Character Recognition Based on Phenomenological Attributes

B. Blesser, R. Shillman, C. Cox, T. Kuklinski, J. Ventura, M. Eden

A theoretical approach is suggested for describing upper-case letters not in terms of the physical attributes of their archetypes but in terms of more general descriptions of their underlying representations. A method is presented for finding these general descriptions through the study of ambiguous characters. Functional attributes are the describers of the underlying representations of letters. The relations between the physical attributes of the input character and the functional attributes that specify its identity are given in part by graphical context rules which incorporate the stylistic consistency within the character itself and its neighbors. The implications of our theory of characters to the areas of computerized character recognition and type design are noted.

I. *Introduction*

The literate human is adept at recognizing both hand and machine printed characters. Even though there are well over 1000 different type styles in general use in the United States (Karch, 1952), these variations do not affect our ability to correctly identify characters. Rarely used artistic fonts and hand printed characters, with their extreme idiosyncratic and variable forms, are easily recognized in the context of a word or when standing alone, even if one has never seen those particular variations before.

The ease with which humans identify characters misled many scientists into thinking that reading machines could be built based upon relatively straightforward descriptions of character archetypes. In contrast to human performance, the computer systems designed over the past 15 years are accurate only when they are restricted to recognizing characters which are limited to at most a small number of typefaces; changing the character's size, orientation, line thickness, serifs, etc., usually requires that the machine be retrained (for a survey of machine performance, the reader is referred to papers by Nagy [1968] and Harmon [1972]). An ideal machine would

Visible Language, VII 3 (Summer 1973), 209–223.

Authors' Address: Cognitive Information Processing Group, M.I.T., Cambridge, MA 02139.

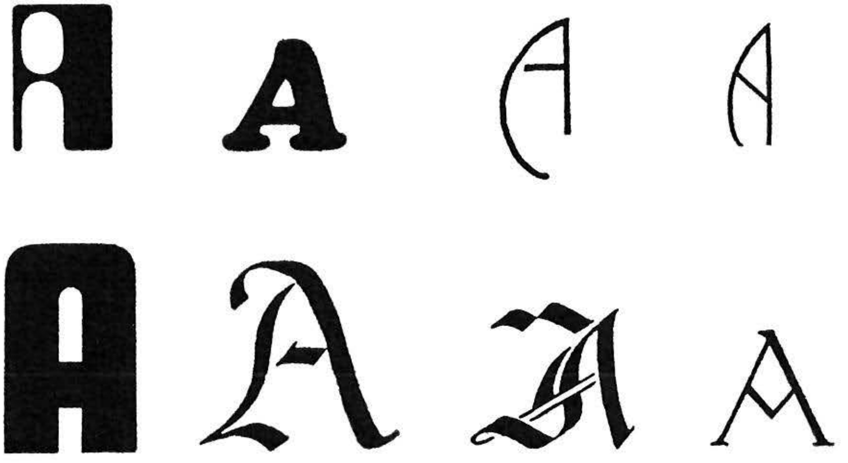


Figure 1. Various examples of the letter **A**.

duplicate human performance without the requirement of frequent retraining or updating.

Humans have the ability to generalize, to consistently group physically different objects into a common class. Although the physical similarities of the examples shown in Figure 1 are minimal, all literate people of our culture recognize them as the letter **A**.[1] Each of these allographs must possess certain invariants in order to account for this result; the discovery of these invariants has plagued psychologists and computer scientists alike. In the absence of knowledge of these invariants, we can neither explain how the human recognizes the examples in Figure 1 nor hope to design a computer recognition system which will duplicate human performance. Neisser (1967) stated the problem quite clearly: "Without some definition or criterion of similarity, no empirical prediction is possible; we are left to guess whether any particular stimulus will be

1. Henceforth, a letter printed in sans serif boldface will be used to represent the class of all characters that can be assigned that letter name; for example, **A** represents all characters that are called "ay."

recognized or not. Without an explicit model or mechanism, the notion of 'similarity' is only a restatement of the observed fact that some inputs are recognized while others are not." Similarity of letters can only be defined in terms of abstractions; the examples in Figure 1 are all A's despite the fact that they are not similar in terms of shape or size. Therefore, we make a sharp distinction between the physical image with all its allowable variability and an abstract underlying representation. We postulate a model for similarity based upon the presence of functional rather than physical invariants.

Functional similarity, however, depends on the nature of the question being asked. The allographs in Figure 1 are all similar in terms of the letter label, yet they are all dissimilar when considered as representative of their type style. Each allograph is a unique representation of its type style, e.g., Countdown (Letraset), Futura Display (Bauer), etc.

A model of character recognition can be called complete only when it provides a set of similarity relations for describing each letter such that the model predicts human performance. In our investigation, we seek an experimental strategy for finding these relations which will ultimately be stated as a set of "grammar-like" rules and which will be based upon a set of functional invariants or attributes.

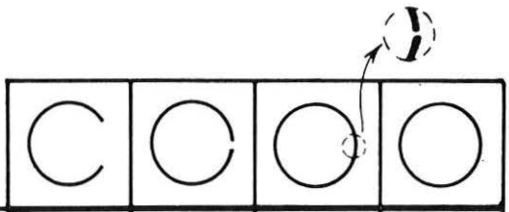
II. *The Relationship between Physical, Perceptual, and Functional Attributes*

We postulate the existence of three classes of attributes for each character.[2] Physical attributes are the geometric parts which make up the character—such as lines, angles, and physical closures. Perceptual attributes are the parts that are perceived by an observer as being present in the character. Functional attributes are the underlying describers of letters.[3]

The difference between functional, perceptual, and physical attributes is illustrated in Figure 2, for the attribute closure. The state of the attribute is indicated by a plus sign (+) for closed and by a

2. We shall use the term "character" to denote the specific physical printed sample to which an alphabetic (or numeric) label will be assigned.

3. "Letter" is used to represent a class of characters all of which are assigned the same alphabetic label.



Functional Closure	−	+	+	+
Perceptual Closure	−	−	+	+
Physical Closure	−	−	−	+

Figure 2. Three levels of the attribute closure: functional, perceptual, physical.

minus sign (—) for not closed. The first character has no closure in any of the three senses: it is physically open, it is perceived as being open, and it functions as open—since the character is identified as a C. In contrast, although the second character is physically open and perceptually open (a gap is visible on the right side), it functions as if it were closed—since it is reported more often as an O than as a C. The third character is physically open but is perceived as being closed, because the gap is too small to be resolved when viewed at a normal reading distance. The fourth character is closed in all three senses.

Physical closure is determined by examining the character according to a topological test for the existence of a completely surrounded white region. Perceptual closure is determined by experimentally answering the question, "Is it closed?" Functional closure can only be determined by answering the question, "Is it an O or a C?"[4]

Functional attributes are concise descriptors of the underlying representation of letters. Removing or changing a functional attri-

4. Although one might think that functional closure is the answer to the question, "Can you easily imagine the right side being closed?", this is not always the case. The contextual effects, which will be discussed later, are an integral part of the functional question.

bute usually changes a character's identity.[5] Since the change of a functional attribute can shift a letter's identity, a partial change in a physical attribute related to a functional attribute can result in an ambiguous character, one that is labelled equally often as either of two or more letters. Conversely, the ambiguous character is a manifestation of at least one attribute in transition; by observing ambiguities we find functional attributes. For example, Figure 3 shows various A-H ambiguities. Let us consider the attribute closure; by varying the physical image, we alter the degree to which the top of the character exhibits functional closure. However, we are faced with resolving why character 6 is more often judged closed (functionally) than character 14. Here we have a direct contradiction between physical closure and functional closure: character 14 is physically closed but functionally open, whereas character 6 is physically open but functionally closed.[6] Another apparent contradiction is illustrated by the top regions of characters 26 through 30. Although the regions are identical, character 26 possesses A-like functional closure whereas character 30 lacks it. These examples clearly indicate that observations on a region alone are not always sufficient to determine whether or not that region is functionally closed. Other portions of the character or neighboring characters provide a context within which the state of the functional attribute is resolved. In characters 26 through 30, the lower region provides a context for the characterization of the top region.

III. *Functional Attributes*[7]

Our method for discovering functional attributes is based on characters which have an ambiguous identity. We began the process of discovering functional attributes by creating the matrix shown in Figure 4 which contains 180 different ambiguous characters

5. Functional attributes are analogous to phonetic distinctive features in linguistic theory. In both a linguistic theory and a character theory, the abstractions are chosen to describe human performance in the most compact and predictive way possible.
6. This apparent contradiction can be resolved by using graphical context as shown in Section IV.
7. We purposely avoid using the term "feature" because its meaning in the literature is often unclear; it could refer to either the physical or the functional level.

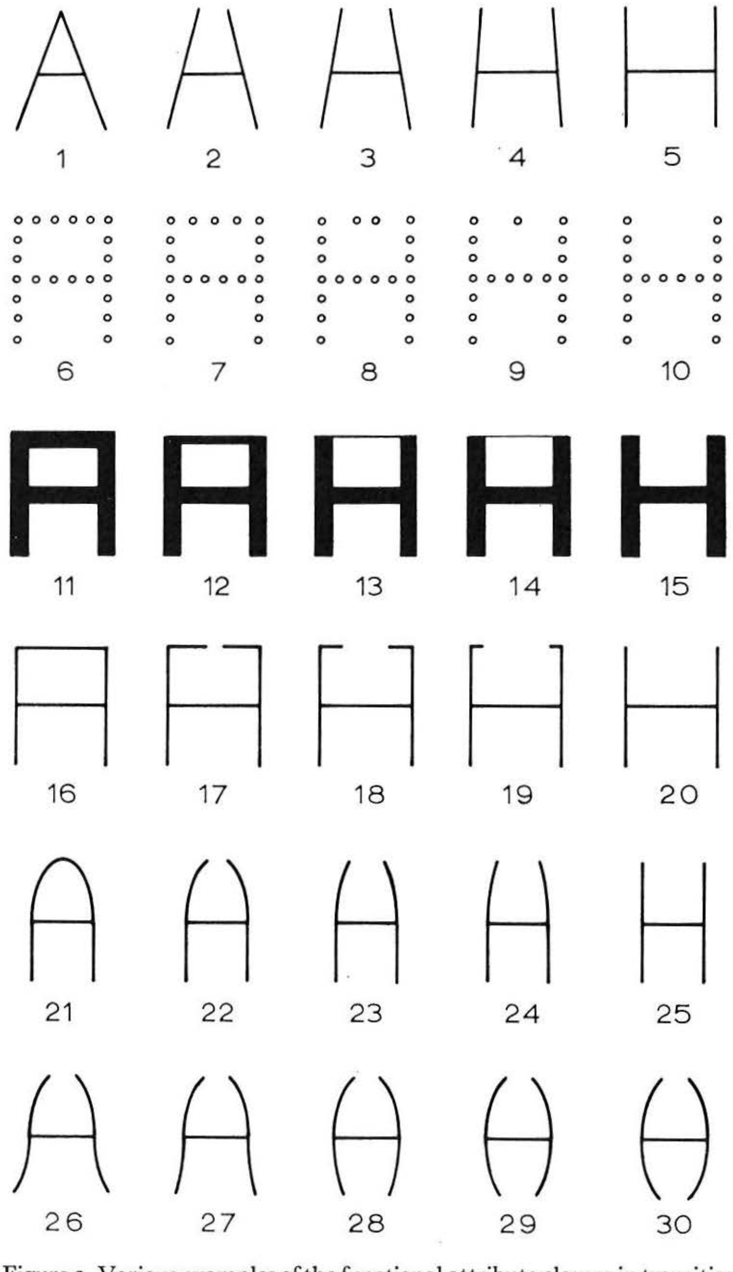


Figure 3. Various examples of the functional attribute closure in transition (left to right).

(Shillman and Blesser, 1973). According to our theory, such characters have one or more functional attributes whose state is not clearly defined.

To date, the functional attributes we have found are: closure, line extension, line addition, symmetry of intersection, symmetry (about the vertical or horizontal), smoothness, and verticality.[8] Examples of ambiguities involving these seven attributes are shown in Figure 5. The first character in Figure 5a is an **O** if the state of closure is a plus (+) and is a **C** if the state is minus (−). Similarly, the identity of each of the other characters in the figure depends upon the state of its functional attribute.

As we discovered new ambiguous characters, we found that they could often be described in terms of previously found functional attributes. Various combinations of the seven major functional attributes can describe over 90% of the ambiguous characters of Figure 4.

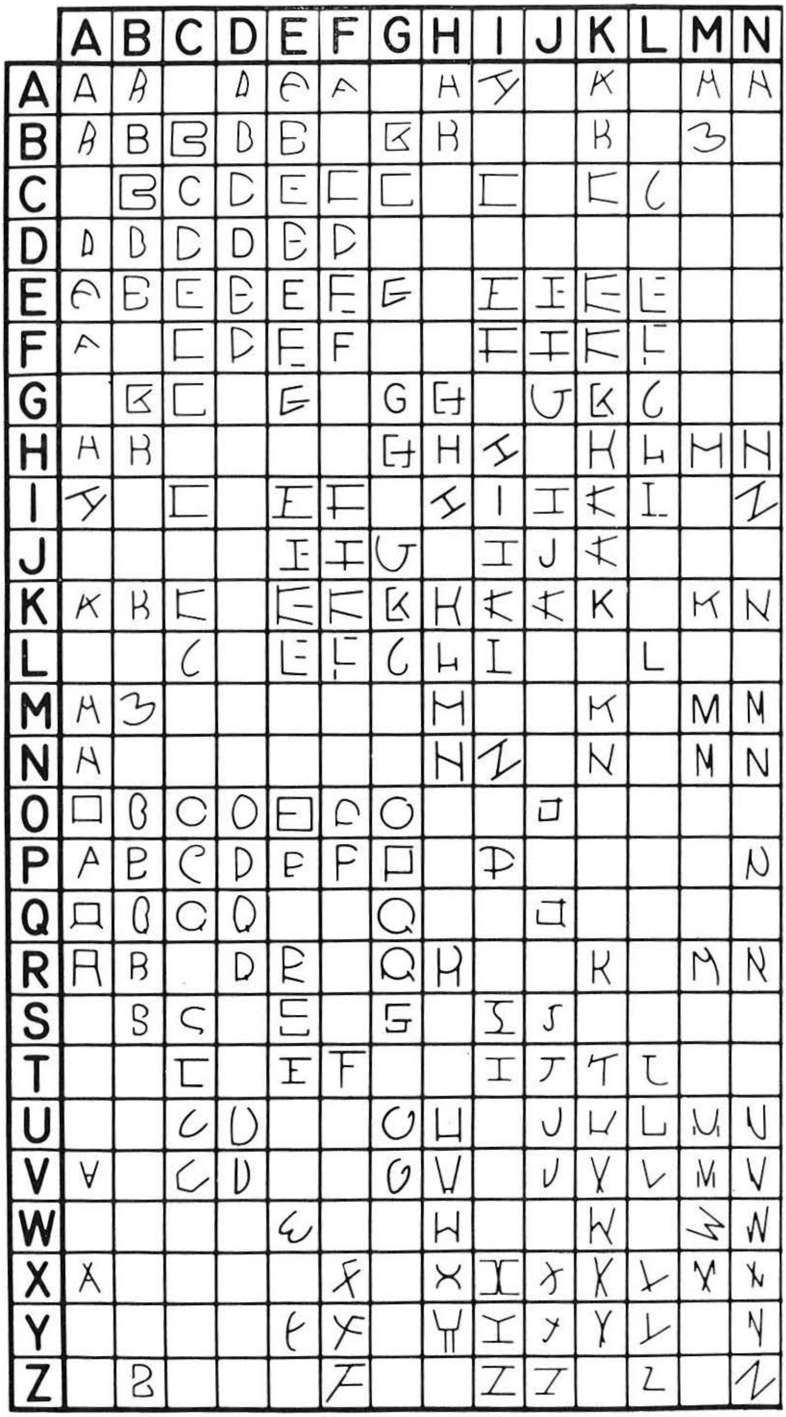
The underlying representation of each upper-case letter of the roman alphabet will ultimately be described in terms of the functional attributes that resolve the ambiguities involving that letter. Figure 6 contains a partial list of functional attributes of the letters **A**, **B**, **H**, **K**, **P**, and **R**; the list was prepared after considering 13 of the 56 ambiguities we have found involving these six letters.

IV. *Graphical Context*

In order to identify which functional attributes are present in a given character, one must be able to map from the given physical attributes of the sample to the set of functional attributes. We shall refer to these mapping rules as graphical context rules. It is our contention that these rules are a basic part of character recognition because they provide a basis for the mapping between the physical and the functional.

Many types of context[9] have been suggested as an aid to character recognition, e.g., semantic, syntactic, letter n-gram statistics, and

8. We do not argue that our list of attributes is complete, but we note that any counter example may itself generate the necessary additional attributes which must be added to our grammar.
9. We are using a broad definition of context; information from the general used to analyze the specific.


A B C D E F G H I J K L M N
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

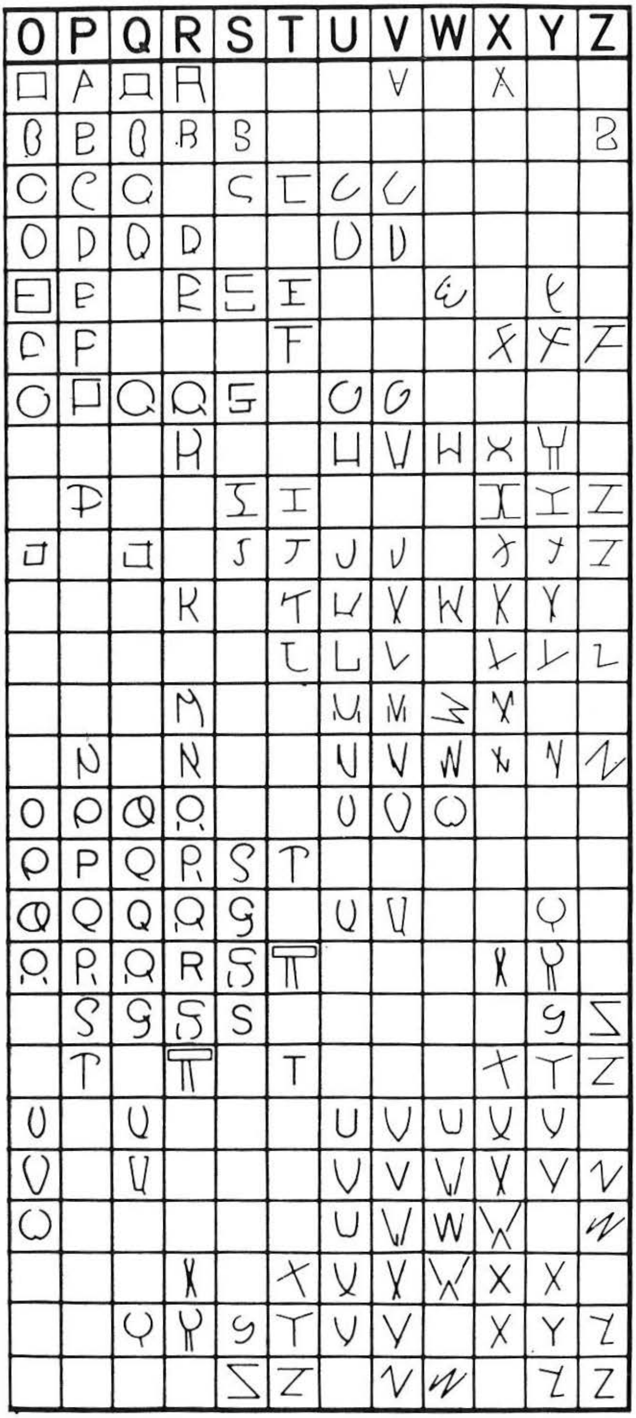


Figure 4.
A matrix of
ambiguous characters.

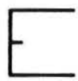
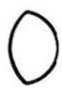
Functional Attribute	Character	Character Identity (state = +)	Character Identity (state = -)
Closure		O	C
Line Extension		Y	V
Line Addition		E	C
Symmetry of Intersection		A	R
Symmetry (vertical)		O	D
Smoothness		D	B
Verticality		N	Z

Figure 5. The seven major functional attributes and examples of characters whose identity depends upon them.

	A	B	H	K	R	P
Closure (upper)	+	+	–	–	+	+
Closure (lower)	–	+	–	–	–	o
Symmetry of Intersection	+	–	+	–	–	o

A zero (o) indicates that the functional attribute is not relevant for that letter

Figure 6. A possible coding of letter identity using functional attributes.

word probability of occurrence (Harmon, 1972). We shall restrict ourselves to a discussion of graphical context which is the particular type of context that depends upon the consistency of stylistic design within a character itself and with its neighbors.

In contrast to most character recognition techniques which implicitly take stylistic consistency into account, we will explicitly incorporate graphical context rules into our theory. The rules do not exist a priori; they must be derived from the given samples. They can be inferred by examining neighboring characters or through the examination of other parts of the character under consideration. We refer to these two types of inference as inter-character and intra-character graphical context respectively.

Incorporating inter-character context entails using information derived from the consistency of stylistic design among neighboring characters. For example, the character of Figure 7b is ambiguous, it could be an **N** or a **Z**. In the context of Figure 7a, the same character is easily identified as an **N**, whereas in the context of Figure 7c the character becomes a **Z**. Figure 7 shows how the mapping of the attribute verticality from the physical to the functional domain is achieved by using inter-character context; Figure 8 shows a similar

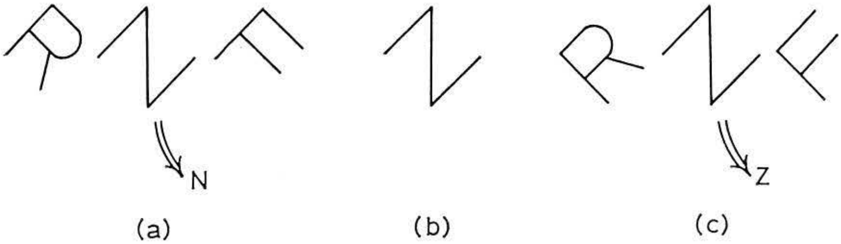


Figure 7. The effect of inter-character context upon the functional attribute verticality.

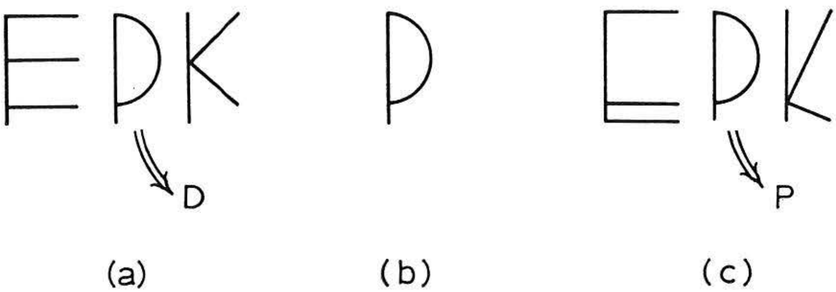


Figure 8. The effect of inter-character context upon the functional attribute line extension.

Figure 9. The effect of intra-character context upon the determination of functional closure.

effect for the attribute line extension. In both figures, inter-character consistency can be used to identify an otherwise ambiguous character.

Intra-character context can be viewed as a localized version of inter-character context; it entails using the consistency of stylistic design within the character itself. To illustrate this concept, consider the confusing examples discussed in Section II which are reproduced in Figure 9. Using intra-character context, the apparent contradiction between physical and functional closure can be resolved by considering the definition of line within the stylistic design of each of the characters. Figure 9a is an A since a sequence of open circles defines line within that particular character; Figure 9b is an H because the definition of line demands a thick physical line, hence the thin upper horizontal line is rejected as being a stylistic flourish or an artifact even though it is a line in the physical sense. In both cases, intra-character context determines the state of functional closure.[10]

10. According to our theory, graphical context is necessary for character recognition; we do not claim it is always sufficient. The following example illustrates how ambiguity can persist when inter- and intra-character context conflict. The identity of the first character remains in question since inter-character context implies that top horizontal lines can be thin whereas intra-character context implies that this line is an artifact.

V. *Summary*

We have proposed a theory of characters in which each upper-case letter of the roman alphabet will be represented by a unique set of underlying attributes. We have presented a technique for finding these functional attributes through the study of ambiguous characters. The characterization of the alphabet is not yet complete but evidence indicates that the number of functional attributes will be small. We do not argue that our set of functional attributes is the only possible set; we will choose that set of attributes which yields the most compact description of letters.

The relationship between the physical attributes of a character and the functional attributes which specify its identity are given by the graphical context rules. These rules do not exist a priori: they must be derived from the samples under consideration.

Experimental data relating physical attributes to functional attributes in specific contextual situations have been obtained (Sefkow, 1973). In addition, we are currently investigating the graphical context rules for several different type styles. Experimental results in support of our technique are forthcoming.

VI. *Conclusions*

Our theory of characters has implications both to the area of computerized character recognition and to the design of new type fonts.

Because our technique for describing letters is based upon the underlying functional attributes of letters as opposed to the physical features of archetypes, we feel that a computer implementation of our technique will have a higher recognition rate and be capable of handling a wider variety of input text than currently available character recognition systems.

According to our theory, the design of a new type style can be viewed as character recognition in reverse. The design of a new type font is the creation of a set of *inverse* graphical context rules that operate on the invariant set of functional attributes that define letter identity to yield new graphemes.

REFERENCES

Arkadev, A. G., and Braverman, E. M. *Computers and Pattern Recognition.* Washington, D.C.: Thompson, 1967.

Harmon, Leon C. Automatic recognition of print and script. *Proc. IEEE*, 1972, *60*, 1165–1176.

Jaspert, W. P., Berry, W. T., and Johnson, A. F. *The Encyclopaedia of Type Faces.* London: Blandford Press, 1970.

Karch, R. Randolph, *How to Recognize Type Faces.* Bloomington, Ill.: McKnight and McKnight, 1952.

Nagy, George. State of the art in pattern recognition. *Proc. IEEE*, 1968, *56*, 836–862.

Neisser, Ulric. *Cognitive Psychology.* New York: Meredith, 1967.

Sefkow, T. Symmetry as an aspect of character recognition. B.S.thesis, M.I.T., January, 1973 (unpublished).

Shillman, R. J., and Blesser, B. A. The use of ambiguous characters in measuring functional invariants. *Quarterly Progress Report* (M.I.T., Research Laboratory of Electronics) 1973, *109*, 155–158.

This work was supported by the National Science Foundation under Grant GK-33736X.