

Typographs

Frank Harary

Typographs are defined as graphs derived from letters. A method for classification of letters by their underlying graphical structure is illustrated. A classification of the upper-case roman letters is presented, and tables of typographs are compiled as well for the roman numerals and the lower-case script letters. The possibility of devising new alphabets is mentioned.

The shape of letters has always had a fascination for literate mankind. With the advent of the computer came the need for the development of rapid methods for pattern recognition. The patterns formed by individual letters are studied in the science of graphology and the art form known as calligraphy. It is interesting to note that both calligraphy and graphology contain the word "graph" as the root. It is not at all unnatural to associate a graph with each letter. We have learned, in fact, that others are independently pursuing similar studies. We shall call these associated graphs *typographs*. Only capital roman letters will be studied in this paper. Other alphabets including the Hebrew, Greek, and Cyrillic may be investigated in later communications.

Graphical Preliminaries

Before presenting the typographs of the roman letters, it is necessary to explain quite a bit of graph theoretic notation and terminology. The following definitions, included here for the sake of completeness, may be found in [1].

A *graph* G is defined to be a finite, nonempty set of points, along with a set of unordered pairs of points. The pairs of points in the graph are called the *lines* of G . To illustrate, all the graphs with four points are shown in Figure 1. Those to the right of the dashed line are the *connected graphs* with four points.

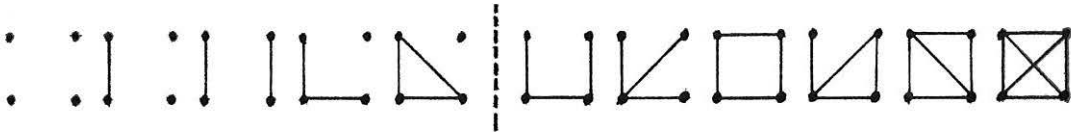


Figure 1. Graphs with Four Points.

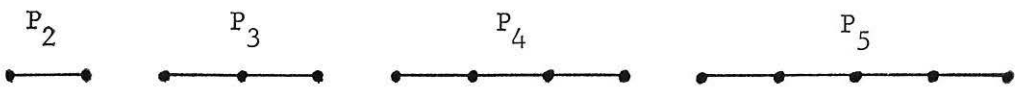


Figure 2. Paths.



Figure 3. Cycles.

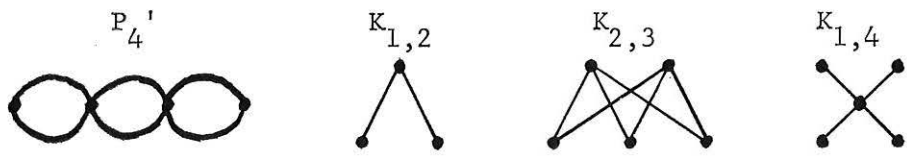


Figure 4. Types of Graphs.

We show the smallest graphs known as *paths* in Figure 2 and the smallest cycles in Figure 3.

An *acyclic graph*, also known as a *forest*, has no cycles; there are just six such graphs with four points. A *tree* is a connected acyclic graph; there are two trees with four points.

Strictly speaking, as in [1], the first two cycles shown in Figure 3 are not proper graphs; the first being known as a *loop* and the second as two lines *in parallel*. Just as P_n in Figure 2 denotes the path with n points, let P'_n be the graph derived from P_n by replacing each line by a pair of parallel lines.

The graph $K_{m,n}$ consists of two disjoint sets of points; one set with m points and the other with n . If two points appear in different sets, they are joined by a line, otherwise they are not. A *star* is a graph of the form $K_{1,n}$. Examples of the preceding types of graphs are shown in Figure 4.

Graphs of the Roman Capital Letters

In Table I we give the typograph of each letter. We have chosen letters similar to those used by Telingater [2].

The letters are classified in Table II as treelike, pathlike, unicyclic, starlike, or other. Note that the only letter in the "other" category is B, whose typograph is P_2' . In the same table, we have grouped together the letters which have the same typograph. Also, the letters are arranged in one of four groups depending on the frequency of occurrence of their typographs.

Table I. Typographs of Capital Roman Letters.

Letter	Typograph	Name	Letter	Typograph	Name	Letter	Typograph	Name
A			J		$K_{1,3}$	S		P_2
B		P_3'	K		$K_{1,4}$	T		$K_{1,3}$
C		P_2	L		P_3	U		P_2
D		P_2'	M		P_5	V		P_3
E			N		P_4	W		P_5
F			O			X		$K_{1,4}$
G		$K_{1,3}$	P			Y		$K_{1,3}$
H			Q			Z		P_4
I			R					

Table II. Classification of Typographs of Capital Roman Letters.








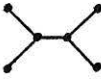



















Frequency	Typograph	Letter	Classification
1		E	tree
		F	tree
		A	unicyclic
	P_2^1	D	unicyclic
		O	unicyclic
		P	unicyclic
		Q	unicyclic
		R	unicyclic
P_3^1	B	other	
2		H, I	tree
	P_3	L, V	path
	P_4	N, Z	path
	P_5	M, W	path
	$K_{1,4}$	X, K	star
3	P_2	C, S, U	path
4	$K_{1,3}$	G, J, Y, T	star

Table III. Typographs of Arabic Numerals.

0	1	2	3	4	5	6	7	8	9
									
									

Conclusions

It is necessary to point out that the typograph of a letter definitely depends on the particular way it is drawn in the plane! Clearly two topologically different presentations of a letter can produce different typographs with possibly different classifications. As an illustration, we take the arabic numerals 0 through 9; occasionally giving two drawings, and thus two typographs.

Cuspless printing, that is printing with rounded rather than pointed characters, provides another illustration of ambiguity in presentation. In Table I, if we had used . . . rather than A, we would have had typograph (a) rather than (b) in Figure 5.

Here one should also note the relation between the capital script \mathcal{Q} and the capital printed Q. The typographs of the two are identical and are shown in Table I. Tables for both lower- and upper-case script roman letters similar to Table I for capital typed roman letters can be constructed and used for comparative purposes.

Other figures which have been analyzed and classified are the roman numerals and the lower-case script alphabet. To clarify the notation of Table IV, we define the *union* of two graphs G_1 and G_2 to be the graph $G_1 \cup G_2$ containing every point and line in either G_1 or G_2 . Figure 6 gives examples of unions of graphs.

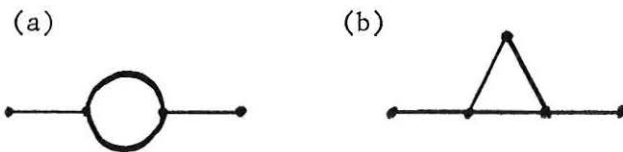


Figure 5. Two Typographs of A.

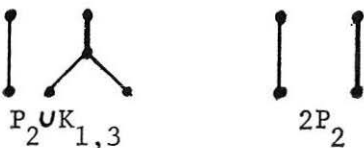


Figure 6. Unions of Graphs.

The first nine roman numerals are given below. The classification of more complicated numerals follows directly, and is left to the reader.

When analyzing the script lower-case roman letters, it is especially important to recognize the difference in typograph due to individual embellishments in penmanship. One set of typographs of the script lower-case Roman letters is presented in Table V.
















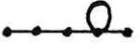
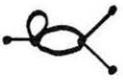

Amongst other things, our approach in this paper suggests the possibility of changing and improving our own alphabet, by exploiting typographs. Instead of using a letter, we could use its typograph, and in the cases where more than one letter belongs to the same typograph, we could choose a different plane embedding of the graph to correspond to each letter. For example, we could choose the embeddings (a) and (b) of Figure 7 of their (isomorphically) common typograph to represent H and I.

Clearly graphical classification offers limitless possibilities for modification and expansion of existing alphabets, and paves the way for the construction of entirely new alphabets.

Table IV. Classification of Roman Numerals.

Numeral		Classification	Numeral		Classification
I	P_2	path	VI	$P_3 \cup P_2$	forest
II	$2P_2$	forest	VII	$P_3 \cup 2P_2$	forest
III	$3P_2$	forest	VIII	$P_3 \cup 3P_2$	forest
IV	$P_2 \cup P_3$	forest	IX	$K_2 \cup K_{1,4}$	forest
V	P_3	tree	X	$K_{1,4}$	star

Table V. Typographs of Lower-Case *Script* Roman Alphabet in a Standard Presentation.

Letters	Typograph	Letters	Typograph
a		m, n	
b, h		r, s, v	
c		o	
d, e, l		p	
f		u	
g, q		t, x	
i		w	
j		y	
k		z	



(a)



(b)

Figure 7. Two Embeddings of a Typograph.

REFERENCES

1. Harary, F. *Graph Theory*. Reading, Mass. : Addison-Wesley, 1969.
2. Telingater, S. B. The standardization of alphabetic graphemes. *Journal of typographic research* (now *Visible language*), II (1968), 235-240.

Added in proof. This article was received for publication on January 11, 1970, at which time the name of this journal was as in reference [2]. It was circulated in preprint form and inspired the insightful and interesting paper: J. A. Bondy, The "graph theory" of the Greek alphabet. *Graph theory and applications* (Y. Alavi, et al., eds.), Berlin: Springer, 1972. Pp. 43-54.

This article was prepared with the assistance of Linda Bidelman.