Title

A Unified Theory Of Cognitive Reference Frames

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 4(0)

Author

Leyton, Michael

Publication Date

1982

A Unified Theory of Cognitive Reference Frames

Michael Leyton

Department of Psychology
University of California, Berkeley

The term reference frame is used in a wide variety of studies to describe a remarkably diverse set of phenomena in the field of Human Cognition. No unified theory exists. This paper elaborates such a theory and applies it to a number of examples in the following main areas: (1) Categorization and Prototypicality (2)Visual Shape Perception; (3) Auditory Reference; (4)Motion Perception; (5)Linguistic Deixis.

Cognition as the modeling of logical processes

In a lengthy study of perceptual organization (Leyton, 1974), I concluded that perception is an attempt to represent the world as a set of logical languages. Any such language consists of four components (S,F,A,P)

S= a set of primitive symbols
F= a set of rules of formation
A= a set of axioms
P= a set of rules of procedure

Essentially, the rules P are applied to the axioms A to produce a further set of formations which are called theorems. I argued that perception attempts to distinguish in any environment an axiom set of stimulus formations and derive the other stimuli as theorems generated from A, via perceptual operations P. I argued further that because a logical language is equivalent to a machine (Minsky, 1972), perception is inherently an attempt to give a machine-like (or computational) account of the environment. Because it seemed to me that perception, as a descriptive mechanism, exhibited, in the above respects, general properties of all descriptive processes, I argued that all information or description is inherently a computational account.

Although my argument in Leyton (1974) used purely cognitive evidence, in Leyton (1981a&b),I arrived at the same conclusion using theoretical-biological and statistical-mechanical arguments: Perceptual mechanisms were developed to identify, in the environment, machines to which the organism could couple itself to extract work. Thus, in claiming that all perception is the description of machines (or computational processes), I was claiming that all perception is inherently the identification of available work. The present paper elaborates this view further and shows how it explains cognitive reference phenomena.

Machines as the basis of description.

Essentially, any machine M (a state-output machine) can be described as

$$\mathbf{M} = \begin{cases} Q = \text{a set of states (i.e. a state-space)} \\ P = \text{a set of inputs} \\ \text{an action of the input set on the states} \end{cases}$$

The inputs can thus be viewed as transformations causing state-transitions.

I claim that all description (including perception) is an attempt to characterize classes of stimuli as state-spaces of machines. Thus, in particular, I argue that the properties of any single stimulus are split into two classes (1) those properties denoting state, (2) those denoting the object which is undergoing the state (e.g. a falling rock). Thus we have:

Preliminary definition A description of a stimulus set S, is a map from the state-space of a machine onto S; that is, a map

$$D: Q \longrightarrow (S, \emptyset)$$

for some machine M=(Q,P).

(The empty set ∅ is included because S might not yeild the entire state-space).

Example Consider a hexagon. There are 12 transformations (rotations and reflections) which map it to itself: e, r_{60}, r_{120}, r_{180}, r_{240}, r_{300}, t, tr_{60}, tr_{120}, tr_{180}, tr_{240}, tr_{300}.

where e = no transformation (the identity map)
r_n = rotation by n degrees
t = reflection

In the above view of description,(1) the sides are perceived as the states of a machine, and (2) the state-transitions therefore become the above 12 transformations. All twelve transitions map any one side to some other side, or to itself. The resulting diagram is exactly the state-transition diagram of the associated finite-state machine. For clarity, Fig 1 presents only a part of the diagram. Most of the 72 perceived connections are omitted.

The meaning of reference.

I claim that a viable unified theory of reference frames is obtained if one assumes that the brain identifies certain states as initial ones; that is. they are viewed as pre-input or axiomatic. The important result is: Because all other states are then obtained by applying the input operations, P, each state is identifiable by the operation which produced it. Thus the machine description of a hexagon is reduced simply to viewing one edge, e.g. the top edge) as a starting point and viewing the others each as equivalent to only the operations which obtained them from the top. In consequence, the other sides are referred back

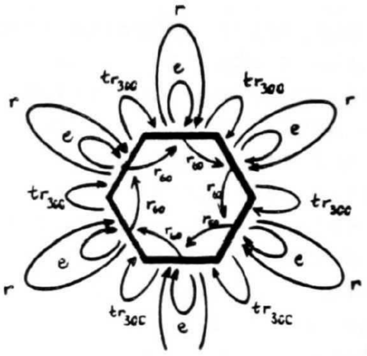


Fig 1. A state-space description of a hexagon

to the top one (Fig 2).(I proposed this view of reference, in mathematical-logical terms, in Leyton, 1974).

We therefore have a revised version of what a description is. It is a map from the inputs (or state-transitions) onto the stimulus set. Thus the individual stimuli are described as follows: 'this stimulus is what I obtained after I applied *such and such an act to the initial* one'.

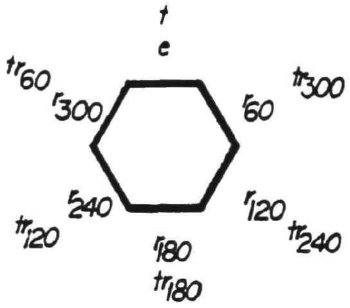


Fig 2. An input description of a hexagon

Group or input descriptions.

The system of state-transitions (or inputs) of a machine obeys a set of conditions which define it to be what mathematicians call a semigroup. We will assume the existence of an extra condition (each input has an inverse input) which makes the system what is called a group. The assumption is psychologically important because it allows the object/state splitting of the stimulus properties.

Thus the input set can be viewed as a group of state-transition functions, or an input group acting on S. But our theory of reference states that all description is the identification of stimuli with members of the input group. Thus we argue that all description is of this form:

Definition: A description of a stimulus set S is the map of the input group G, of a machine,onto S; in fact the map

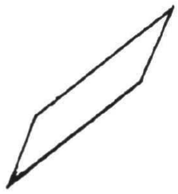
$$\Gamma : G \longrightarrow (S,0)$$

for some machine M=(Q,G)

Therefore, because reference acts with respect to the pre-input or non-transformed state, it acts here with respect to the non-transformation element (the 'identity' element) which every group contains.

The structure of reference.

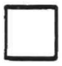
In the usual reference situation, the state-space is multidimensional; that is, it is the product of several one-dimensional component groups. In this case reference acts not just with respect to the identity element of the entire group but with respect to the identity elements of each of the 1-dimensional component groups. In fact, I found (Leyton, in preparation) that reference acts successively across the components. For example: a rotated parallelogram



nich is referred to a non-rotated one:

which is referred to a straightened one, i.e. a rectangle:

which is referred to the non-elongated version i.e. a square:



In the reference process above, the mind first eliminates the group of rotations,i.e. refers back to the identity of the rotation group, then it eliminates the group of shears, i.e. refers back to the identity of the shear group, and finally eliminates the group of elongations i.e. refers back to the identity of the elongation group. In fact, I have shown (Leyton, 1982: Leyton, in preparation) that the ordering in which elimination occurs is that of the perceived increasing stability of the successive group dimensions i.e. inputs. The above rotation is perceived as less stable than the shear, which is perceived as less stable than the elongation

We thus conclude:

Reference involves the mapping of an input group *of a machine* to a stimulus set such that the members of the set become viewed as a generated space of states, identifiable with th inputs that obtained them. The reference process successively factors out the 1-dimensional component groups (or machines) in order corresponding to their increasing perceived stability. The reference point in each dimension is the group identity element (i.e. giving the pre-input state).

APPLICATIONS

1. Protypicality and reference.

Rosch (1975) has proposed that natural categories - such as colors, line-orientations and numbers - have reference point stimuli - such as focal colors, vertical and horizontal lines,and number multiples of 10 - with respect to which other catergory members are judged. For example, pink is referenced to red, a leaning object to the vertical, and 99 to the number 100. The reverse references do not happen.

Using the above theory of descriptions, I claim that :

> A prototype is a stimulus which is labeled by the identity element of the associated input group.

It is for this reason, for example, that a giraffe is judged as an animal with a long neck, whereas the neck of a more prototypical animal, such as a dog, is not even mentioned. In our theory, the giraffe is viewed as needing a transformation to be obtained (in fact being equivalent to that transformation) whereas a dog is not, i.e. the dog is at the initial (axiomatic, pre-input) state of the associated dynamical system. Again 99 is obtained by moving 1 down from 100 (i.e. applying the subtraction transformation) whereas 100 is obtained by 'just staying there'.

2. Shape perception

2.1 Shape and orientation

As is now well documented, the perception of shape depends on the assigment of orientation (Rock, 1973) A famous example (Fig 3) is the perceived difference between a square and a diamond, which depends on how the perceiver places a reference coordinate system over the same underlying figure.

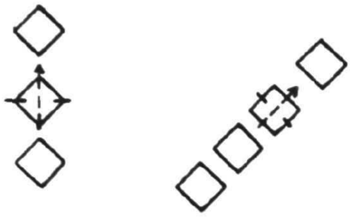


Fig 3. Assigned direction effecting perceived shape.

Leyton (1974, 1978) and Palmer (1981) have independently proposed a theory of shape perception, in terms of the internal symmetry transformations. However, while their view accounts for several important effects, it is clear that it does not account for the effects of orientation on form perception. I claim that the present view does, because it maps the input group directly down onto the stimuli, thus identifying the stimuli totally by the transformations (i.e. inputs) which obtain them. (Note that internal symmetries allow a range of alternative symmetrically related descriptions which do not violate interpretation.) Thus a definite element (or range of elements) has necessarily to be identified as the starting point of the associated machine. Furthermore, specific subsets have definitely to be identified with specific component 1-dimensional groups. A change of interpretation of a figure then becomes an alteration in the elements which perception allows to be labeled by the identity input, or an alteration in the subsets which receive the component groups, or a total change of group. For example,the main perceived axial structure of a square implies that it is interpretable as generated foom initial parts such as those in Fig 4. However, the main perceived axes of a diamond imply that amongst the allowable generators are the stimuli in Fig 5. Thus interpretations change with the set of allowable generators.

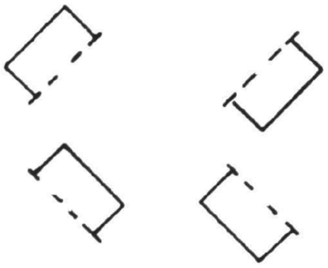


Fig 4 Allowable generators of a square.

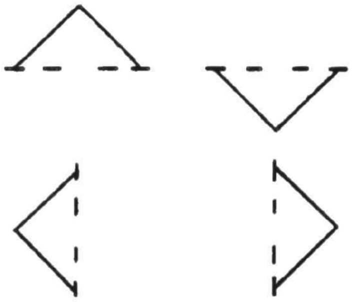


Fig 5 Allowable generators of a diamond.

2.2 What is shape?

A shape is an interaction between two state spaces: its internal state space (e.g. the input description of the hexagon, given in Fig 2) and its external state space; i.e. what the figure can do (e.g. rotate). We have seen that a square and a diamond are distinguished by the mappings of their internal input groups. However, I claim that they are distinguished also between their external input groups. For a square, the more stable input group includes the state transitions in Fig 6. However, for a diamond, it includes more stably the state transitions in Fig 7. Squashing across the corners is not allowed stably for the square. I have identified (Leyton, In preparation) that an important aspect of the interaction of the internal (symmetry) state space and the external one is that the axes of symmetry in the former become identified as the axes of flexibility of the latter (i.e. become the 1-dimensional component groups in the latter).

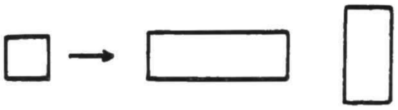


Fig 6 Allowable external inputs of the square

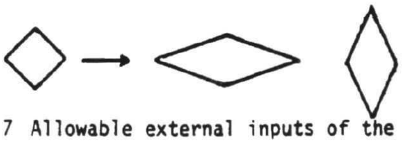


Fig 7 Allowable external inputs of the diamond

In their external descriptions, figures are also clearly identified as particular members of a state transition group, because reference also exists with respect to the initial point. For example, Wiser (1981) found that even if objects such as that in Fig 8 were presented in a non-vertical orientation, they were nevertheless recognized faster when presented again in the vertical orientation than in the initial one. Thus her results show that (1) the stimulus properties are clearly partitioned into those denoting state and those denoting the object

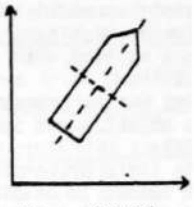


Fig 8 From Wiser (1981)

undergoing the state and (2) that the state is in fact identified as a transformation with respect to a referent initial state.

2.3 Pattern goodness

The relation of goodness to reference takes two important forms: Type 1, where a pattern such as Fig 9 is judged as less good, and is referenced to, its completed version; and Type 2, where a pattern such as Fig 10 is judged as less good, and is referenced to its non-deformed version, a square.

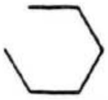


Fig 9

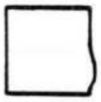


Fig 10

Our theory explains the two phenomena thus: Type 1: The goodness rating in Fig 9 is clearly based on the fact that the sides are perceived as needing more input transformations. That is, the entire machine has not been given. That is, a larger set of internal inputs is assumed for the figure. Thus, pattern goodness, in this case, is evaluated by the ratio

$$\frac{|\Gamma^{-1}(S)|}{|G|}$$

i.e. the proportion of the internal input group G used in the description map, Γ. Type 2: The goodness rating in Fig 10 is clearly based on the positioning of the figure in an external space of inputs (i.e. of deformations) and referencing it to an identity or pre-input element (which we have shown, constitutes the prototype).

We emphasize: Type 1 goodness verifies our postulation of an internal input group, and Type 2 goodness verifies our postulation of an external input group.

2.4 The Marr/Nishihara Shape Description Theory.

Marr and Nishihara (1978) claimed that the perceptual description of shape (e.g. the shape of animals) is given by viewing the figure as a concatenation of approximately cylindrical modules (Fig 11) with specific relative widths and lengths (Fig 12). These are obtained by assigning a collection of object-centered local reference frames (axes) to the parts of the stimulus configuration. The relationship between the frames is given by the coordinate system (p,r,θ,i,ϕ,s) where symbols are as shown in Fig 13. By applying our theory, we see that each of the figures in the Marr/Nishihara paper describes one of the points in an input space. The generation of a module (Fig 11) by translating a circle through space along an axis and by varying the diameter is the perception of external inputs to the circle. (Note that they become internal inputs of the module). The relative position of one module to another, as described by their coordinate system (Fig 13), is clearly a state-space, where the module positioning is essentially a state under the associated group of transformations along these parameters e.g. lowering arms lengthing legs, waving the hand, nodding the head, etc). Thus the figure is a point in the multidimensional group input space described by the interactions and shapes of the modules. The reference points in this space would be the prototypical animals and prototypical positions identified by the theory and techniques of Rosch (1978). Recall also that we claimed that an important interaction between the internal and external groups is that the invariant axes of the former become the component groups (directions of action) of the latter. This is clearly evidenced in the Marr/Nishihara description: the central axis of a module i.e. the invariant line under internal rotation of the module, becomes the direction along which it can be stretched.

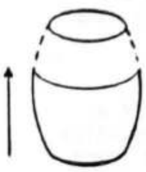


Fig 11 Generating a vase.

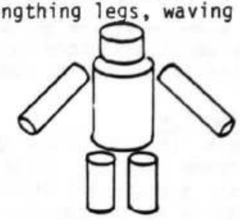


Fig 12 An ape.

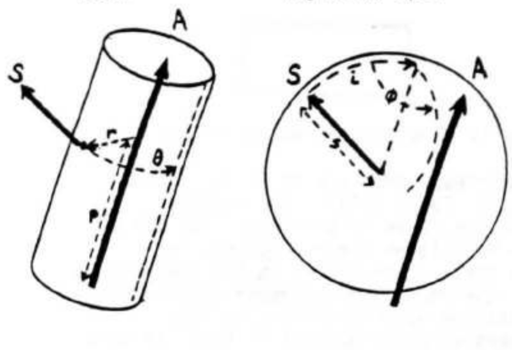


Fig 13 The Marr/Nishihara coordinates for relating two modules (After Marr & Nishihara, 1978).

3. Audition

3. Autitory Streams

Auditory input, e.g. a rapid sequence of tones is segregated perceptually into what Bregman and Campbell(1971) call, 'primary auditory streams'. These streams are groupings or frames and any tone can be allocated to only one of them.

Our theory of the situation is as follows: Bregman (1981) himself argued that an auditory stream corresponds to the object in visual perception. Leyton (1974) described the group-theoretic and logical language structure of music. In particular, he showed that musical transposition (change) of pitch is modeled by a group. This group allows the tones of a melody to be perceived as a single tone (object) being moved into different states under an input group. Therefore, the segregation of auditory stimuli into streams is, in our view, the description of the latter as a disjoint set of machines.

3.2 Musical Reference to the tonic.

If my hypothesis is correct that a stimulus becomes identified not just as a state of an object but as the operation (in an input group) which achieves that state, then there must be a stimulus which is labeled as the identity element of the group. This conjecture is amply evidenced by music: the reference point is called the tonic.

4. Relative motion

An example is the following: When a rectangular frame (Fig 14) is moved relative to an observer, and a point inside the frame is fixed relative to the observer, the point is nevertheless perceived to be the moving object (e.g. Rock, 1975). Our theory describes the above in this way: the set of possible velocities clearly defines the relevant state space (which is two-dimensional). However, the reference judgement enters when one identifies each velocity with the transformation which obtains it from the zero velocity, i.e. it is perceived as an increase (or decrease) of speed by a certain amount. This allows it to be referenced back to the '0' or identity element of the input group. The latter element is then assigned to perceptually the most stable object in the field, i.e. the rectangular frame.

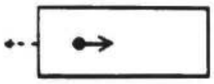


Fig 14 The induced motion effect

5. Linguistic deixis

Deixis (Bühler, 1934) is a term used to denote those linguistic aspects which locate or point to the object of speech; e.g. 'here', 'there', 'this', 'that', 'then'. Bühler claimed that these aspects create a coordinate system, centered on the referent (Fig 15).

Fig 15 The deictic field

The theory, which I have proposed, appears to model Buhler's concept. The deictic field clearly is a dynamical view of the space centered at the origin. "Put the book in front of my chair" means "One can find the place to put the book by inputing a translation forward from my chair's location". Thus, the coordinate system (Fig15) is - as I belive all coordinate systems are - labeled by the internal inputs (i.e. transformations) which move location with respect to the origin and axes. When an individual gives the pointing gesture, 'there', he is literally translating the deictic input group from himself to another point, such that the axes are properly aligned. As with the gravitational frame, these axes are representations of the 1-dimensional component groups of the internal input group; i.e. they give discrete labels for movement, not for physical packets of stimuli.

General conclusions.

The above presents a large-scale view of cognition. The view is corroborated by the several examples considered. In particular, the examples confirm the following principles

(1) Cognition is the attempt to model the environment as a union of machines.
(2) A reference frame as a machine with initial conditions defined.
(3) Referencing a stimulus is the process of
    (i) deciding on an object/state split of its properties
    and(ii) identifying the state properties with the input needed to obtain the stimulus from the initial conditions of the associated machine

The substantiation of this view of reference corroborates also our proposal that description is a mapping of an input group of a machine, onto a stimulus set.

References

Bregman, A.S. Asking the "What For" Question in Auditory Perception. In: Perceptual Organization, M.Kubovy &J.R. Pomerantz (Eds).Hillsdale, N.J.: Lawrence Erlbaum 1981

Bregman, A.S. & Campbell, J. Primry auditory stream segregation and perception of order in rapid sequence of tones. J. Exp. Psychol 1971, 89, 244-249.

Buhler, K. Sprachtheorie: die Darstellungsfunktion der Sprache. Jena: Gustav Fischer, 1934.

Leyton, M. Principles of Artistic Method: Algebraico-Logical Postulates in the Foundations of the Science of Perception. Research Report, Mathematics Institute, University of Warwick, Coventry, England 1974*

Leyton, M. Artistic Structure: Group-Theoretic, Differential-Geometric, and Mathematical-Logical Factors in Perception. Research Report, Mathematics Institute, University of Warwick, Coventry, England, 1978*

Leyton, M. Artistic Activity and Human Survival: Volume 1. Unpublished book. 1981a*

Leyton, M. Do Structural and Statistical Evaluations of Information Vary Inversely or Directly. Research report. 1981b*

Leyton, M. Description, categorization, and reference - a unified theory. Seminar given at the Department of Psychology University of California, Berkeley, March, 1982*

Marr, D. & Nishihara, H.K. Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London. 1978, B200, 169-294.

Minsky, M. Computation: Finite and Infinite Machines. Englewood Cliffs NJ: Prentice Hall, 1972.

Palmer, S.E. Transformational structure and perceptual organization. Proceedings of the 3rd annual conference of the Cognitive Science Society. 1981.

Rosch, E. Cognitive reference points. Cog. Psychol. 1975, 7, 532-547.

Rosch, E. Principles of Categorization. In: *Cognition and Categorization*, E. Rosch & B.B. Lloyd (Eds) Hillsdale, NJ: Lawrence Erlbaum. 1978

Rock, I. *Orientation and Form*. New York: Academic Press, 1973.

Rock, I. *Introduction to Perception*. New York: Macmillan, 1975.

Wiser, M. The role of intrinsic axes in shape recognition. *Proceedings of the 3rd annual conference of the Cognitive Science Society.* 1981.

*Copies available from:
Michael Leyton,
Department of Psychology,
University of California,
Berkeley, CA 94720.