (1940) 335–59; "The Localization of Metrical Word Types in the Greek Hexameter," *YCS* 8 (1942) 103–78. O'Neill defined the outer metric as "the syllabic, or quantitative, pattern of a verse form"; the inner metric as consisting of "the principles which govern the composition of words into verses, within the limitation of... a particular verse form" (1940, p. 336).

[2] J. La Roche, "Zahlenverhältnisse im homerischen Vers," *WS* 20 (1898) 1–69; "Untersuchungen über den Vers bei Hesiod und in den homerischen Hymnen," *WS* 20 (1898) 70–90; "Der Hexameter bei Vergil," *WS* 23 (1901) 127–29.

[3] In the whole of the *Aeneid*, G. E. Duckworth, "Variety and Repetition in Vergil's Hexameters," *TAPA* 95 (1964) 9–65, found 119 lines which La Roche had wrongly scanned and 121 errors caused by transpositions in the listings. In addition there were 14 lines for which the Ribbeck edition used by La Roche gives different readings and different scansion patterns from the Oxford text of Hirtzel.

facilitate it, I should like to propose a two-digit numerical code to replace the traditional scansion pattern. Based on binary-octal arithmetic,[4] it is well suited to computer programming. Because of its brevity and simplicity it would minimize human errors in transcribing and keep down the costs of publication. In this paper I shall define the code and describe some of the uses to which it might be put.

In binary arithmetic the base is two, rather than ten as in the decimal system. There are only two symbols, zero and one. In counting, the next number after 1 is 10 (one-zero) instead of 2. Each additional 0 at the right doubles the sum instead of multiplying it by ten. Thus 100 = four, 1000 = eight, 10000 = sixteen, and so forth.

Any series of units that can be dichotomized—described as either A or B—can be written as a binary number. The Greek or Latin hexameter provides an excellent example. Each of the first five feet must be either a dactyl or a spondee. The sixth foot, which never has more than two syllables, is conventionally treated as a spondee, no matter what the quantity of the final syllable. So the possible combinations of dactyls and spondees are 32 in number, or 2 to the fifth power.

To write an hexameter as a binary number, let 1 = a spondee and let 0 = a dactyl (or "non-spondee"). With this notation, the first line of the *Iliad*

Μῆνιν ἄειδε, θεά, Πηληϊάδεω Ἀχιλῆος

would be written 0 0 1 0 0 1. In the same way, the second line

οὐλομένην, ἣ μυρί᾽ Ἀχαιοῖς ἄλγε᾽ ἔθηκε,

becomes 0 1 0 1 0 1. The third line

πολλὰς δ᾽ ἰφθίμους ψυχὰς Ἄϊδι προΐαψεν

is 1 1 1 0 0 1.

[4] Number systems with bases other than ten are a prominent feature of the "new math." and are now explained in elementary textbooks of arithmetic. The best account of the binary-octal system that I have been able to find was published by the IBM Corporation: *IBM 7094: Principles of Operation* (IBM Corporation Customer Manuals; Dept. B98, Poughkeepsie, N. Y., 1963) 148–52. For the development of the binary-octal code I am much indebted to the criticism and suggestions of my colleagues at the Institute for Psychological Research, particularly John A. Hanson, Florence E. Gray, and Charles Budrose.

Binary numbers are too bulky for ordinary use and are usually converted either into decimals or into octals. Octals, which are formed with a base of eight, have certain advantages over decimals. Only the digits from 0 to 7 are used, and since 8 is the third power of 2, conversion from binary to octal is very easy. It is not necessary to understand octal arithmetic in order to make the conversion; simply divide the binary number into groups of three digits and substitute for any of the 8 combinations of 0 and 1 the corresponding octal number as shown in the following table:

000=0	100=4
001=1	101=5
010=2	110=6
011=3	111=7

To convert back from octal to binary, reverse the procedure. The first line of the *Iliad* is 11 in octal (0 0 1 = 1 in each half of the line). The second line, divided into two groups of three, is 0 1 0 1 0 1, which gives an octal number of 25. In the same way, the third line, 1 1 1 0 0 1, converts to 71. Examples of each of the 32 patterns together with their frequencies in the first 1000 lines of the *Iliad* are given in Table 1.

Besides being easier to write out and remember, octals have all the essential properties of binaries. Each octal number identifies and describes one of the 32 possible hexameter patterns.[5] By a simple operation, it gives the number of syllables in the line, the number of spondees and dactyls, and tells which foot or feet the spondees are in.[6] I have found the system useful as a kind of metrical shorthand for annotating my own texts. The two digits convey as much information as a scansion pattern and take up a great deal less space. (For Latin hexameters, the system can also be used to note homodyne-heterodyne patterns, which dichotomize the line in a different way.)

[5] There are 32 instead of 64 patterns because the last foot is always treated as a spondee. If it seemed desirable to distinguish the lines that end in a trochee this information could be given by using a zero as the last binary digit whenever the foot was not a true spondee. In the conversion to octal the four even digits, 0, 2, 4, 6, would then be used, giving a total of 64 code numbers.

[6] Because the binary-octal code is truly descriptive, more operations can be performed with it than with a set of arbitrary symbols.

TABLE 1. METRICAL PATTERNS IN THE *Iliad*

Code Number		Example	Frequencies in First 1000 Lines	
Octal	Binary	Book and Line	Jones	La Roche
01	000001	1.10	190	185
03	000011	1.21	12	12
05	000101	1.16	99	87
07	000111	1.226	2	2
11	001001	1.63	40	38
13	001011	1.472	4	2
15	001101	1.7	9	7
17	001111	1.339	1	1
21	010001	1.20	153	159
23	010011	1.14	11	11
25	010101	1.59	56	56
27	010111	2.167	2	2
31	011001	1.15	34	30
33	011011	1.11	1	3
35	011101	1.28	7	9
37	011111	9.137	0	0
41	100001	1.5	150	154
43	100011	1.107	7	7
45	100101	1.19	59	61
47	100111	2.123	2	2
51	101001	1.45	29	29
53	101011	2.104	1	1
55	101101	1.79	16	15
57	101111	13.428	0	0
61	110001	1.4	69	79
63	110011	2.345	2	2
65	110101	1.6	33	34
67	110111	2.388	1	1
71	111001	1.3	7	6
73	111011	8.472	0	0
75	111101	1.66	3	5
77	111111	23.221	0	0

Putting the octal numbers on punch cards for use in a computer program would be relatively simple. The standard IBM card has space for 36 two-digit numbers plus space for identification. The whole of the *Iliad* would thus require 436 cards, the *Aeneid* 275. Once the cards had been punched and proof-read any statistical operation available to a computer could be performed with the data and the results printed out. The operations might be merely descriptive like cataloguing and indexing. They might summarize frequencies within various subdivisions of the poem—speeches, narrative, formulaic lines, suspected passages, and the like. Or they might employ analytical techniques for correlating the frequencies in two or more samples.

A useful measure for testing homogeneity of pattern is the correlation coefficient (or r).[7] The size of r ranges from +1 for a perfect positive correlation to −1 for a perfect negative correlation. One would expect a high (but not perfect) correlation between two samples taken from the *Iliad* and a low (but positive) correlation between the *Iliad* and the *Aeneid*. This proves to be the case when the first 500 lines of the *Iliad* are compared with the second 500 lines and then with the first 500 lines of the *Aeneid*. For the comparison of *Iliad* with *Iliad* the correlation coefficient is .97;[8] for the comparison of *Iliad* with *Aeneid* it is only .10. (The data are given in Table 2.) With a computer, correlations can be obtained very quickly between samples chosen to test a great variety of hypotheses.

The binary-octal code was devised for the outer metric alone. It can, however, be advantageously adapted to a statistical treatment of the inner metric as well. The inner metric can be fully defined by numbering the syllables in the line, listing the numbers of the final syllables in each word except the last, and adding the octal code number in parentheses. (The last word does not need to be included in the numbering since the information it gives is already contained in the code number.) Thus in *Iliad* 1.1, inner as well as outer metric is defined by combining the sequence numbers 2, 5, 7, 12 with the code number (11). Two numbers plus the index number are all that are needed to define any word, phrase, or formula that is contained in a single line. This notation gives as much information about the inner metric as either of the other systems with which I am familiar.[9] For a

[7] The Pearson "product-moment" correlation, which measures the relationship between two variables, is put to frequent use in the physical and social sciences. Discussions can be found in almost any textbook of statistics, e.g. W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis* (New York 1951) 162–69.

[8] A comparison of *Iliad* 1.1-500 with *Iliad* 10.1-500 (using La Roche's figures) gives a correlation coefficient of .98.

[9] And gives it in a briefer space. O'Neill (see above, note 1) used a numerical system of notation which is similar to the system used by P. Maas (*Greek Metre*, translated by H. Lloyd-Jones [Oxford 1962]). The numbers indicate the position of a syllable within the line and within the foot but not within the word. When O'Neill's system is used

computer program it would be more economical than the other notations since it would require less time for card punching and offer less opportunity for error.

TABLE 2. FREQUENCY OF METRICAL INDICES: THREE SAMPLES COMPARED

A. *Iliad* 1.1–500; B. *Iliad* 1.501–2.389; C. *Aeneid* 1.1–500

Index	Frequency A	Frequency B	Frequency C	Index	Frequency A	Frequency B	Frequency C
01	104	86	3	41	75	75	10
03	7	5	0	43	6	1	0
05	44	55	38	45	23	36	28
07	1	1	0	47	0	2	0
11	26	14	18	51	8	21	18
13	2	2	0	53	0	1	0
15	5	4	70	55	9	7	47
17	1	0	0	57	0	0	0
21	78	75	22	61	39	30	12
23	6	5	0	63	0	2	0
25	27	29	52	65	14	19	37
27	0	2	0	67	0	1	0
31	15	19	24	71	4	3	15
33	1	0	0	73	0	0	0
35	3	4	69	75	2	1	37
37	0	0	0	77	0	0	0

Correlations: A vs. B, r=.97; A vs. C, r=.10

to transcribe single words or phrases the numbers have to be combined with the conventional scansion pattern in order to indicate all of the quantities. J. T. McDonough for his IBM "Iliad project" ("Homer, the Humanities and IBM," *Literary Data Processing Convention Proceedings*, September, 1964, Modern Language Association, New York, pp. 25–36) devised a system of coding in which numbers arbitrarily assigned identify a syllable as long or short and as occurring at the beginning, middle, or end of a word, or in a monosyllable. They do not give the position of the syllable in the line. Each of these systems calls for a minimum of 12 numbers to describe an hexameter line. R. D. Sweeney in a paper entitled "Computer Analysis of Classical Verse" presented at the 1966 meeting of the C. A. N. E. at Exeter, New Hampshire, described a system of notation for use with a computer program which gives a great deal more information about a line of verse than do any of the other systems. It requires a great deal more space, however. (Sweeney's paper is now published in the *Sixty-First Annual Bulletin* of the C.A.N.E., Hanover, N.H., 1966.)