COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 37, 20–53 (1987)

Nested Structures of Control: An Intuitive View

MICHAEL LEYTON *

Harvard University, Cambridge, Massachusetts 02138 and State University of New York, Buffalo, New York 14260

Received May 20, 1986

An intuitive exposition is provided of a theory (rigorously elaborated elsewhere) that perceptual organization is the description of phenomena as nested structures of control. The theory is used to explain a number of perceptual-organizational phenomena such a complex shape prototypification, generalized-cone analysis, the importance of curvature singularities, grouping phenomena, motion phenomena, etc. The nested-control theory maintains that any perceptual organization is given a stratification where each level of the stratification is operated on (1) asymmetrically, and (2) as a whole, by the level in which it is embedded. The levels-within-levels hierarchy is divided into two subhierarchies, one describing the stimulus set as generated via a sequence of deformations, and the other describing the initial prototype as generated from a subset of itself. The structure of the latter sub-hierarchy yields the grouping structure of the stimulus set. Principles are proposed that determine the way these two sub-hierarchies are interrelated: Symmetry axes in the grouping structure are converted into lines of flexibility in the prototype modifications. Using these concepts, solutions are offered to some classical problems in perceptual organization, and a theory of complex shape is developed.

1. INTRODUCTION

1.1. The Main Purpose

In a number of recent papers, I have argued that perception organizes stimulus sets into *nested structures of control* (Leyton [14–18, 21]). A nested structure of control is a stratification in which each level is operated on (1) asymmetrically, and (2) as a whole, by the level in which it is embedded. In addition, there are a number of further constraints both on the organization of any level and on the relationships between levels. I have argued that a wide variety of organizational phenomena—e.g., complex shape prototypification, generalized cone structure, the organization of curvature extrema, grouping phenomena, various motion phenomena–are all examples of this single type of highly restricted structure. The purpose of the present paper is to give an intuitive exposition of the argument, avoiding the mathematical concepts and notation of the previous papers. In addition, the last section presents new results on complex shape.

1.2. Separate Organizational Theories

Let us begin by observing that the field of perceptual organization, at present, contains a variety of quite *separate* theories which are constructed to handle apparently quite separate organizational situations. Fig. 1 represents a number of these theories by the situations that they were developed to handle. For example,

*Requests for reprints should be sent to Departments of Computer Science and Psychology, State University of New York, Buffalo, N.Y. 14260.

0734-189X/87 $3.00

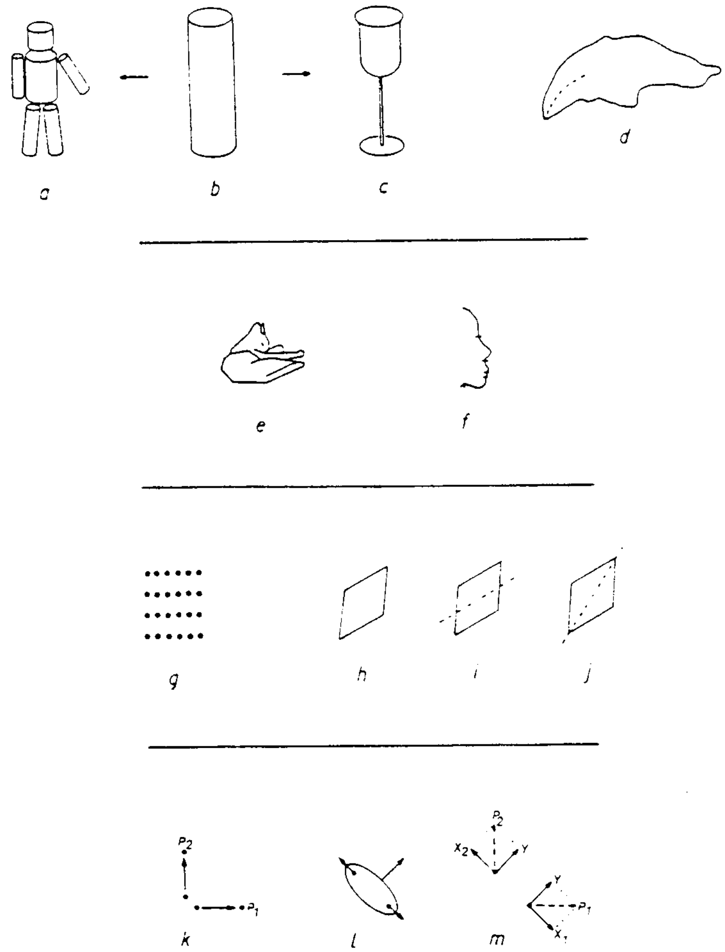


FIG. 1. (a–c) Multi-resolution hierarchies in which (b) a cylinder resolves into (a) a concatenation of generalized cylinders typical of a man, or (c) a concatenation of generalized cylinders typical of a goblet. (d) A local symmetry analysis of the head of a dolphin. (e) Attneave's cat. (f) The contour of a face segmented at negative curvature extrema. (g) A grid of dots grouped into rows by the proximity principle. (h–j) An illustration of the orientation and form problem: (k–m) The Johansson motion phenomenon.

Fig. 1a shows a complex shape, a man, described as the concatenation of generalized cylinders—an analysis advocated by Binford [2] and Marr and Nishihara [23]. The latter pair of researchers also propose a multi-resolution hierarchy of such analyses. Thus, at the coarsest level of the hierarchy, one has a single cylinder (Fig. 1b). This cylinder can resolve into the concatenation of cylinders typical of a man, as in Fig. 1a, or that typical of a goblet as in Fig. 1c. Analyses that more thoroughly describe the symmetry structure, upon which the former analysis must ultimately be based, are represented in Fig. 1d, where the curved symmetry axis (the dotted line) could have been created either by the symmetric axis transform (SAT) of Blum [3]; the smoothed local symmetry (SLS) of Brady [5]; or the process-inferring symmetry analysis (PISA) of Leyton [21]. Recently, Pizer, Oliver, and Bloomberg [25] have developed an algorithm that combines important aspects of the above approaches; i.e., it performs an SAT-based multi-resolution hierarchy.

On the second row of Fig. 1, we see research that has demonstrated the importance of curvature extrema in perception. For example, Attneave [1] showed that such extrema carry the maximal information of points along the contour; e.g., subjects, asked to represent a contour by a small number of points, choose these extrema. Attneave's famous drawing, shown in Fig. 1e, is constructed by connecting together curvature extrema using only straight lines. The figure is instantly recognizable as a cat.

Curvature extrema have been used in a different type of analysis by Hoffman and Richards [10] and Richards and Hoffman [27]. These researchers propose that a contour is perceptually segmented at points of negative curvature extrema (points of "maximal indentation"). For example, if one partitions the outline of a face, Fig. 1f, at such points, one obtains the crown of the head, the nose, the lips, and the chin; i.e., the segments turn out to be the natural parts of the figure. Hoffman and Richards give an explanation of this using differential topological concepts, and they also give a 5-fold classification of such segments—thus establishing a 5-letter alphabet out of which any curve can be constructed.

On the third row of Fig. 1, we see situations that were examined by the German gestalt movement (e.g., Wertheimer [30]). Figure 1g shows a rectangular grid of dots, which is grouped into rows due to the proximity principle. (Greater vertical proximity would have caused grouping into columns.) Again, in Fig. 1h, we see an illustration of another Gestalt phenomenon—the orientation and form phenomenon: That is, Fig. 1h has perceptual ambiguity since it can either be seen as a sheared square, Fig. 1i, or a stretched diamond, Fig. 1j. The different interpretations are due to the assignment of differently oriented Cartesian frames (see Goldmeier [9] and Rock [28], for further discussion).

Finally, in the bottom row of Fig. 1, we have the situation discovered and analyzed by Johansson [12]: A pair of point stimuli are moved perpendicularly from each other, in simple harmonic motion, and in phase; as illustrated in Fig. 1k. However, subjects do not see the stimulus set as doing this. Instead they see the two dots moving directly away from each other (see Fig. 1l), as if the dots are the ends of a diagonal rod. Furthermore, they see the "rod" as moving through space along the opposite diagonal. Johansson explained the phenomenon in this way: He argued that subjects resolve the motion vectors P_i, shown in Fig. 1k, into two sets of components: (1) the vectors X_i (see Fig. 1m), which constitute that part of the motion that is the dots' relation to each other, and (2) the vector Y (see Fig. 1m), which is that part of the motion that is common to the two dots.

1.3. Separate Perceptual Situations?

The several theories illustrated in Fig. 1 were developed quite separately under the assumption that the situations are very different from each other. We shall argue, in the present paper, that this assumption is not correct, and that a single set of principles will explain all these situations. The situations are, we shall claim, all examples of nested control structures in a highly restricted sense.

There are two ways in which this can be argued. On the one hand, one can propose a simple neuronal architecture that has a highly specific nested control structure and will nevertheless result in each of these situations when presented with the respective stimulus inputs. This approach was taken in Leyton [15]. In contrast, one can argue purely from *principles of representation*. It is this latter approach that we shall take in the present paper.

In fact, the research that will be elaborated is part of an ongoing investigation into *cognitive* representation generally (Leyton [16–18]). This program surveys four areas of cognition, (1) perceptual organization, (2) categorization, (3) linguistics, and (4) planning, to see if *general* principles of cognitive representation can be extracted from the research that is currently being done in those areas; i.e., to see if one can extract *domain-independent rules of representation*. For example, a survey of several highly developed theories of linguistic grammar reveals that, even though these theories are essentially constructed in opposition to each other, they all crucially hold to two basic principles: that *grammatical space is stratified* and that *each level is asymmetrically organized*. In fact, these two principles will also emerge in the perceptual organizational theory offered in the present paper. Thus, I have argued that these principles are domain-independent representational ones (see Leyton [16, 17], for a detailed exposition).

It therefore seems appropriate to arrange the present paper as follows: Part I elaborates the domain-independent representational principles while nevertheless illustrating them with perceptual phenomena. Part II develops the perceptual organizational theory from those principles. Finally, in Part III, we return to each of the apparently separate situations shown in Fig. 1, and explain them by this single organizational theory.

PART I

2. COGNITION AND MACHINES

The basic principle of this work is the proposal that cognition is an attempt to represent the environment as a collection of machines; i.e., as deterministic processes.

Observe that this proposal describes a relationship between cognition and machines that is very different from that which is currently held to be the basis for cognitive science. The currently held view can be called the *processing computer analogy*; and it can be expressed by saying that the cognitive *system* has the structure of a machine.

In contrast to this analogy, that which is being described here states that cognition structures the *environment* as a collection of machines. We shall call this the *representational* or *content computer analogy* (see Leyton [16–18], for detailed discussion).

Machines are currently studied by two major disciplines with respect to two very different structural interests: (1) Computer science studies a machine as a computational structure; i.e., as a structure of *operations* that completely determine behavior. (2) Dynamical systems theory studies a machine from the point of view of the *stability* of behavior.

We shall now investigate the representational consequences of the above two types of structure; i.e., the *operational* and *stability* structures of a machine. It will be argued that, in human representations, the two types interact to form nested systems of control, and that such systems determine the perceptual phenomena shown in Fig. 1. We begin by examining operational structure, and later incorporate stability considerations.

3. SYMMETRIC VS ASYMMETRIC REFERENCE FRAMES

3.1. Two Types of Operational Description

A machine consists of (1) an object, (2) a set of operations that can be applied to the object, and (3) the set of states into which the object is pushed under the operations.[1]

The set of operations satisfies two simple properties. (1) There is a trivial operation, e, which does nothing when applied to the ~~machine~~. This operation is object called an *identity element*. (2) The second property, associativity, is just as simple, but will be given only in footnote 2 as our discussion will not explicitly refer to it.[2]

Many machines, but not all, satisfy the following additional property: The effect of any operation can be reversed by applying some other operation—its "inverse." When this third property is satisfied, the collection of operations forms what is mathematically called a *group*. We will make the assumption that, in cognition, the set of operations, used to represent the environment as a deterministic structure or machine, is a *group*.

Let us now illustrate these notions. We will begin by considering how a simple stimulus set—a hexagon—can be structured as a machine. In fact, we will examine *two* ways in which this stimulus set can be so described. The two methods will be used to characterize two very different sets of *psychological assumptions*, as we shall see in the next section.

In both methods, the following three conditions will hold:

(1) the *object*, which is pushed into different states, is a *side* of a hexagon;

(2) the *states*, into which a side can be pushed, are six possible positions of a side in a hexagon;

(3) the *operations*, which force a side into the different positions, are of the following three types:

(i) clockwise rotations, r_θ, of a side, where the amount of rotation θ can be any multiple of 60°;
(ii) a reflection, t, which reflects a side about its bisecting axis;
(iii) all possible multiples of these operations.

This set of operations, has, in fact, twelve elements and these form a group called $\mathbf{D}_6$:

$$\mathbf{D}_6 = \{e, r_{60}, r_{120}, r_{180}, r_{240}, r_{300}, \\ t, tr_{60}, tr_{120}, tr_{180}, tr_{240}, tr_{300}\}.$$

The upper row lists the rotations by multiples of 60°. The lower row is the reflection, t, times each of those rotations.

METHOD 1. The first method of describing a hexagon as a machine is illustrated in Fig. 2a. The operations are represented by the arrows that go in between the sides. For example, given any side, there is an arrow emerging from it, labeled r_{60},

[1] One can regard any output as a part of the state (see, e.g., [4]).
[2] It is that a string of operations, $w_1 \cdot w_2 \cdot w_3$, can be bracketed in any way; that is $(w_1 \cdot w_2) \cdot w_3 = w_1 \cdot (w_2 \cdot w_3)$.

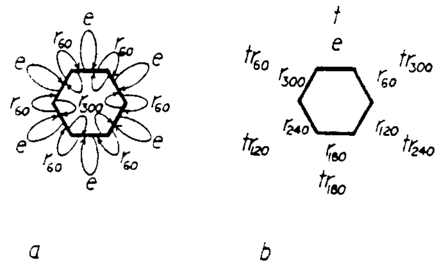


FIG. 2. (a) A symmetric description of a hexagon as a machine. (b) An asymmetric description of a hexagon as a machine.

that goes from that side to the next side in the clockwise direction. It is important to note that this arrow is supposed to represent a *perceived relationship* between the side and its neighbor. Again, given any side, there is an arrow marked r_{300} (drawn inside the figure) that moves from the side to the side that precedes it with respect to the clockwise direction. Similarly, given any side, there should be twelve arrows—corresponding to the twelve operations in $\mathbf{D}_6$—that emerge from the side and go to other sides. That is, there should be a total of 72 arrows in the diagram. However, for legibility, Fig. 2a shows only three arrows per side. (Note that the figure is, in fact, the *state-transition* diagram of the associated machine.)

, that about the perfectness, and concatenation?

METHOD 2. In the second method of describing a hexagon as a machine, a particular side is distinguished as the *initial* one from which the others are obtained by applying the operators. Figure 2b illustrates this method. In the figure, the top side has been chosen as the starting position. Thus the upper right side is obtained from the top one by applying r_{60}, rotation by 60°. In fact, the upper right side can be obtained from the top one also by applying tr_{300}. The important point about this method of description is that each side can be characterized by the operations that obtained it from the top side. Thus the operations can act as the descriptions of the side. In fact, one can say that *each stimulus is described by its referential relation to the top side*.

Before we compare the psychological implications of these two methods, we should observe that we are *not* going to be dealing only with highly regular and simple structures such as hexagons. One of the outcomes of this paper is an analysis of highly complex shape, such as the shapes of animals, plants, etc. shown in Fig. 17. However, it is necessary first to extract a number of principles from the analysis of regular structures.

3.2. *Different Psychological Implications*

Let us return to the two forms of description illustrated in Figs. 2a and b. What we see now is that they embody two very different sets of assumptions about psychological representations, as follows:

Consider Fig. 2a. Observe that any side in that figure is descriptively *identical* to any other side (e.g., it has the same set of arrows emerging from it as any other side). Thus, in this type of representation, the notion of *invariant* captures a crucial sloppy: side invariant under permutation of sides.

phenomenon; i.e., a side is an invariant under the action of the operations. Correspondingly, the representation is entirely symmetric or homogeneous.

It is important now to observe that this first type of description can be used to characterize the major existing transformational theories of perception, e.g., those of Piaget [24], Gibson [7, 8], and Hoffman [11], because these theories assume the fundamental importance of invariants and symmetrically acting transformations.

Let us now, in contrast, consider Fig. 2b. The first thing to observe is that, every side is descriptively *distinct* from every other side; e.g., only the upper right side carries the labels r_{60} and tr_{300}. Thus the notion of invariant does not have a significant role in this type of description. Correspondingly, the space is structured inhomogeneously.

but we could have started anywhere.

Intimately related to the inhomogeneity is the fact that reference is *asymmetric* in this form of description. For example, observe that, in Fig. 2b, the label r_{60} means "this side was obtained by the application of clockwise rotation by 60°." The label is therefore defined with respect to the top side as the starting position. However, consider the label, e, of the top side itself. Because the label says "do nothing," it is not defined with respect to the upper-right side. Thus the labeling is asymmetric: the upper-right side refers to the top side, but not vice versa.

This second method of description can be used to understand various representational theories outside perceptual psychology. For example, in Chomsky's Standard Theory [6], sentences are characterized by the transformations that generated them from the kernel, e.g., a question is characterized by the question transformation. Thus reference is asymmetric; e.g., a question is defined in terms of the kernel but not vice versa. Again, in Rosch's theory of categorization, non-prototypical members are referred to the prototypes but not vice-versa (e.g., Rosch [29]).

Because of the weight of evidence from linguistics and categorization, and the psychological data to be reviewed later, I have argued in Leyton [16, 17] that the second method of description is the appropriate one on which to base a theory of perceptual organization. More precisely, I have argued that it may be appropriate to invoke invariants in bottom-up *processing* theories, but that is not appropriate to invoke them in *representational* theories. Accordingly, the present paper will be based upon the second type of description. These issues are discussed in detail in Leyton [16, 17].

4. THE STRATIFICATION OF REFERENCE FRAMES

It is important to recall that, in our chosen method (Method 2), the operations act as *referential relations*, e.g., they refer sides to the top side. Thus the group G of operations can be regarded as a collection of referential relations. In fact, because the members of a group are inter-related (e.g., r_{120} is twice r_{60}), the group G forms

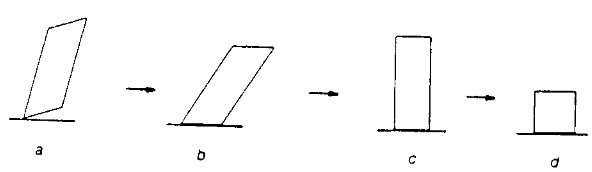


FIG. 3. An example of successive reference.

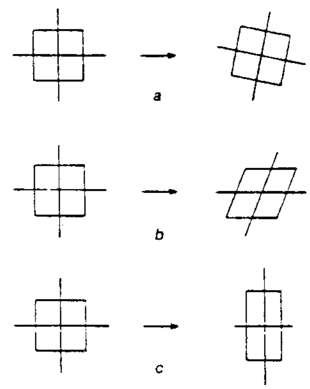


FIG. 4. (a) Rotation; (b) shear; (c) stretch.

an *interconnected system* of referential relations. We can thus consider the group to be a *reference frame* that organizes the stimulus set; i.e., that defines the referential structure between its members.

What I will argue now is that, in human cognition, a reference frame G is given a *stratification*, in a particular way to be illustrated as follows:

In a converging series of five experiments, I found that subjects, when presented with a rotated parallelogram, Fig. 3a, reference it to a non-rotated one, Fig. 3b, which they then reference to a rectangle, Fig. 3c, which they then reference to a square (Fig. 3d; see Leyton [17], for details[3]).

Let us interpret these results in terms of the system we have been developing. It is reasonable first to assume that the relationship between the initial stimulus, the rotated parallelogram, and the last stimulus, the square, is given by a linear transformation. (A linear transformation is simply a transformation in which straight subspaces are mapped to straight subspaces; e.g., straight lines through the origin are mapped to straight lines through the origin.) Thus the group of possible relations between the first and the last stimulus can be assumed to be the group **Lin** of linear transformations.[4]

Now, it is a mathematical fact (e.g., Lang [13]) that any linear transformation can be decomposed into three primitive linear transformations that are illustrated in Fig. 4. The first is simply a *rotation* (Fig. 4a). The second is a *shear*; e.g., it is created by pulling the top and bottom sides of a square horizontally in opposite directions (Fig. 4b). The third operation is a *stretch* along perpendicular directions (Fig. 4c). Thus the group **Lin** can be decomposed as follows:

$$\mathbf{Lin} = \mathbf{Stretches} \cdot \mathbf{Shears} \cdot \mathbf{Rotations}.$$

[3] For example, in one of the experiments, subjects were presented with only the non-rotated parallelogram, Fig. 3b, and were asked which other figure came to mind. Their answer was a rectangle. When presented, then, with a rectangle, they were asked the same question again, and this time their answer was a square. It should be observed therefore that the subjects dictated the sequence.

[4] In fact, we will make the assumption that **Lin** is the group of linear transformations without reflection, i.e., the group that is usually denoted by $\mathbf{GL}_2^+ R$.

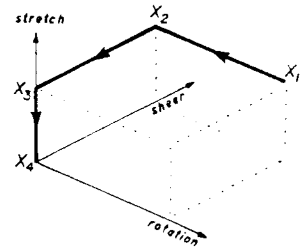


FIG. 5. The successive reference, in Fig. 3, modeled as a trajectory in **Lin**.

Therefore, **Lin** can be represented by the 3-dimensional space shown in Fig. 5, where the three axes correspond to (1) the amount of stretch, (2) the amount of shear, and (3) the amount of rotation.[5]

In this space, the initial stimulus, the rotated parallelogram, can be represented by some point x_1 away from the origin (since the shape has non-zero values in each dimension). Furthermore, the first stage of reference (rotated parallelogram → non-rotated parallelogram) can be modeled by the bold trajectory that goes from x_1 to x_2 in Fig. 5, because the point x_2 has zero value on the rotation axis, and therefore represents the non-rotated parallogram. Similarly, the next stage of reference (parallelogram → rectangle) removes shear and can be represented by the bold trajectory that goes from x_2 to x_3, the latter point having zero value on the shear axis and thus representing the rectangle. Similarly, the final reference stage (rectangle → square) removes stretch and can thus be represented by the trajectory from x_3 to x_4 where the square is situated.

The structure we have just described, **Lin** = **Stretches** · **Shears** · **Rotation**, is an example of the general structure that will be important to our perceptual organizational theory: A group G, such as **Lin**, is given a stratification $G = G_1 \cdot G_2 \dots G_n$ into levels where each level G_i is itself a group.[6] Reference, in the group, proceeds not in one step to the overall reference point (identity element), but *successively*, first through level G_n to its own reference point e_n, then through G_{n-1} to its own reference point e_{n-1}, and so on down to G_1 and to its own reference point e_1.

An important question however needs to be answered: What determines (1) the decomposition, and (2) the order of removal?

To answer this question, recall that, in Section 2, we noted that machines, as they are studied today, are investigated essentially with respect to two kinds of structure, their operational (computational) structure and their stability structure. We have, up till now, been concentrating only on operational aspects. The next section begins to consider stability.

5. STABILITY AND REFERENCE FRAMES

To answer the question of what determines the decomposition $G = G_1 \cdot G_2 \dots G_n$, and the order in which the components are removed, I will propose the following principle which is based upon stability considerations.

[5] The amount of stretch can be the determinant of the stretch matrix; the amount of shear can be the non-zero off-diagonal member of the shear matrix; and the amount of rotation can be the angle of rotation.

[6] More precisely, any member g of G can be written as a product $g = g_1 \cdot g_2 \dots g_n$, where $g_i \in G_i$.

COGNITIVE STABILITY PRINCIPLE. *The cognitive system removes the properties of a stimulus in sequence starting from those properties ranked as most unstable and progressing successively to those properties ranked as most stable.*

Our claim is that this principle answers our question concerning decomposition and ordering: That is, the decomposition $G = G_1 \cdot G_2 \ldots G_n$ is a stratification of G into *levels of stability*, and the order of removal is that which starts from the most unstable and progresses to the most stable.

what is stability - mechanical?

To corroborate this, let us return to the decomposition **Lin = Stretches · Shears · Rotations**. In a further series of experiments (Leyton [17]), I found that subjects rated rotation to be the most unstable property of the initial percept, shear to be the next most unstable, and stretch to be the most stable. Thus the order of removal is that predicted by the above principle.

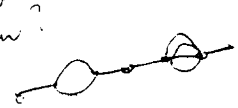
The general structure that has been developed here—a decomposition $G = G_1 \cdot G_2 \ldots G_n$ of a group into stability levels, where there is a removal of these levels in order of increasing stability—is the central concept of this paper. We will see in Part III, that the several types of perceptual organizations, listed in Fig. 1, all are explained as examples of this structure. A rigorous definition of such structures, in terms of Lie groups and fiber bundles is given in Leyton [16, 17]. Furthermore, in Leyton [16], extensive examples of this type of structure in other branches of cognitive science are documented.

6. THE PRINCIPLE OF NESTED CONTROL

We end our discussion of the general cognitive principles by showing that a structure with the above type of stability stratification has an important property: it forms a *nested system of control*. By this, we mean that, given a single operation in one level, the operation acts on the *entire* level below it; i.e., it transforms that level *as a whole*. We shall start illustrating this concept within a few pages, when we begin to look at perceptual organization in detail. However, at this point, a general explanation of the phenomenon is offered. Since the explanation is rather abstract, some readers may wish to omit it—and they can do so without loss of comprehension to the rest of the paper.

We proceed by induction. Consider an example of the above type of stability stratification, where the decomposition is of order two; i.e., one has $G = G_1 \cdot G_2$. The cognitive system will reference down a trajectory in the second factor G_2, to the identity element e_2 of G_2, before any significant movement has occurred in the first factor G_1. This means that during this first reference stage, the system holds a fixed value g_1 in G_1. Thus, to restate this situation in a slightly different way: Given each value g_1 in G_1, there is a particular trajectory in G_2 that is associated with g_1. In other words G_1 parameterizes trajectories on G_2. Now, in control theory, a variable that parameterizes trajectories on a space is called a *control parameter*. Thus, G_1 acts as a control parameter with respect to trajectories on G_2. By induction, therefore, we have that, for a stability stratification $G = G_1 \cdot G_2 \ldots G_n$ of length n, each factor G_i is cognized as controlling the trajectories on the subsequence $G_{i+1} \cdot G_{i+2} \ldots G_n$ to the right of G_i in the overall sequence. Hence, we conclude that a stability stratification $G = G_1 \cdot G_2 \ldots G_n$, of the above type, is cognized as a *nested structure of control*.

what about parallelism?



PART II

The above discussion elaborated a set of general principles of cognitive representation. We can now turn to an analysis of perceptual organization.

7. PERCEPTUAL ORGANIZATION—THE BASIC PROPOSAL

In this section, the basic proposal of the perceptual organizational theory will be presented. For explanatory purposes, it is advisable to introduce the proposal using a very simple example, a square. Part III will give detailed illustrations using *complex* organizational phenomena.

Observe, from the proceeding discussion, that we can identify *two* generative descriptions associated with a square:

(1) The first is the analog of the description given for the hexagon in Fig. 2b. That is, a square can be generated from a side by the application of the eight rotation and reflection operations of $\mathbf{D}_4$, just as a hexagon was generated from a side by the twelve rotation and reflection operations of $\mathbf{D}_6$ (Sect. 3.1). A description of this type will be called an *internal* description because the group $\mathbf{D}_4$ generates the figure from a subset of itself: i.e., from *inside* itself.

(2) The second type of description is **Lin** = **Stretches** · **Shears** · **Rotations**, as illustrated in Fig. 3. Here, the operations act on the square *as a whole*. Because of this action on the whole percept, **Lin** will be said to form an *external* description. The external group can describe *deformation*, or, as we shall see later, it can describe *motion*.

Now observe that, in relation to the external group, **Lin** = **Stretches** · **Shears** · **Rotation**, the square is a *prototype*. Thus the internal group is responsible for generating the prototype; and the external group then acts on the prototype, changing it into a non-prototypical version (e.g., the rotated parallelogram).

Because the internal group generates the prototype, the group must use only particularly simple types of operations: those that generate spheres, polygons, cylinders, etc.—shapes that are classically regarded as prototypes. We shall claim that this means that the internal group is always a subgroup of what is called the Euclidean group; the group consisting of translations, rotations, and reflection. → Scaling

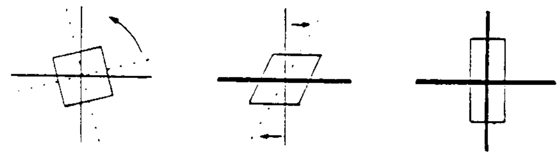
The basic proposal of our perceptual organizational theory can now be stated:

PROPOSAL 1. *A perceptual organization is the* **interaction** *of two generative descriptions*:

(1) *One is given by an* **external group**, *which generates the stimulus set from a prototype*.

(2) *The other is given by an* **internal group**, *which* (*a*) *generates the prototype from a subset of itself and* (*b*) *is a subgroup of the Euclidean group*.

The following sections will develop an understanding of this proposal. For example, we shall see that an examination of the internal generative structure leads to a theory of *grouping*. Again, we shall examine how the internal and external structures interact, and shall see that the interaction is given by a particular law that is a powerful determinant of each of the phenomena shown in Fig. 1.



a *b* *c*

FIG. 6. The bold lines represent the eigenspaces of the groups (a) rotation, (b) stretch, and (c) shear.

8. WHY THE ORDER: ROTATION, SHEAR, STRETCH?

Let us now try to understand why rotation was rated to be less stable than shear, and shear to be less stable than stretch.

Consider a linear transformation of a plane. Under such a transformation, lines move and undergo various actions. However, it can be the case that some lines do not change their position. If such a line goes through the origin, it is called, in mathematics, an *eigenspace-line*; i.e., an eigenspace-line is a line through the origin that does not change its position during a linear transformation. The concept will be illustrated shortly. One should note that some transformations have more such lines than others. The dimension of the space spanned by these lines is called the *eigenspace-dimension*.

What is important to observe is that, because eigenspace-lines are subsets that do not change position, the eigenspace-*dimension* can be regarded as one possible measure of the amount of *stability* associated with the transformation.

With this concept in mind, let us return to the three transformations we have been considering, and establish their respective eigenspace-dimensions. The transformations are shown in Fig. 6. We see, from Fig. 6a, that every line changes position under a rotation. Thus the eigenspace-dimension, in this case, is *zero*. Now consider the shear in Fig. 6b. In this diagram, every line through the origin has changed position except the single line that is the x axis. Thus the dimension of the eigenspace, for a shear, is *one*. Finally, in Fig. 6c, there are two lines, through the origin, that have not changed position: the x and y axes. Thus the eigenspace-dimension, for stretch, is *two*.

well, 1 complex

This means that the order chosen by the subjects—rotation, followed by shear, followed by stretch—is the order of the sizes of the respective eigenspace-dimensions; i.e., zero, one, two. Thus the stability ranking by the subjects accords with the stability ranking based upon eigenspace-dimension. Further evidence for the eigenspace basis of stability will emerge later from other considerations.

9. THE MEANING OF PROTOTYPE

Recall that in Proposal 1 (Sect. 7) it was claimed that a perceptual organization of a stimulus set S is the interaction between an external group G_E that generates S from a prototype of S, and an internal group G_I that generates the prototype from a subset of itself. We have begun to see that G_E has a decomposition sequence (e.g., **Lin = Stretches · Shears · Rotations**) that yields a nested structure of control. What we shall now begin to show is that G_I also has such a structure—and that the latter structure has crucial organizational consequences.

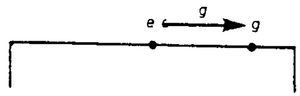


FIG. 7. The representation of a polygon-side as asymmetrically related to the central point.

At this stage it is advisable to consider again a simple percept, a square, for explanatory purposes. *Complex* organizations will be examined in Part III.

We begin by noting that a square is perceived as a nested structure of control, in the following obvious sense: On the first level of organization, a square is simply a collection of points; on the next level, the points are bound together into sides; and on the next level, the sides are bound together into the whole, the square.

However, the situation is much more subtle than this: Each level is in fact an *asymmetric operational structure*, as follows:

(1) It was shown in Leyton [18] that the points of a side are perceptually referenced to the central point on that side. This phenomenon can be captured as illustrated in Fig. 7: The side is structured by the (1-dimensional group) **Translations**, acting on the central point. Thus the central point is characterized by the group identity element, e, and any other point is described as a translation g from the central point. Note that this structure captures the *asymmetric* reference of points with respect to the central point.

(2) Moving up one level, each of the sides can be regarded as generated from the top side by $\mathbf{D}_4$, in the same way that $\mathbf{D}_6$ generated the hexagon from its top side. Experimental support for this view is provided in Leyton [18], where it was found that subjects define the orientations of the sides with respect to the orientation of the top side.

is that so?

Thus the perceptual structure seems to be given by a hierarchy of groups: **Translations** starts with a point and generates a side, and $\mathbf{D}_4$ starts with a side and generates the whole. This sequence of groups, is given thus:

$$\mathbf{D}_4 \cdot \mathbf{Translations} \cdot \mathbf{Points}.$$

Note that the left-to-right order is the top-to-bottom order in the perceptual organization.

Observe also that this sequence is a *nested structure of control*, as follows: The group $\mathbf{D}_4$ acts on **Translations**, moving the latter around; the group **Translations** acts on a **Point** moving the latter around. This is illustrated in Fig. 8. The two points shown on the top side are related by the translation g. The same is true of the two points shown on the right side. The circular arrow is a rotation, in $\mathbf{D}_4$, which sends the translation g on the top side onto the translation g on the right side. In other words, the circular arrow shows that the relationship between the top pair of points corresponds to the relationship between the side pair.

The above discussion therefore illustrates a concept we will see validated throughout this paper: prototypes are nested structures of control. In fact, more

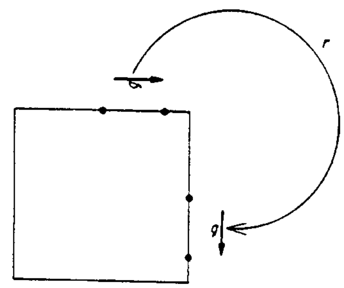


FIG. 8. The nested control hierarchy of a square.

precisely, we have

PROPOSAL 2. (1) *a* **prototype** *is a nested structure of control, where each level of control is a subgroup of the Euclidean group (i.e., translations, rotations, reflections).* (2) *A* **shape prototype** *is a* **connected** *nested structure of control by Euclidean subgroups.*

By *connected*, in part (2) of the proposal, we mean that any two points in the organization can be connected by a continuous path which itself consists only of points in the organization. Clearly, a square satisfies part (2) of the proposal because any two points of a square are path-connected, and the square is a nested hierarchy of subgroups of the Euclidean group. Other stimulus sets that satisfy part (2) are shown in Fig. 9. One can see from these figures that the proposal does indeed seem to yield what are psychologically taken to be shape prototypes.

how is connectedness (of closure) guaranteed

10. THE MEANING OF GROUPING

The most important point about internal structures can be introduced by considering the simple example in Fig. 10. It shows a rectangular grid of dots that can be perceived either as a set of rows (Fig. 10a) or as a set of columns (Fig. 10b). In fact, as we shall shortly see, both percepts involve the *same* groups. However the central difference between the two percepts is that these groups are *ordered* differently in

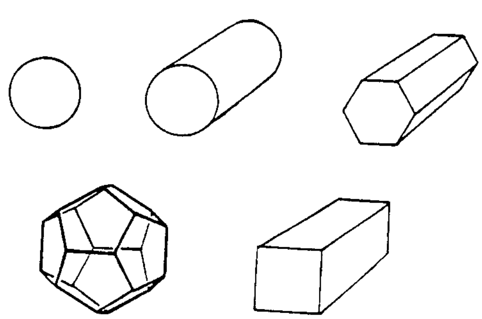


FIG. 9. A collection of shape prototypes, according to Proposal 2.

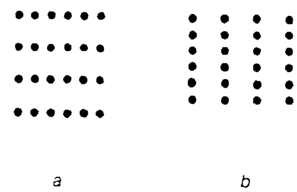


FIG. 10. A grid of dots grouped into rows or columns.

the respective nested control hierarchies, as follows:

Let **Horizontal** be the group of horizontal translations by regular intervals, and **Vertical** be the group of vertical translations by regular intervals. Then Fig. 10a is generated as follows: Start with a point. Apply **Horizontal** to generate a row. Then apply **Vertical** to generate the set of rows. In contrast, Fig. 10b is generated as follows: Start with a point. Apply **Vertical** to generate a column. Then apply **Horizontal** to generate the set of columns. That is, the nested control structure in Fig. 10a is

Vertical · **Horizontal** · **Point**

whereas the nested control structure in Fig. 10b reverses the order in which the translation groups are applied, thus

Horizontal · **Vertical** · **Point**.

The question that must now be settled is this: What criterion makes **Vertical** · **Horizontal** · **Point** the choice in the first case, and **Horizontal** · **Vertical** · **Point** the choice in the second case? It will be claimed that the criterion that is used is this:

Similarity

STABILITY EVALUATION PRINCIPLE. *Higher levels of control are assigned to stimuli that are more distant from each other, and lower levels of control are assigned to stimuli that are nearer to each other.*

For example, because the column stimuli are more distant from each other than the row stimuli in Fig. 10a, the group **Vertical** controls (i.e., "moves around") the group **Horizontal**. Correspondingly, the levels of control are reversed in Fig. 10b.

The above principle explains several of the Gestalt grouping criteria. In fact, in Leyton [16, Sect. 5.5], I list eleven of the Gestalt criteria and argue that half of them are explained by this principle. How are the remaining half explained?

Recall that Part I of this paper began by noting that machines embody two kinds of structure: operational structure and stability structure. The above evaluation principle corresponds to the latter type. There is in fact another evaluation principle that corresponds to the former type. This other principle is based on the rather technical notation of a group presentation, and will therefore not be given here. However, it should be noted that a conclusion in Leyton [17] is that half of the Gestalt principles are explained by an *operational* evaluation principle, and half by a *stability* evaluation principle.

Let us return to the *stability evaluation principle*. It is important to observe that this principle involves a more profound phenomenon than is described by the

Gestalt criteria. For example, observe first that the Gestalt criterion of proximity is derivable from the principle (the criterion states: proximal stimuli are grouped over distant stimuli). In fact, it is the proximity criterion that has been conventionally used to explain the percepts resulting from Figs. 10a and b. However, the use of the proximity criterion does not recognize the fact that both percepts are in fact *nested structures of control*. That is, separate groupings on one level are themselves elements that are grouped on the next level, and so on.

Thus the Gestalt school failed to realize that *grouping* is ultimately *a nested control phenomenon*. In fact, our system provides a theory of grouping in the following way: The internal group of a perceptual organization is given a decomposition $G = G_1 \cdot G_2 \ldots G_n$. The stability evaluation principle states that the ordering on the decomposition is such that it assigns more distant stimuli to the higher G_i and more proximal stimuli to the lower G_i. Finally, we require the following proposal:

PROPOSAL 3. *The* **perceptual groupings** *of a stimulus set are those subsets that are described by the right subsequences of the nested control structure* $G_I = G_{I_2} \cdot G_{I_2} \ldots G_{I_n}$.

A right subsequence is one that starts at some point along the sequence and continues to the right-hand end. To illustrate the proposal, consider our description of Fig. 10a:

Vertical · Horizontal · Point.

The right subsequences, in this case, are (1) **Point**, (2) **Horizontal · Point**, and (3) **Vertical · Horizontal · Point**. These correspond respectively to the *points*, the *rows*, the *whole*—which, of course, are the psychological groupings. In contrast, as the reader can easily check, the right subsequences of the description **Horizontal · Vertical · Point**, of Fig. 10b, yield the following sets: the *points*, the *columns*, the *whole*—which, of course, are the psychological groupings in that organization.

11. THE INTERACTION LAW

Recall that Proposal 1, in Section 7, claims that a perceptual organization, of a stimulus set, is an *interaction* between an external group, G_E, that generates the set from a prototype, and an internal group, G_I, that generates the prototype from a subset of itself. We now derive a crucial principle that determines this interaction.

By definition, the external group *acts* on the prototype generated by the internal group. For example, **Lin = Stretches · Shears · Rotations** *acts* on the prototype, a square, which is generated by $\mathbf{D}_4$. Now the interaction principle, to be derived, rests on the following proposal: a transformation in G_E (e.g., a deformation) can be regarded as more stable if it is more *structure-preserving* when it acts on the prototype; i.e., on G_I. What we must do therefore is to understand the nature of structure-preservation. Observe first that the operations in G_I that create the prototype (rotations, reflection) are associated with symmetry axes (i.e., rotational and reflection axes). Thus we can regard structure-preservation of G_I as preservation of symmetry axes. Now, an axis is preserved if it is either (1) rotated, (2) translated, or (3) left over itself. Furthermore, one can regard the third alternative as the least altering. However, a line that is left over itself is an *eigenspace*. In fact, as we shall see, such a line is seen as a line of flexibility. Thus we obtain the following principle:

INTERACTION LAW. *The symmetry axes of the internal group become the eigenspaces ("lines of flexibility") of the most stable external groups.*

To Illustrate. The most stable group in **Lin** = **Stretches** · **Shears** · **Rotations** is the factor **Stretches**. However, we can see from Fig. 3 that, under the transformation from the square to the rectangle, the symmetry axes of the square become the eigenspaces (lines of flexibility).

It turns out that the interaction law explains a considerable variety of perceptual organizational phenomena, both in motion and complex shape perception—as we shall see in Part III of the paper.

A Notational Point. We shall use the symbol $G_E \hat{\imath} G_I$ to describe this relationship between the external and internal groups; that is, $\hat{\imath}$ indicates that the symmetry axes of G_I are the eigenspaces of the most stable groups in the G_E-sequence.

12. CARTESIAN FRAMES

To conclude Part II, we need one further concept. Cartesian reference frames have been invoked as crucial in a number of perceptual organizational theories. For example, in the Marr and Nishihara [23] analysis of complex shape, an important stage is the assignment of Cartesian frames to regions of the 2-dimensional image; e.g., these define what will be taken to be the limbs of an animal. Again, in the orientation and form phenomenon (Figs. 1h, i, j), it is assumed that different percepts arise from the same figure because a Cartesian frame has been assigned differently in each case.

It is currently supposed that Cartesian frames are trivial structures. Nevertheless, no one has attempted to explain why they have such "magical" effects on stimulus sets. For instance, the shape examples just given are completely unexplained. In the present paper, an attempt will be made to develop an understanding of why Cartesian frames have the effects that they do. In the process of developing this understanding, we shall see that Cartesian frames are not the trivial structures that they have been assumed so far to be.

12.1. The External Structure of Cartesian Frames

If the reader were asked to draw a Cartesian frame, it is highly likely that he or she would draw one that possessed the following three properties: (1) the x axis of the frame would be horizontally aligned; (2) the x and y axes would be perpendicular to each other, and (3) units along the x axis would tend to be equal to units on the y axis. Furthermore, it is important to observe that, when one sees a frame that does not have these three properties, one references it to a frame that does have these properties. That is, a frame that is not horizontally aligned, or does not have perpendicular axes, or has axes where the y units are twice the length of the x units, is referenced to the above type of frame.

Such reference must, of course, be created by three operations: rotation, shear, and stretch. Thus we are beginning to see that the external structure of a Cartesian frame cognitively involves the reference structure found earlier in the experiments involving the rotated parallelogram (Fig. 3). In fact, later it will be claimed that the

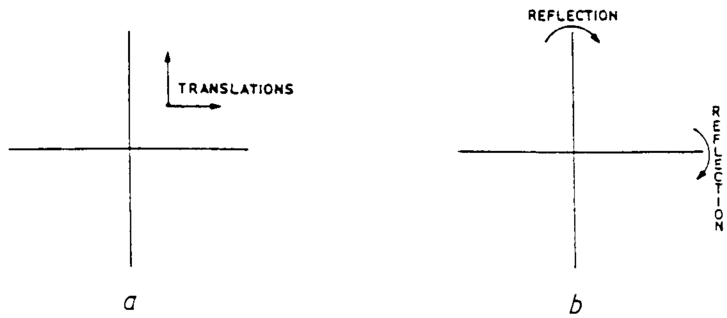


FIG. 11. The Cartesian plane *internally* generated either by translation or by reflections.

reference results for the rotated parallelogram were due to the imposition of a Cartesian frame and the fact that it is the *latter* that has the external structure **Lin = Stretches · Shears · Rotations.**

Let us therefore define the **external structure of a Cartesian frame** to be the structure **Lin = Stretches · Shears · Rotations**, with the factorization ordering defined by increasing eigenspace-dimension. This concept will be very useful later, for example, in the analysis of complex shape.

12.2. The Internal Structure of Cartesian Frames

The next thing to observe is that, just as with shape prototypes, the prototype of a Cartesian frame has an *internal* structure; i.e., it can be generated from a subset of itself. In fact, it has at least two such structures, as follows:

(1) As illustrated in Fig. 11a, a Cartesian frame can be generated from a point by translating the point through all distances in both the x and y directions; that is, by applying the 2-dimensional group, **Translations** $\times$ **Translations**. This is essentially the structure that was defined by Descartes.

(2) However, there is another structure that is involved, and this will be more important to us. As shown in Fig. 11b, the whole plane can be generated from a quadrant by reflections about the x axis and reflections about the y axis; that is, by applying the group **Reflection** $\times$ **Reflection**.

The groups defined in (1) and (2) in fact interact: **Reflection** $\times$ **Reflection** makes copies of **Translations** $\times$ **Translations** such that the directions of the axes are reversed. In mathematics, when one group makes automorphic copies of the other, they can be combined by a product called a semi-direct product denoted by $(\times_{sd})$. Thus we say that the entire internal structure of the Cartesian frame is

$$(\textbf{Reflection} \times \textbf{Reflection})(\times_{sd})(\textbf{Translations} \times \textbf{Translations}).$$

The two groups, just described, **Reflection** $\times$ **Reflection** and **Translations** $\times$ **Translations**, are subgroups of the Euclidean group. The Euclidean group has a remaining subgroup that also internally generates the frame: The group **Rotations** generates the frame from one of its half-axes. Thus we should think of the internal structure of the Cartesian frame as being the entire Euclidean group. However, to avoid confusing the use of **Rotations** as an internal group, with its use as an external one in **Lin**, we shall omit it when talking about the internal structure of the frame.

12.3. The Full Cartesian Frame

We now have two structures for a Cartesian reference frame: (1) an external structure **Stretches** · **Shears** · **Rotations**, and (2) an internal structure (**Reflection** × **Reflection**)($\times_{sd}$)(**Translations** × **Translations**). Furthermore, these structures can be locked together using the interaction law; that is, ensuring that the symmetry axes in the latter system becomes the eigenspace-lines of the former system. Notationally, the resulting structure is

(*)

$$(\mathbf{Stretches} \cdot \mathbf{Shears} \cdot \mathbf{Rotations})\,\hat{\imath}\,((\mathbf{Reflection} \times \mathbf{Reflection})(\times_{sd})(\mathbf{Translations} \times \mathbf{Translations})).$$

In Part III, we shall see that this system of operations is integral to the phenomenon of perceptual organization.

12.4. Imposing a Cartesian Frame

Recall now that we have been developing two structures in our analysis of perceptual organization: (1) The first is the structure $G = G_1 \cdot G_2 \ldots G_n$, which is the sequence of groups defining an organization as a nested system of control. The first j factors constitute the external sequence that alters the organization, and the remaining factors constitute the internal sequence that gives the grouping hierarchy. (2) The other kind of structure is the Cartesian reference frame, which, we have argued, is given by (**Stretches** · **Shears** · **Rotations**)$\hat{\imath}$((**Reflection** × **Reflection**)($\times_{sd}$)(**Translations** × **Translations**)).

The question that will now be addressed is what it can mean to *impose* a Cartesian frame on a stimulus set. Conventionally, the imposition of such a frame is understood to be an act rather like this: A pair of axes are drawn on some transparent sheet and the sheet is lain over the stimulus set. Magically, as a consequence of this act, the stimulus set becomes structured in a certain way that leads to perceptual results, such as the identification of limbs using generalized cylinders, and the orientation and form phenomenon. A new view of the imposition of Cartesian frames will now be given, such that these results no longer appear magical.

We have two group sequences, $G = G_1 \cdot G_2 \ldots G_n$ and (**Stretches** · **Shears** · **Rotations**)$\hat{\imath}$((**Reflection** × **Reflection**)($\times_{sd}$)(**Translations** × **Translations**)). The crucial claim to be made is this: *The imposition of a Cartesian reference frame means that the latter group sequence becomes a subsequence of the former*. A helpful metaphor, with which to understand this concept, is that of genetic splicing: The latter sequence is *spliced* into the former.

What does this mean in perceptual organizational terms? First, it means that the operations that describe the frame *generatively* are a subset of the set of operations that describe the organization generatively. In particular, this means: (1) The external structure of the perceptual organization (e.g., the prototypification or motion structure) contains the operations **Stretches** · **Shears** · **Rotations**. Furthermore, it means: (2) The internal structure of the perceptual organization (i.e., the grouping structure), contains the operations (**Reflection** × **Reflection**)($\times_{sd}$)(**Translations** × **Translations**). This view will be corroborated several times in Part III.

(*) implicit seems to be 2D. is it visual impression to s[illegible] organized?

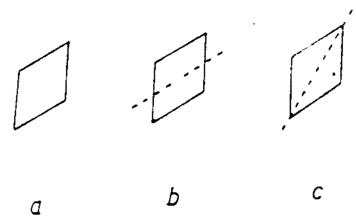


FIG. 12. The figure in (a) can be interpreted either as a sheared square (b) or as a stretched diamond (c).

PART III

Our intention now is to go back to the examples given in Fig. 1 and show that they are not perceptually separate situations, but that they are all organized according to the same principles; i.e., as nested structures of control in the highly restricted sense described in Part II.

13. SOLUTION TO THE ORIENTATION AND FORM PROBLEM

We shall begin by looking at the orientation and form problem—a problem that has remained unsolved for the 90 years since it was discovered by Mach [22]. In fact the problem, as it exists today, is a large collection of inexplicable data.

The first point to make is that the characterization "orientation and form phenomenon" is based on a current misunderstanding of that phenomenon. Consider the following simple example: Fig. 12a has perceptual ambiguity. It can be viewed either as a sheared square, Fig. 12b, or as a stretched diamond, Fig. 12c. This means that, not only is orientation involved in creating the ambiguity, but stretch and shear are also involved. In fact, using the concepts defined in the present paper, we propose

PROPOSAL 4. *What has conventionally been called the* **orientation and form phenomenon** *is this: Different assignments of the external group to a stimulus set prescribe different assignments of the internal group.*

The external group corresponds to what has conventionally been called "orientation" or rotation—but, as we have seen, it involves other operations. The internal group corresponds to what has conventionally been called "form"—which, according to our system, is the generative structure of the associated prototype.

Recall now that, according to the last section, the imposition of a Cartesian reference frame is the inclusion of the structure $(\mathbf{Stretches} \cdot \mathbf{Shears} \cdot \mathbf{Rotations}) \wr ((\mathbf{Reflection} \times \mathbf{Reflection})(\times_{sd})(\mathbf{Translations} \times \mathbf{Translations}))$ into the overall generative structure of the figure. One of the benefits of this view is that it predicts what we have just observed: that, by imposing a Cartesian reference frame, one is going to have more problems than with just orientation; i.e., stretches and shears are also going to be crucial—hence one obtains a phenomenon like Fig. 12. Before we can consider these factors any further it is necessary to show how our theory of perceptual organization presents a solution to the orientation and form problem.

Consider again the above statement of the problem: Different assignments of the external group prescribe different assignments of the internal group. Why should this be the case? The answer to this question is given, in fact, by the interaction law, for the law states that the external and internal groups *interact*. Hence, if one determines one of the two groups, the other is thereby determined. Furthermore, the interaction law predicts exactly how this co-determination takes place: The eigenspaces of the external group are the symmetry axes of the internal group.

This means that if we define the external actions involved in, for example, the two interpretations, Figs. 12b and c, then we can predict the respective prototypes in each case, as follows: The interpretation, Fig. 12b, is created by a shear that has an eigenspace bisecting a pair of opposite sides (i.e., the line shown). However, the interaction law converts eigenspaces into symmetry axes. This means that the symmetry axes of the prototype must have been bisectors of opposite sides. Thus the prototype must have been a square. Conversely, Fig. 12c is created by a stretch that has its perceived eigenspaces bisecting pairs of opposite *angles*. By the interaction law, eigenspaces are converted into symmetry axes, which means that the symmetry axes of the prototype must have been angle bisectors. Thus the prototype must have been a diamond. (Perceptually, the difference between a square and a diamond is that the salient symmetry axes of the former are side bisectors and the salient symmetry axes of the latter are angle bisectors; see Leyton [18]).

Finally, I show in Leyton [16, Sect. 5.5] that a variant of the stability evaluation principle (Sect. 10 of this paper) predicts the extent to which one interpretation is chosen over the other, as follows: Observe, in Fig. 12b, that although shear has a non-zero value, the value of stretch is zero. Conversely, in Fig. 12c, although stretch has a non-zero value, the value of shear is zero. Now, since stretch is a more stable parameter than shear, this means that the second interpretation should be chosen over the first because the larger value is assigned to the more stable parameter, stretch. However, observe that the second interpretation has slightly greater rotation than the first; and rotation is more unstable than shear and stretch. This means that the two interpretations are roughly equal with respect to stability ranking; which means that the two interpretations will be chosen to approximately an equal number of times. Finally, observe that if the amount of rotation were changed so that it would be zero either in the first or in the second case, then the case in which the rotation were zero would be the chosen interpretation, because rotation would be significantly large in the other case (see Figs. 13a and b).

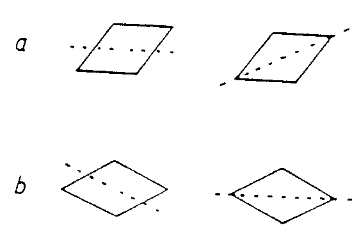


FIG. 13. The shape in Fig. 12, re-oriented such that the eigenspace of shear is horizontal in (a), and the eigenspace of stretch is horizontal in (b).

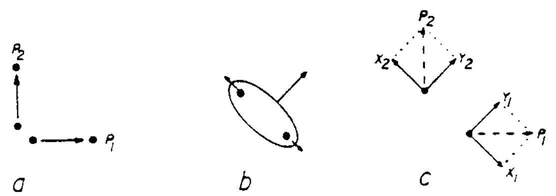


FIG. 14. The Johansson [12] motion phenomenon.

14. THE JOHANSSON [12] MOTION PHENOMENON

The Johansson [12] motion phenomenon is shown again in Fig. 14. Two points are moved perpendicularly from each other, in simple harmonic motion, and in phase (Fig. 14a). However, the subject does not see them as such. Instead, as shown in Fig. 14b, what is seen is the two points moving directly away from each other along one diagonal, as if the points were the ends of an invisible rod that is changing length; furthermore the "rod" is seen to be moving, as a whole, along the opposite diagonal. The analysis offered by Johansson [12] is this: The velocity vectors P_i (in Fig. 14a) are partitioned into (1) the vectors X_i, which describe the motion of the dots purely *relative* to each other, and (2) the vector Y which describes the motion *common* to the dots (Johansson [12]).

However, let us now give an analysis of the percept as a nested system of control —for we will find subsequently that important organizational features emerge.

Consider first the *internal* structure of the rod. Recall that an internal group generates a stimulus set from a subset of itself. Since we have here only two points, the pair must be generated from one of its members. Note that one dot can be reflected onto the other about the line bisecting the "rod." Thus the pair can be generated by applying the group **Reflection** to one of the dots. Again, notice that the pair is reflectionally symmetric also about the line that runs along the rod. Thus the internal structure of the pair contains the group **Reflection** × **Reflection**.

Consider next the *external* structure of the rod. First the rod is acted on by the group **Stretches**, and, second it is moved through space under the group **Translations**. Thus, fitting together the internal and external structure, we obtain

$$(\textbf{Translations} \cdot \textbf{Stretches}) \,\hat{\imath}\, ((\textbf{Reflection} \times \textbf{Reflection}) \cdot \textbf{Point}).$$

An important thing to observe about this structure is that it obeys the interaction law; i.e., symmetry axes of the internal group are converted into eigenspaces of the external group, as follows: Recall that the pair of dots is reflectionally symmetric about the rod's bisecting line. However, this axis becomes the eigenspace known as the common-motion vector. Again, recall that the dot-pair is reflectionally symmetric about the line that runs along the rod itself. However, this axis becomes the eigenspace that is known as the relative motion vector. Another important feature of this structure will emerge later.

15. GENERALIZED CYLINDERS

Generalized cylinders, defined by Binford [2], are a basic tool in computer vision, for the analysis of complex shape. Essentially, on natural shape, such a cylinder

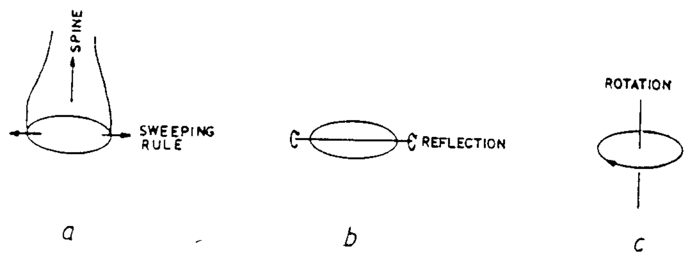


FIG. 15. The structure of a generalized cylinder.

corresponds to a limb (e.g., Fig. 1a). Its structure is shown in Fig. 15a: A cross section is deformed (widened and narrowed) as it is moved through space along a curve. The deformation is called the *sweeping rule*, and the curve is called the *spine* of the cylinder.

Now let us analyze this structure as a nested system of control, since important organizational features can emerge subsequent to such analysis. We start with the internal structure of the cross section, and then move onto the external structure.

One can generate the cross section by taking a single point and applying the rotation group to that point, as shown in Fig. 15c. However, observe that the cross section is reflectionally symmetric about itself, as shown in Fig. 15b. That is, reflection makes copies of the rotational structure. Thus, group-theoretically, one says that the reflection and rotation structure are combined by a semi-direct product, $(\times_{sd})$. In conclusion, then, the internal structure of the cross section is given by **(Reflection**$(\times_{sd})$**Rotation)** $\cdot$ **Point**.

Now let us turn to the external structure of the cross section. The cross section undergoes the action of two groups: **Stretches**, which defines the sweeping rule, and **Translations**, which defines the spine.

Thus, putting together the internal and external structures, we obtain the following system of groups:

$$(\textbf{Translations} \cdot \textbf{Stretches}) \wr ((\textbf{Reflection}(\times_{sd})\textbf{Rotations}) \cdot \textbf{Point}).$$

The first important thing to observe about this structure is that it obeys the interaction law; i.e., symmetry axes are converted into eigenspaces, as follows: One symmetry axis is the rotation axis in Fig. 15c, and this becomes the eigenspace that is usually called the spine (Fig. 15a). The other symmetry axis is the reflection axis in the plane of the cross section (Fig. 15b). This axis becomes the eigenspace that describes the direction of the sweeping rule (Fig. 15a).

Something remarkable now emerges. Let us consider both the sequence that was proposed as the structure of the Johansson motion situation, and the sequence just given for Binford's generalized cylinder:

Johansson situation:

$$(\textbf{Translations} \cdot \textbf{Stretches}) \wr ((\textbf{Reflection} \times \textbf{Reflection}) \cdot \textbf{Point})$$

Binford situation:

$$(\textbf{Translations} \cdot \textbf{Stretches}) \wr ((\textbf{Reflection}(\times_{sd})\textbf{Rotations}) \cdot \textbf{Point}).$$

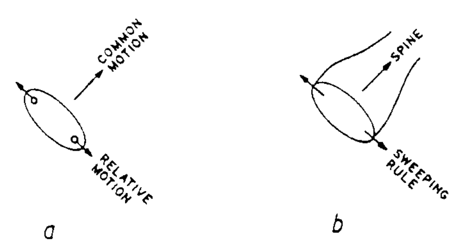


FIG. 16. Matching the eigenspaces of the Johansson motion phenomenon and the eigenspaces of the generalized cylinder.

What is remarkable is that closer examination reveals that the two sequences are identical—except that the latter is a 3-dimensional version of the former. Both begin at the far left with a point. Then, progressing to the next factor, the Johansson case has a reflection group, and the Binford case has a rotation group. However, if a generalized cylinder were used to describe a 2-dimensional situation, the rotational symmetry would be replaced by a reflectional one, thus giving the same component, at this point, as the motion situation. Then, progressing on to the next factor to the left, both sequences have a further reflectional component. Finally, the external groups, in both cases, consist of a stretch factor and a translation factor.

Our analysis therefore shows that the two situations, which are currently understood as quite different, are in fact structurally the same. Furthermore, using our constructs, it is instructive to see how they correspond, as follows: Let us line up eigenspaces. In the Binford cylinder, the eigenspaces are the sweeping rule and the spine. By matching these to the eigenspaces of the Johansson situation, we find that the cylinder's *sweeping rule* corresponds to the *relative motion vector* and the cylinder's *spine* corresponds to the *common motion vector*, as shown in Fig. 16.

Finally, it is worth pointing out the following phenomenon. Recall that it was claimed in Section 12.4 that the imposition of a Cartesian frame is the presence of the sequence (**Stretches** · **Shears** · **Rotations**)$\textcircled{w}$((**Reflection** × **Reflection**) $(\times_{sd})$(**Translations** × **Translations**)) along the full sequence $G_1 \cdot G_2 \ldots G_n$ of the perceptual organization. Considering now the motion and the cylinder situations, we see that certain factors of the Cartesian frame do indeed occur along the sequences developed for those percepts. In fact, we shall find in the next section that, as predicted, *all* the Cartesian group factors occur along the latter—when the latter is fully specified. An understanding of this will be crucial to the analysis of complex shape.

However, at this stage, consider the following: In the analysis by Marr and Nishihara [23], the Cartesian frames that are assigned to various parts of the shape-image are *2-dimensional*. By the above analysis, we see that this means that the shape outline of each part is structured in exactly the same way as the Johansson motion phenomenon. In other words, what results is a generalized "ribbon" rather than a generalized cylinder. Now, to form the cylinder, i.e., a limb, one rotates the ribbon about its axis. However, in this rotation, the axis remains over itself; i.e., it becomes an *eigenspace*. In other words, the symmetry axis of the Cartesian frame becomes the eigenspace in the rotational movement that creates the limb. Thus we see that the *interaction law* is crucially involved in converting a Cartesian frame into a limb.

16. COMPLEX SHAPE

The remainder of this paper is concerned with the analysis of complex shape. It might first appear that there is a barrier to extending the preceding analysis to complex shape, because the operations and principles that were proposed earlier were applied *globally*. For example, in the prototypification of a parallelogram to a rectangle, the single linear transformation, shear, was understood as acting on the *entire* figure. In contrast, consider the prototypification of the seal shown in the second column, third row of Fig. 17. The action of a global linear transformation, like shear, cannot straighten (i.e., prototypify) the back. Furthermore, the prototypification of the back involves a different type of transformation from that of the head. Thus, the prototypification of complex shape requires operations that are *local* in two senses:

LOCAL, SENSE 1. *Properties are confined to particular regions.*
LOCAL, SENSE 2. *Properties vary smoothly over a region.*

However, what we shall find, surprisingly, is that the same principles that have been developed in the previous sections *do* apply to complex shape; but they are used locally in the two senses just defined. A crucial example is evident in the following experiment: In Leyton [18], subjects were presented with the 22 outlines of complex natural and abstract shapes shown in Fig. 17. Their task was to give, at each of 4 points in each shape, the direction of *maximal perceived flexibility* of the region in which the point was situated. With considerable statistical significance, the subjects chose local symmetry axes. This means that they were converting local symmetry axes into local eigenspaces.

The conclusion of this experiment is that the interaction law applies to complex shape and that is applies locally. Thus we have a crucial clue in the analysis of local prototypification, because (1) it can be assumed that prototypification occurs along lines of maximal perceived flexibility and (2) we find that, in accord with the interaction law, such lines are the local symmetry axes.

However, we have a problem. Because, in complex shape, the interaction law converts local symmetry axes into local eigenspaces, we require, as input, a local symmetry analysis. There are three natural local symmetry analyses in use at present; they are the *symmetric axis transform* (SAT) of Blum [3], the *smoothed local symmetry* (SLS) of Brady [3], and the *process-inferring symmetry analysis* (PISA) of Leyton [21]. These can be regarded as *natural* local symmetry analyses because they are the result of the reflectional symmetries yielded by the tangent vectors of the contour. To illustrate, consider Fig. 18. The bold curve in this figure is a segment of contour. The vectors $\mathbf{t}_A$ and $\mathbf{t}_B$ are tangential to the curve at A and B, respectively. That is, they *locally* represent the curve at A and B. The points A and B are paired because the two vectors are reflectionally symmetric about a central vector $\mathbf{t}$ (shown in the figure). That is, one can place a local mirror along $\mathbf{t}$, and it will reflect the vector $\mathbf{t}_A$ onto the vector $\mathbf{t}_B$. Thus, $\mathbf{t}$ describes the *local* reflectional structure between A and B. Now each of the three schemes, the SAT, SLS, and PISA, fixes a different point on the local mirror to be the "incidence point" for the imaginary ray

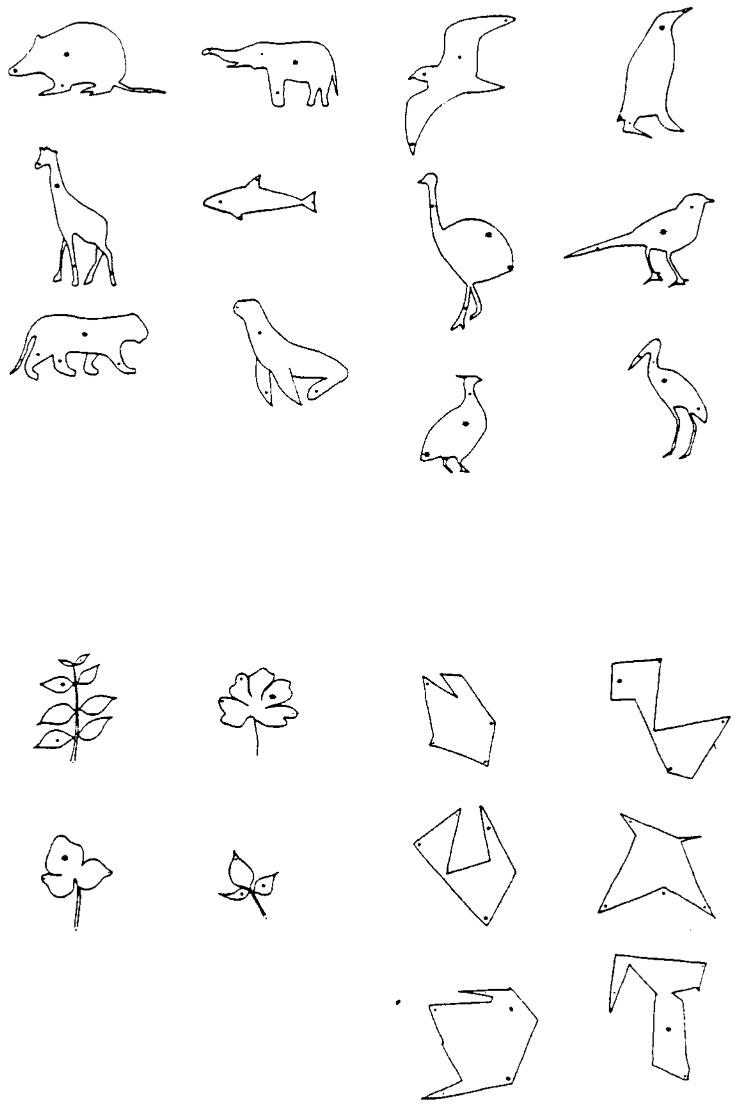


FIG. 17. Given these 22 complex shapes, subjects converted local symmetry axes into local eigenspaces.

of light that defines the reflection. To simplify the discussion, we shall choose the Brady point, which is the midpoint along the line AB. Finally, the *symmetry axis* of the entire curve is the locus of such midpoints (i.e., the dotted line shown), which results from taking all equivalent constructions along the contour.

Recall now our main problem. We wish to understand the *local* prototypification of shape. In preparation for this, we have, in the previous paragraph, begun to look at local structure that accords with *Local, Sense 2*, above. However, we also need to understand local structure that accords with *Local, Sense 1*. In other words, we need to understand what a *local region* is.

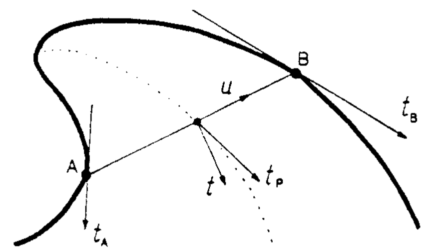


FIG. 18. Local symmetry analysis: A and B are paired when their tangent vectors, t_A and t_B, are reflectionally symmetric, i.e., about a local mirror t.

To settle this issue, we shall invoke an analysis provided by Hoffman and Richards [10] and Richards and Hoffman [27]. The central claim of their analysis is that contours are perceptually segmented at points of negative curvature extrema ("maximal indentation"). An example of this is shown in Fig. 1f, where such a segmentation of the face yields the crown of the head, the nose, the lips, and the chin.

Hoffman and Richards call a segment where endpoints are curvature minima, a *codon*. Hence, they introduce the concept of a *codon representation* of a curve; i.e., the curve's description as a sequence of codons. Such a representation has two important advantages: (1) Any curve has a *unique* codon representation; and (2) there are only five types of non-trivial codons. This means that any curve can be given uniquely as a string of letters taken from an alphabet of size five. The five codon types are shown in Fig. 19. They are distinguished by the number of curvature zeros (flat points) on the codon, and whether the endpoints have positive or negative curvature. In Fig. 19, the flat points are represented by dots, and the endpoints by slashes.

Is there any reason why codons should be made our required definition of *local region*? In other words: Are codons related to the interaction law (symmetry axes → lines of flexibility) and hence to the problem of prototypification?

The answer, I claim, is given by the following theorem that I proposed and proved in Leyton [19]. It relates symmetry structure to curvature extrema.

SYMMETRY-CURVATURE DUALITY THEOREM (LEYTON [19]). *Any segment of smooth planar curve, bounded by two consecutive curvature extrema of the same type (either both maxima or both minima), has a unique symmetry axis (SAT, SLS, or PISA), and this axis terminates at the curvature extremum of the opposite type.*

COROLLARY. *The SLS of a codon is unique, and terminates at the point of maximal curvature on the codon.*

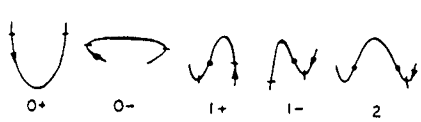


FIG. 19. The five non-trivial codons.

It is important to observe that this theorem relates two branches of perceptual research that were previously believed to be quite separate: (1) *symmetry* research, starting with the Gestalt movement and going up to modern AI symmetry-extraction programs, and (2) *curvature* research, starting with Attneave's [1] demonstration that information is maximized at curvature extrema (Fig. 1e) and going up to a recent mathematicization of Attneave's results by Resnikoff [26].

The above theorem justifies defining a local region to be a codon, because the theorem implies that if one travels along a contour, past the endpoint of a codon, a new symmetry axis appears. Thus a codon not only has a unique symmetry axis, it is the *maximal* region with respect to which this uniqueness holds. This allows us to optimally combine *Local, Sense 1* and *Local, Sense 2*, in that, according to the theorem, a codon is that region which is maximal with respect to *Sense 1*, using symmetry in *Sense 2*.

We can now return to the prototypification of complex shape. Consider Fig. 18, again. The segment of contour shown is, in fact, a codon. It is for this reason that it was drawn with a unique symmetry axis (the dotted line), and such that the axis terminates at the point of maximal curvature. Recall now that the vectors **t** and **u** are symmetry vectors: **t** is the axis about which $\mathbf{t}_A$ and $\mathbf{t}_B$ are symmetric; and **u** is a symmetry axis within the 3-dimensional cross section, as shown in Fig. 15. Now by the interaction law, the two axes are converted into eigenspaces, i.e., directions of flexibility. That is, they provide a local coordinate system for the action of linear transformations (i.e., a basis of eigenvectors). In fact, as this coordinate system moves along the curved axis, it undergoes (1) stretch, because the cross section widens and narrows; (2) shear, because **t** maps to $\mathbf{t}_P$, the tangent to the curved axis; and (3) rotation, because the axis curves.

Therefore, the structure of the codon conforms exactly to that given earlier as the full group structure of the Cartesian frame:

$$(\textbf{Stretches} \cdot \textbf{Shears} \cdot \textbf{Rotations}) \wr ((\textbf{Reflection} \times \textbf{Reflection})(\times_{sd}) \\ (\textbf{Translations} \times \textbf{Translations})).$$

It is this structure that finally provides us with a method of prototypification: for one can remove the factors, rotation, shear, and stretch, as was done in the global case.

Let us therefore find out what shapes result from this process. Figure 20 shows the structure of prototypification. The top node represents an arbitrary shape. The middle nodes represent shapes with one level of prototypification, i.e., where either stretch, shear, or rotation have been removed. The bottom nodes represent a second level of prototypification, where a further factor has been removed.

Let us consider first the case where the transformations represent *global* actions (as in Fig. 3). It is important to observe that, in this case, all the nodes represent mathematically realizable shapes. For example, if the top node is a rotated parallelogram, the middle nodes represent, from left to right, respectively: a rotated rhombus, a rotated rectangle, and a parallelogram. Furthermore, the bottom nodes represent, from left to right, respectively: a rhombus, a rotated square, and a rectangle. In fact, the experimental results (summarized in Fig. 3) established that subjects take the right-hand sequence of nodes, even though shapes exist at all other positions.

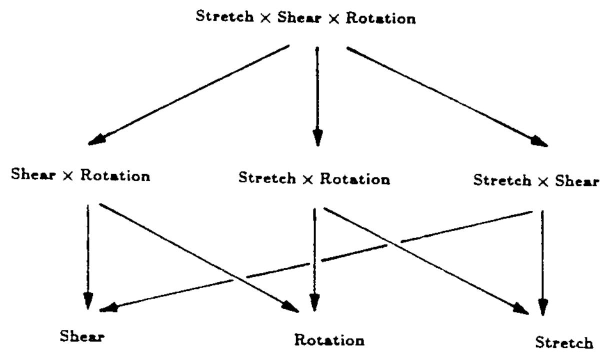


FIG. 20. All possible factorizations of **Lin**.

The question we must now consider is what shapes exist when the transformations are used *locally*, for example, as in Fig. 18. To answer this question, I recently proved

THEOREM (Leyton [20]). *Let a shape be locally characterized by the decompositon* **Lin = Stretches · Shears · Rotations**, *defined by the coordinate system yielded by the SLS symmetry vectors under the interaction law. Then the removal of one of the factor subgroups necessarily involves the removal of one of the other factor subgroups.*

What the theorem states is that one level of prototypification is impossible. That is, as shown in Fig. 21, no shapes can exist at any of the middle-level nodes. Furthermore, in Leyton [20], I show that no shape can exist at the bottom left node.

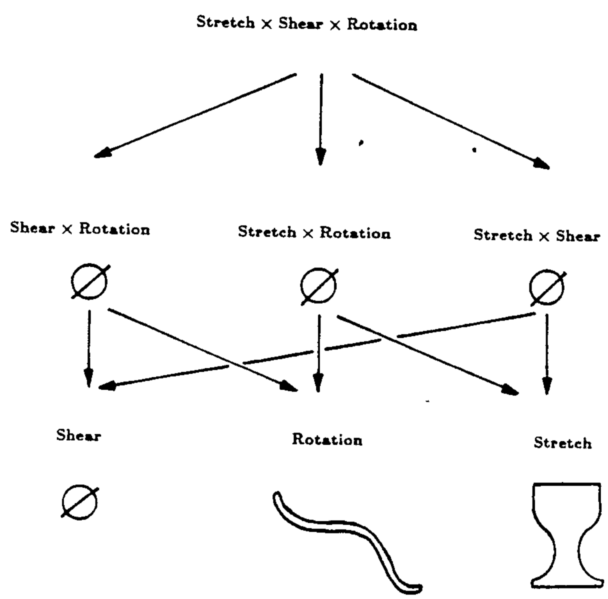


FIG. 21. The only possible shapes that result from factorizations of **Lin**, when **Lin** is used locally to characterize the SLS.

This leaves only two possible nodes: the rotation-only and stretch-only nodes on the bottom level. The former node corresponds to shapes that are flexed symmetries (e.g., worms; see Fig. 21); and the latter corresponds to shapes that are global symmetries (e.g., goblets; see Fig. 21).

It is important to observe that these two types of shapes are, psychologically, *highly salient as prototypes*. This need not have been the case. For example, we could have gone through this long process; i.e., applied the Brady SLS, applied the interaction law locally such that the SLS is locally represented by local linear transformations, performed the factorization—and nevertheless found that the resulting shapes were not significantly prototypical. There are many points where mistakes could have been made. For example, in Leyton [20], I show that if the same scheme is followed, except that a coordinate system is used that is different from that determined by the interaction law, the resultant shapes are not psychologically prototypical.

We shall now bring into consideration a further group. Besides the Cartesian factors, **Stretches**, **Shears**, and **Rotations**, the organization of a codon also involves **Translations**, which actually moves the frame down the axis of the codon. This was seen in our analysis of generalized cylinders, where a translation factor was gained, corresponding to the movement of the cross-section down the spine (recall Fig. 15a, and the group sequence that was generated for the cylinder). Thus, the external structure of the frame defining the codon is

Translations · **Stretches** · **Shears** · **Rotations**.

Now observe that while the removal of the factors **Stretches**, **Shears**, and **Rotations**, symmetrizes the codon, the removal of **Translations** causes the codon to shrink until it disappears. This is because the removal of **Translations** makes the axis diminish until it has no length, and, by the symmetry-curvature duality theorem, this must simultaneously result in the disappearance of the extremum at the tip of the codon. However, the codon is a protrusion, or indentation, on a still larger shape. Thus the removal of the **Translations** factor causes the overall shape to lose a protrusion or indentation, i.e., to become smoother.

We conclude now by showing that the analysis in this section provides a particular scheme, for the prototypification of complex shape. This scheme can be demonstrated using Fig. 22a, which is the contour of a dolphin under water.

Local Prototypification

Stage 1. Apply a *local* symmetry analysis. Thus, for example, the head of the dolphin (Fig. 22a) is given a curved symmetry axis.

Stage 2. use the interaction law *locally* on codons (protrusions or indentations), converting *local* symmetry axes into *local* eigenspaces in a *local* use of **Lin** = **Stretches** · **Shears** · **Rotations**.

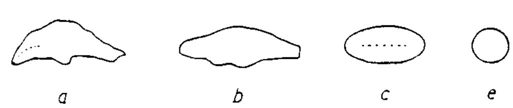


Fig. 22. A dolphin under water.

Stage 3. *Locally* prototypify by removing the factors, **Stretches, Shears, Rotations.**

TRANSITION: LOCAL → GLOBAL. Remove **Translations.** This is equivalent to removing the protrusion or indentation by shrinking its axis. According to the symmetry-curvature duality theorem, this removes the curvature extremum, thus causing a reduction in curvature variation.

GLOBAL PROTOTYPIFICATION. The reduction in curvature variation that results from the removal of all protrusions, in the above way, leads to a *globally* symmetric structure; e.g., the oval in Fig. 22c. This allows the following stages:

Stage 1. Apply a *global* symmetry analysis. For example, this extracts the axis of the oval shown in Fig. 22c.

Stage 2. Use the interaction law *globally*, converting *global* symmetry axes into *global* eigenspaces in a *global* use of **Lin** $\dot{=}$ **Stretches · Shears · Rotations.**

Stage 3. *Globally* prototypify by removing the factors, stretch, shear, and rotation. Hence one obtains, for smooth shape, the circle shown in Fig. 22d. The circle is *internally* generated by a subgroup of the Euclidean group (i.e., by rotations). This corroborates Proposal 2 (Sect. 9) that prototypes are generated in this way.

Essentially, the scheme is that of the local application of (1) a symmetry analysis, (2) the interaction law, and (3) stretch—shear-rotation—removal; followed by the global application of these three stages—where translation-removal intervenes between the local and global application of the three stages. In fact one can best specify this system by saying that the three local stages and translations-removal are cycled through recursively from the smallest to the largest scales.

17. SUMMARY

The basic principle of this system is the proposal that cognition is the attempt to represent the environment as a collection of machines, i.e., as deterministic processes. In order to understand the representational consequences of this proposal, it is necessary to understand that any machine has two kinds of structure: an *operational* (computational) structure and a *stability* structure. We began by examining the first kind of structure. In fact, we examined *two* ways in which a stimulus set can be organized in terms of the operational structure of a machine. The two methods were used to characterize two very different sets of *psychological assumptions*, which were illustrated using a hexagon. In one method, the stimulus set is structured symmetrically, and each stimulus (e.g., a side of the hexagon) is a perceptual invariant. In the other method, the stimulus set is given an asymmetric referential structure, and the notion of invariant ceases to play an important role. Because of the weight of evidence from linguistics and categorization, and the psychological data reviewed in this paper, we argued that the second method of description is the appropriate one on which to base a theory of perceptual organization.

It was then argued that the group G, which is used to structure the stimulus set, is given a stratification $G = G_1 \cdot G_2 \ldots G_n$ into levels where each level G_i is itself a group. Reference, in the group, proceeds not in one step to the overall reference point (identity element), but *successively*, first through level G_n to its own reference point e_n, then through G_{n-1} to its own reference point e_{n-1}, and so on down to G_1 and to its own reference point e_1. In order to answer the question of why this decomposition and ordering occurs, we proposed that the decomposition is a stratification of G into *levels of stability*, and the order of removal is that which starts from the most unstable and progresses to the most stable.

We ended our discussion of the general cognitive principles by showing that a structure with the above type of stability stratification has an important property: it forms a *nested system of control*. By this, we mean that, given a single operation in one level, the operation acts on the *entire* level below it; i.e., it transforms that level *as a whole*.

The basic proposal of our perceptual theory is that a perceptual organization is the *interaction* between two generative descriptions: (1) One is given by an *external group*, which generates the stimulus set from a prototype. (2) The other is given by an *internal group*, which (a) generates the prototype from a subset of itself, and (b) is a subgroup of the Euclidean group.

We then examined the structure of the internal group and argued that the group is a nested structure of control. Note that it follows that a prototype is a nested control structure where each level of control is a subgroup of the Euclidean group. We also argued that a psychological grouping corresponds to a right subsequence in $G = G_1 \cdot G_2 \ldots G_n$ (i.e., the sequence of groups up to a particular level of control). An evaluation principle was proposed that chooses between the alternative control hierarchies assignable to a stimulus set. The principle states that higher levels of control are assigned to more distant stimuli, and lower levels to nearer stimuli.

A central proposal to the organizational theory is the interaction law which relates the internal structure of a perceptual organization to the external structure, by proposing that the symmetry axes of the former become eigenspaces ("lines of flexibility") of the most stable external groups.

We then moved on to an analysis of Cartesian frames and expressed such frames as a sequence of groups. It was claimed that the visual imposition of a Cartesian frame is actually the "splicing" of the Cartesian group-sequence into the full group-sequence of the perceptual organization.

The above principles were strongly corroborated in several areas of perceptual organization. For example, we then expressed what has classically been called "the orientation and form problem" in this way: Different assignments of the external group determine different assignments of the internal group. This phenomenon becomes clearly explicable by the interaction law (symmetry axes $\rightarrow$ eigenspaces).

Again, in building up an analysis of the Johansson [12] motion phenomenon, as a nested structure of control, we found that the symmetry axes of the "rod" become the eigenspaces currently understood as the common motion vector and the relative motion vector. Similarly, in building up an analysis of generalized cylinders as nested structures of control, we found that the symmetry axes of the cross section become the eigenspaces of what are more commonly known as the sweeping rule and the spine. Furthermore, a remarkable thing emerged from these analyses: we found that the Johansson phenomenon and the generalized cylinder have the same structure, and that, matching up eigenspaces, the common-motion vector corresponds to the spine and the relative motion vector corresponds to the sweeping rule.

Finally, we analyzed complex shape. We saw that such shape requires analysis that is local in two senses: (1) the properties are confined to particular regions, and (2) the properties vary smoothly over a region. Psychological results, using 22 complex shapes, were cited showing that the interaction law holds locally (in both senses) for complex shape; i.e., local symmetry axes are converted into local eigenspaces. We claimed that the optimal local region of a shape is a codon because the symmetry-curvature duality theorem shows that a codon has a unique symmetry axis. The symmetry analysis of a codon was then converted into a moving frame using the interaction law, where the external action of the frame is defined by the sequence **Translations** · **Stretches** · **Shears** · **Rotations**. A theorem was used to show that the removal of one of these factors (in the right-to-left direction in the sequence) necessarily involves the removal of one of the other factors, and that this symmetrizes the codon (i.e., protrusion or indentation). Finally, the removal of **Translations** removes the protrusion completely, by shrinking along the axis. Furthermore, by the symmetry-curvature duality theorem, the removal of translations causes the removal of the associated curvature extremum. Thus it was proposed that the prototypification of complex shape involves the successive removal of these four factors in a cyclic recursion from the smallest to the largest scales.

18. GUIDE TO RELATED PAPERS

A number of people have asked me to include a guide to my other papers concerning nested control.

Leyton [17] and [18] give a much more detailed treatment of both the cognitive and perceptual material. Lengthy discussion is included on the structure of cognitive reference frames and the relationship between the perceptual phenomena and reference frames. These papers can be understood by anyone who has read the present paper.

Leyton [16] applies the cognitive principles to other levels of the human cognitive system. The paper is a rather lengthy one, 120 journal pages, and a substantial proportion of it is devoted to linguistics, surveying various grammatical theories and drawing data not only from English but from other languages such as African Bantu languages and Philippine languages. Familiarity with Leyton [17] is advisable before tackling this paper.

Leyton [14] gives a concentrated and technical elaboration of the perceptual theory. Leyton [15] gives an exposition of the perceptual material, not from representational principles, but by describing a single neuronal architecture that will result in each of the perceptual phenomena.

Leyton [19] proves the symmetry-curvature duality theorem, and a number of other results relating the Brady (SLS) to curvature extrema. Leyton [20] gives a proof of the last theorem quoted in Section 16. Finally, Leyton [21] develops a grammar for the process-analysis of shape, where the rules of the grammar are expressed purely in terms of the curvature extrema.

ACKNOWLEDGMENTS

The writing of this paper was supported by NSF Grants IST-8418164 and IST-8312240, and by AFOSR Grant F49620-83-C-0135. The paper was written while the author was at Harvard University.

REFERENCES

1. F. Attneave, Some informational aspects of visual perception, *Psychol. Rev.* **61**, 1954, 183–193.
2. O. B. Binford, Visual perception by computer, presented at *IEEE Syst. Sci. Cybern. Conf.*, 1971, Miami, Fla.
3. H. Blum, Biological shape and visual science (Part 1), *J. Theor. Biol.* **38**, 1973, 205–287.
4. L. S. Bobrow and M. A. Arbib, *Discrete Mathematics: Applied Algebra for Computer and Information Science*, Hemisphere, Washington, D.C., 1974.
5. M. Brady, Criteria for representations of shape, in *Human and Machine Vision* (A. Rosenfeld and J. Beck, Eds.), Erlbaum, Hillsdale, N.J., 1983.
6. N. Chomsky, *Syntactic Structures*, Mouton, The Hague, 1957.
7. J. J. Gibson, *The Senses Considered as Perceptual Systems*, Houghton-Mifflin, Boston, Mass., 1966.
8. J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton-Mifflin, Boston, Mass. 1979.
9. E. Goldmeier, Uber Anlichkeit bei gehenen Figuren, *Psychol. Forsch.* **21**, 1936, 146–209; English transl. by E. Goldmeier, Similarity in visually perceived forms, *Psychol. Issues* **8**, No. 1, 1972; Monograph 29, International Universities, New York.
10. D. D. Hoffman and W. A. Richards, Parts of recognition, *Cognition* **18**, 1984, 65–96.
11. W. C. Hoffman, The Lie algebra of visual perception, *J. Math. Psychol.* **3**, 1966, 65–98.
12. G. Johansson, *Configurations in Event Perception*, Almqvist & Wiksell, Stockholm, 1950.
13. S. Lang, $SL_2 R$, Addison–Wesley, London, 1975.
14. M. Leyton, Perceptual organization as nested control, *Biol. Cybern.* **51**, 1984, 141–153.
15. M. Leyton, Generative systems of analyzers, *Comput. Vision Graphics Image Process.* **31**, 1985, 201–241.
16. M. Leyton, Principles of information structure common to six levels of the human cognitive system, *Inform. Sci.* **38**, No. 1, 1986.
17. M. Leyton, A theory of information structure. I. General principles, *J. Math. Psychol.* **30**, 1986, 103–160.
18. M. Leyton, A theory of information structure. II. A theory of perceptual organization, *J. Math. Psychol.* **30**, 1986, 275–305.
19. M. Leyton, Symmetry-curvature duality, *Comput. Vision Graphics Image Process.*, in press.
20. M. Leyton, Prototypification of complex shape by local Lie group action, submitted for publication.
21. M. Leyton, A process grammar for shape, submitted for publication.
22. E. Mach, *The Analysis of Sensations*, 1897; English transl., Dover, New York, 1959.
23. D. Marr and H. K. Nishihara, Representation and recognition of the spatial organization of three-dimensional shapes, *Proceedings, R. Soc. London B* 200, 1978, 169–294.
24. J. Piaget, *The Mechanisms of Perception* (transl. by G. N. Seagrim), Basic Books, New York, 1969.
25. S. M. Pizer, W. Oliver, and H. Bloomberg, Hierarchical shape description via the multiresolution of symmetric axis transform, tech. report, Univ. Northern Carolina, 1986; submitted for publication.
26. H. L. Resnikoff, *The Illusion of Reality: Topics in Information Science*, Springer-Verlag, New York, 1987.
27. W. Richards and D. D. Hoffman, Codon constraints on closed 2D shapes, *Comput. Vision Graphics Image Process.* **31**, 1985, 265–281.
28. I. Rock, *Orientation and Form*, Academic Press, New York, 1973.
29. E. Rosch, Cognitive reference points, *Cognit. Psychol.* **7**, 1975, 532–547.
30. M. Wertheimer, Laws of organization in perceptual forms. (English translation (1938) in *A Source Book of Gestalt Psychology* W. D. Ellis, Ed.), Harcourt Brace, New York, 1923.