

A computer program to generate and measure random forms¹

RONALD L. KNOLL AND HERBERT H. STENSON

BEHAVIOR RESEARCH LABORATORY, ANTIOCH COLLEGE

A computer program is described that generates random form stimuli by a method parallel to that described by Attneave & Arnoult (1956) for construction forms with both angles and arcs in their perimeters. The program also performs a physical analysis of the forms by computing the values of a large number of physical variables describing each form. Already existing forms also may be analyzed by the program. Thus, it has the capability to perform physical analyses of approximations to natural forms.

The methods devised by Attneave and Arnoult (1956) for the generation of form stimuli allow one to sample a large, well defined population of forms and have been used extensively to generate stimuli for studies of form perception. Attneave and Arnoult's Method 4 allows both angles and arcs to appear in the perimeters of forms and, therefore, specifies a stimulus domain of more general interest than Method 1, which allows only angular forms. However, the manual generation of forms of the Method 4 type becomes laborious when the random variables involved in form construction are allowed to vary over a wide range. The measurement of the geometric characteristics of the Method 4 forms also is difficult and involves a prohibitive amount of manual calculation if more than just a few, simple physical descriptors are desired.

This paper describes a computer program that was written to reduce the efforts of constructing form stimuli of the Method 4 variety and to allow the computation of a large number of physical measures of their shapes. The program can also be used to measure the physical characteristics of already existing forms by giving the program certain data from the forms.²

PROGRAM OPERATION

The program logic parallels that of Method 4 developed by Attneave and Arnoult (1956) for the construction of closed forms with both straight and curved segments in the perimeter. That is, points are located randomly on a square grid and connected to form a closed figure whose perimeter consists entirely of straight line segments. Some of the angles in the perimeter then may be replaced by arcs, with the sizes and positions of the arcs randomly determined.

The program allows the user to specify the particular subpopulation of forms to be sampled through appropriate use of two groups of parameters. Values

for the first group of parameters must be given to the program by the user. This group of parameters controls (1) the number of starting points and the number of arcs, (2) the types and numbers of physical variables that are to be measured on the forms, and (3) the number of forms to be generated. These parameters are discussed later in the section on program input.

The second group of parameters controls four general characteristics of the forms generated by the program: (1) the grid size for locating the initial points; (2) the length of individual straight segments in the perimeter prior to the construction of arcs; (3) the magnitude of the angles in the perimeter prior to the construction of arcs; and (4) the magnitude of the radius of an arc. Values for these parameters are defined by the program, but may be altered by the user with appropriate input to the program. For example, one parameter of this group is used to test the magnitudes of the angles in the perimeter of the form. The existence of an angle of magnitude θ , such that $175^\circ < \theta < 185^\circ$, results in the printing of an "error" statement that an angle of the form has a magnitude that belongs to the restricted range 175° to 185° . After the printing of the error statement the program ceases calculations for the form with the angle in the restricted range, and begins calculations for another form. The restriction imposed on angle magnitudes can be altered by the substitution of a different value, ϕ , for this parameter so that forms with angles of magnitude θ , where $\phi < \theta < (360^\circ - \phi)$, are not permitted.

The initial straight segments in the perimeter of a form are restricted in size by another parameter in this second group. The existence of a straight segment with a length less than $1/10$ the average length of all the straight segments causes the printing of an error statement and the start of calculations for another form.

The size of the grid is defined by the program to be 100 units by 100 units. The length of each arc's radius is restricted to be less than 3 times the grid size by another parameter. Hence, if the grid size were to be changed by input to be 200 by 200 units, then no arc radius could have a length equal to or greater than 600 units.

At the same time that the spatial configuration of the form is being determined, the program accumulates physical descriptor data such as the

lengths of the arcs and straight segments in the perimeter. A description of the physical data that is part of the program output is given below in the section on output. It is also possible to analyze the physical characteristics of already existing forms with this program, and a later section describes this process.

A noteworthy feature of the physical analysis that is performed on the forms is the calculation of what we have called the "area moments" of the forms. A form is treated as a bivariate density function of unit height within the form and of zero height everywhere else. Moments along two arbitrary reference axes and product-moments are calculated for this bivariate density function defining the form. Moments and product-moments along the two principal axes of the form also are calculated using the center of gravity of the form as the origin. Zusne (1965) discusses the use of the principal axis moments of forms as physical variables to describe shape, and Hu (1962) shows some interesting properties of moments when they are used in this manner.

An integral method for calculating area moments of the forms constructed by our program was developed, and a description of the method is given in a later section.

Input

Certain general characteristics of the forms to be constructed are specified by a group of input parameters. One parameter specifies how many forms are to be constructed at one time, and two other parameters control the number of angles and arcs to be allowed in the forms. Four different conditions are possible for assigning the numbers of angles and arcs.

Condition 1. This condition places the least restriction on the number of angles and the number of arcs. The program selects at random the number of starting points within limits specified by the input. The number of angles that are to be replaced by arcs also is selected at random, and may be equal to the number of starting points, zero, or any intermediate value.

Condition 2. For this condition the number of starting points and the number of arcs, for each form to be constructed, are specified by the user and given to the program.

Condition 3. The number of starting points for a group of forms is given to the program, and the number of arcs for each form is determined randomly by the program as in Condition 1. Therefore, all the forms have the same number of starting points, but have different numbers of arcs.

Condition 4. Both the number of starting points and the number of arcs are constant for a group of forms. The two numbers are given to the program only once.

Additional parameters control the types and the numbers of physical variables that are measured on each form that is constructed. For example, central moments of the distributions of arc lengths, straight-segment lengths, and angle sizes are calculated for each form. The number of moments to be calculated for each distribution is determined by a control parameter. Likewise, another parameter controls the number of area moments to be computed.

Output

Both physical-variable data and coordinate data for form construction are included in the output of this program. The sampled forms can be drawn easily with the use of the coordinate data, and the physical variables provide a basis for constructing more sophisticated measures of shape.

Physical variables. The values of the following list of variables are calculated for each form. The units for the data are grid units unless otherwise noted.

- (1) Number of turns (angles plus arcs)—equal to the number of starting points.
- (2) Number of arcs.
- (3) Number of convex angles—the number of angles whose magnitude, as measured interior to the form, is less than 180° .
- (4) Number of convex arcs—the number of arcs that have replaced convex angles.
- (5) Perimeter—the length of the perimeter of the form.
- (6) Area—the area of the form.
- (7) Mean Straight Length—the mean of the distribution of the lengths of all straight segments in the perimeter.
- (8) Straight-Length Central Moments—central moments of the distribution of the lengths of the straight segments.
- (9) Mean Arc Length—the mean of the distribution of the lengths of all curved segments in the perimeter.
- (10) Arc-Length Central Moments—central moments of the distribution of the lengths of the curved segments.
- (11) Mean Internal Angle—the mean (in radians) of the distribution of the magnitudes of the internal angles, including those angles that were replaced by arcs.
- (12) Internal-Angle Central Moments—central moments (in radians) of the distribution of the magnitudes of the internal angles, including those angles that were replaced by arcs.
- (13) Area Raw Moments—area moments with respect to an arbitrary coordinate system whose reference axes are parallel to the grid axes.
- (14) Area Central Moments—area moments with respect to the coordinate system whose axes are parallel to the grid axes and whose origin is the center of gravity of the form.

- (15) Principal-Axes Area Moments—area moments with respect to a particular center of gravity coordinate system. (See section on Area Moments.)
- (16) Orientation Angle—the angle (radians) that the first principal axis of the form makes with the positive horizontal axis of the grid on which the form is drawn.
- (17) The length of each straight segment in the perimeter.
- (18) The length of each curved segment in the perimeter.
- (19) The magnitude (radians) of each interior angle, including those angles that were replaced by arcs.
- (20) The length of the radius of each arc.

Coordinate data. The following data are provided for each form that is constructed by the program. These data are sufficient to draw the form.

- (1) The coordinates of the vertex of each angle before any angles are replaced by arcs.
- (2) The coordinates of the clockwise endpoint of each arc.
- (3) The coordinates of the center of curvature for each arc.
- (4) The coordinates of the center of gravity of the form.

Existing Forms

The program can calculate values of physical variables for Method 4 forms that already exist, when provided with the appropriate coordinate data. Sufficient data to completely determine the spatial characteristics of the form are the number of starting points, the coordinates of the starting points, the number of arcs that replace angles, the coordinates of one endpoint of each arc, and a code for each angle to assign the position of the arcs. The program calculates additional coordinate data and the physical variable data, and produces the same output as for the forms that are constructed by the program.

A particular value of an input parameter instructs the program that data is to be generated for an existing form. The flow diagram in Fig. 1 shows the use of four input parameters by the program to allow the calculation of data for an existing form. Figure 1 also shows the four methods for determining the number of starting points and the number of arcs when forms are to be constructed as discussed in the input section.

The diamonds in the flow diagram refer to tests made by the program on the magnitude of the param-

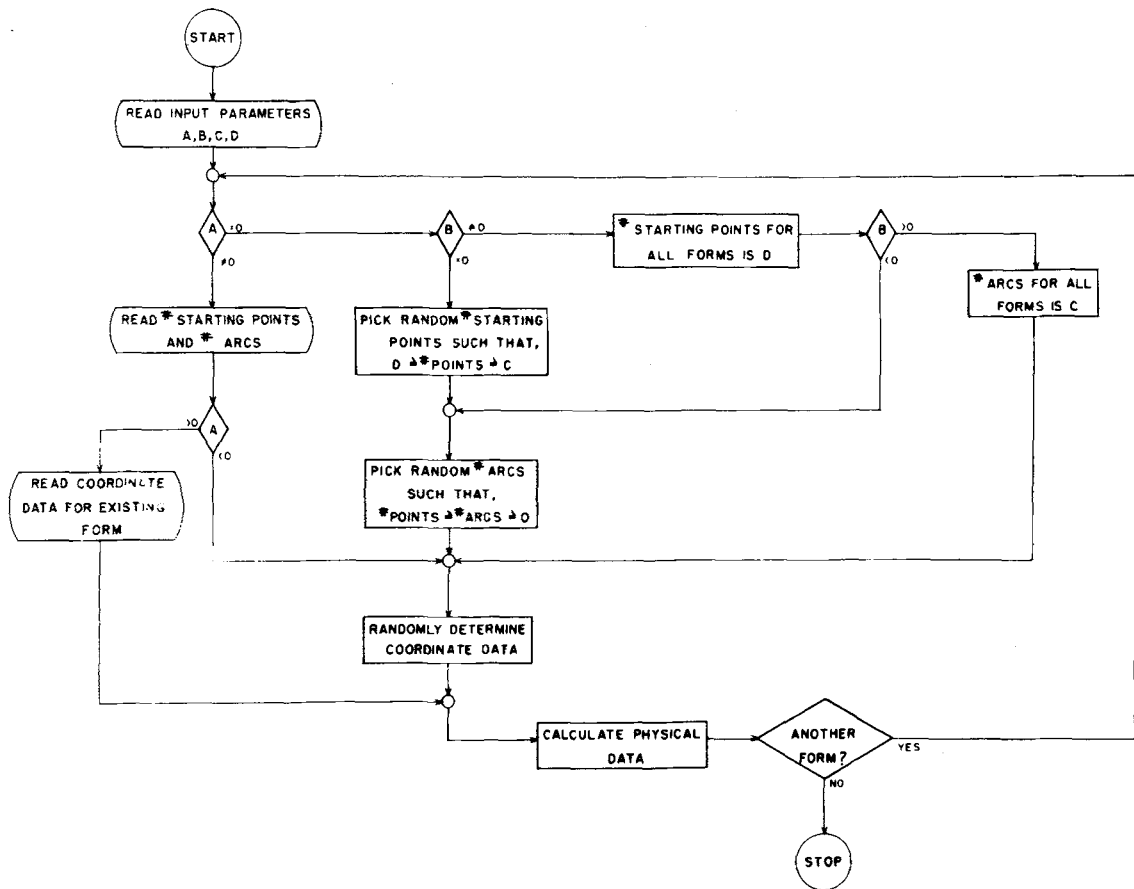


Fig. 1. Flow diagram showing the use of input parameters to analyze existing forms or to construct and analyze forms according to one of four conditions described in the text.

eter inside the diamond. The sequence of the program is determined by setting the parameter less than, equal to, or greater than zero.

AREA MOMENTS

If a form is considered as a bivariate density function of unit height interior to and at the perimeter of the form, and of zero height exterior to the form, then a general moment, $M_{nx:my}$, can be defined in the XY plane by the integral equation,

$$M_{nx:my} = \int_S f(x, y) x^n y^m dS / \int_S f(x, y) x^0 y^0 dS, \quad (1)$$

where the integration extends over the entire plane containing the form and dS is the differential area element. Confining the integration to the perimeter and interior of the form and recognizing that the denominator of (1) is simply the area of the form, Equation (1) may be written

$$M_{nx:my} = \int_{S_f} x^n y^m dS_f / \text{Area}, \quad (2)$$

where dS_f is a differential area element within the form.

Every area moment can be approximated by dividing the form into small subareas and considering only one coordinate product for each subarea. Each coordinate product is weighted by the magnitude, a_i , of the area it represents. In this approximation the general area moment is given by

$$M_{nx:my} = \sum_i a_i x_i^n y_i^m / \text{Area}, \quad (3)$$

where i refers to a particular subarea and the point, (x_i, y_i) , is the center of gravity of the i th subarea.

When exact area moments are calculated for non-overlapping subareas of a form, then the moments for the entire form are given exactly by the equation

$$M_{nx:my} = \sum_i \int_{S_i} x^n y^m dS_i / \text{Area}, \quad (4)$$

where each integration value is an area moment for a particular subarea of the form, weighted by the magnitude of the subarea.

The program calculates area moments for a form using a slight modification of Equation (4). Area-moment products for small areas exterior to the form are included twice in the summation, with opposite signs for the two occurrences. The resultant values of the area moments of the form are identical to those that would be calculated by

not including the area-moment products for the exterior areas in the summation, since the area-moment products for each exterior area cancel. The process can best be explained with reference to a particular example such as the form shown in Fig. 2. Area-moment products are calculated for two types of subareas: the area bounded by the sides of a triangle and the area bounded by an arc and a chord to that arc.

Figure 2 illustrates the division of a particular form. One subroutine calculates area-moment products for all subdivisions of the arc type. If an arc is concave (an arc whose center of curvature is exterior to the form), such as at angle 8, then the area of the subdivision is exterior to the form and the calculated area-moment products are multiplied by -1 . The form then defined by straight lines connecting angles without arcs (1, 2, 3, 5, 6, and 7) and the endpoints of the arcs 4 and 8, is divided into triangles as shown by the dotted lines. Area-moment products are calculated for each triangle by a second subroutine. Moments for the entire form are then given by Equation (4). Since the triangular areas overlap the concave arc areas, which are exterior to the form, the area associated with a concave arc and the moments of that area are included twice in the summation. However, the net effect of these exterior areas on the moments for the entire form is zero because the sign of the area-moment product is opposite for the two occurrences. The logic required to calculate moments for each type of subarea is complex, and this technique avoids the necessity of a subprogram for a third type of subdivision (i.e., a subdivision whose area is bounded

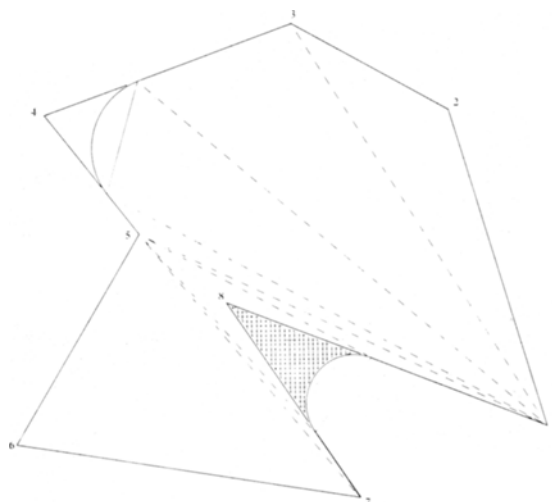


Fig. 2. The division of a typical form into subareas for the calculation of area moments, as described in the text. Eight starting points were connected to give a closed form after which angles 4 and 8 were replaced by arcs. The significance of the cross-hatched portion of the form is discussed in the text.

Table 1.

Sample Times for the Calculation of Data for Existing Forms and for the Simultaneous Construction and Analysis of Random Forms

Program type	No. of forms	No. area moments	Mean no. starting points + arcs	Computer system	Total time (sec)	Mean time per form ^a (sec)
Analysis only	50	15	19	7040/7094	209	4.1
Construct & analyze	82	91	19	7044/7094	3247	63.3
Construct & analyze	25	91	18	7040/7094	1959	78.4
Construct & analyze	19	15	18	7044/7094	82	6.9
Construct & analyze	15	15	18	7040/7094	74	4.9

^a The mean time per form has been adjusted to the IBM 7040-7094 system which is approximately 5/8ths as fast as the IBM 7044-7094.

by an arc and the two tangents of the arc, shown cross-hatched in Fig. 2).

The discussion above refers to raw area moments, but central area moments are derived by our program from the raw area moments. Central area moments are defined as the moments calculated with respect to a coordinate system whose origin is located such that the first area moments ($M_{1x:0y}$ and $M_{0x:1y}$) are equal to zero. This system of coordinates sometimes is referred to as a center of gravity system. Another reference frame that has common usage is the principal-axes coordinate system, which is a center of gravity coordinate system with the added restriction that the first cross-product moment ($M_{1x:1y}$) is equal to zero. The principal-axes frame of reference can be obtained from any central-axes frame by a rotation of the central axes in the plane of the axes.

Our program calculates central moments in the principal-axes frame as well as in the central-axes frame whose axes are parallel to the construction grid. Hu (1962) shows the relationships between the moments with respect to various frames of reference and calculates some area moment combinations which are invariant under rotation and reflection of the coordinate axes. A derivation of area moment measures that are independent of a change of scale on the coordinate axes is also given by Hu (1962).

PROGRAM PARTICULARS

The program described has been successfully compiled and run on the IBM 7040-7094 and 7044-7094 systems located at Wright-Patterson Air Force Base, Ohio. Except for a subprogram written in MAP program language for the calculation of random decimals,³ Fortran IV is the language in which the program has been coded.⁴

The program is 1700 statements in length of which about one-third are nonexecutable comments. The

program length is divided almost evenly between the main program and a block of subroutines that are called by the main program. As presently dimensioned, somewhat less than 5K storage locations are required for data storage, while approximately 19K locations are required for the storage of the program logic. Twenty locations each have been provided for the central moments of the straight lengths, arc lengths, and interior angle magnitudes. Three square arrays of order 22 have been designated for the calculation and storage of area moments. Area moments $M_{nx:my}$ are calculated for all n and m subject to the restrictions that $n, m = 0, 1 \dots 21$ and $n + m \leq 21$.

Physical and coordinate data have been generated for 163 forms and physical data have been calculated for 50 forms that existed prior to the development of the program. Because forms were generated in blocks, no empirical relationships can be constructed for estimating the time dependencies of the program. The logic of the program suggests that the time required to calculate physical and coordinate data should be proportional to the sum of the number of starting points and the number of arcs (for a constant number of area moments). Preliminary results indicate that the time should be approximately proportional to the number of area moments that are calculated, given a constant sum for the number of starting points and the number of arcs. These relationships are not exact because some calculations are performed independently of these variables. Some sample times for the program are given in Table 1.

DISCUSSION

In the preceding sections we were concerned mainly with forms that contained both straight segments and arcs in the perimeter because it is this type of form that is most difficult to construct and measure. However, it is obvious that the program can be used to construct and/or analyze forms with only straight perimeter segments (Attneave & Arnoult, 1956, Method 1). Using Condition 2 as described earlier, the number of arcs can be specified as zero by the user for any number of starting points. The program will then construct Method 1 forms.

The program also can be used to calculate physical data for forms whose shapes do not conform strictly to the requirements of the Attneave and Arnoult methods if the forms can be adequately approximated either by forms with straight segments only, or by forms with straight and arc segments. Given any form whose perimeter is continuous (i.e., the form is not divided into separate parts), the shape can be approximated by connecting the points of the perimeter with straight line segments. The closeness of the approximation could be improved or lessened by shortening or lengthening the dis-

tances between the adjacent points on the perimeter. For some shapes, curved portions of the perimeter could be approximated best by replacement of those portions with arcs of various radii. By using such approximations it would be possible to use our program to measure physical variables describing outlines of natural forms such as buildings, airplanes, automobiles, etc.

The use of natural outlines would present some difficulties in the interpretation of physical variables that have been associated with Attneave and Arnoult forms in the past (see, for example, Brown & Owen, 1967; Stenson, 1966). For example, variables such as the mean length of the straight segments in the perimeter would be highly dependent on the number of points of the perimeter of the natural shape that was designated to accomplish an approximation by straight segments. However, the problem of interpreting the meanings of physical variables for natural shapes might be avoided by the appropriate choice of physical variables. For example, the ratio of the perimeter squared to the area would be relatively independent of the actual approximating form that was employed. Physical variables based on area moments seem particularly attractive in this connection. Area moments describe general aspects of a form rather than particular details, so that the moments of the approximating form and the natural form would have a high correlation when the perimeters of the two were "close."

We have attempted to provide a rather basic list of physical variables as the output of the program so that the experimenter can use them to construct measures of form that are the most appealing to him. All the central moments, for example, are unstandardized. The user may wish to standardize some or all of them by dividing by powers and prod-

ucts of the second central moments, or he may wish to use some of the interesting combinations of moments suggested by Hu (1962). Likewise, perimeter and area measurements are given separately so that they can be used in combinations that, for certain experiments, might be more appropriate than the more conventional "perimeter squared to area ratio."

References

- Attneave, F., & Arnoult, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, 53, 452-471.
- Brown, D. R., & Owen, D. H. The metrics of visual form: Methodological dyspepsia. *Psychol. Bull.*, 1967, 68, 243-259.
- Hu, M-k. A mathematical model for visual perception. In E. E. Bernard & M. R. Kare (Eds.), *Biological prototypes and synthetic systems*. New York: Plenum Press, 1962. Pp. 222-229.
- Stenson, H. H. The physical factor structure of random forms and their judged complexity. *Percept. & Psychophys.*, 1966, 1, 303-310.
- Zusne, L. Moments of area and of the perimeter of visual form as predictors of discrimination performance. *J. exp. Psychol.*, 1965, 3, 213-220.

Notes

1. The research reported in this paper was sponsored by the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, under Contract No. F 33615-67-C-1280 with Antioch College, Yellow Springs, Ohio. This paper has been identified by the Aerospace Medical Research Laboratory as AMRL-TR-67-194. Further reproduction is authorized to satisfy needs of the U. S. Government.
2. The authors will be receptive to requests for a Fortran listing of the program and instructions for its use. Requests should be sent to the second author.
3. The random number subprogram was acquired from the files of the Digital Computation Facility, SESCD, Wright-Patterson Air Force Base, Ohio. It is based on a procedure described under "Random Number Generation and Testing" in IBM manual C20-8011.
4. The program includes a Fortran IV sorting routine that is a duplicate of one used by J. B. Kruskal of Bell Telephone Laboratories in his non-metric, multidimensional, scaling program called MDSCAL. The remainder of the program is original with the authors.

(Accepted for publication October 10, 1967.)