30

Issues in the Theory of Action: Degrees of Freedom, Coordinative Structures and Coalitions

Michael T. Turvey
University of Connecticut
Storrs, Connecticut
and
Haskins Laboratories
New Haven, Connecticut

Robert E. Shaw
University of Connecticut
Storrs, Connecticut

William Mace
Trinity College
Hartford, Connecticut
United States

ABSTRACT

Two major problems for a theory of coordinated movement are considered: The context-conditioned variability in the terminal variables of the motor apparatus, and the management of the large number of degrees of freedom that the motor apparatus attains. One approach to these problems argues that, in general, the free-variables of the motor apparatus are not controlled individually, but are partitioned into a smaller number of collectives, where each collective regulates internally and relatively autonomously a number of degrees of freedom. Another and closely related approach seeks to identify a system in which the responsibility for planning and executing an act is optimally distributed across the components of the system. This paper develops the concept of coordinative structure in concert with the first approach. Using the second approach, the paper lays the ground for showing that the perception and action systems participate in a style of organization in which the operational component and the context of constraint are wedded together into a single, relatively closed system. Such a system is referred to as a coalition, and it is claimed that a coalition is the minimum organization required to solve the problems of context-conditioned variability and degrees of freedom.

THE PROBLEMS OF CONTEXT-CONDITIONED VARIABILITY AND DEGREES OF FREEDOM

A. Two Kinds of Machine

There are two quite dissimilar approaches that might be taken toward designing a machine that acts within and upon an environment (cf. Greene, in press, a). In one approach, the mechanical variables of the machine are programmed so that a desired action is achieved by virtue of a single computation, which in a single instance specifies all the necessary details, including those needed to immunize the machine against perturbing influences. In the other, one begins with the construction of autonomous systems designed, in part, to preserve the stability of the machine in its intended environment. Then, given these systems that perform as they wish, the problem is to organize them in such a fashion that the systems, and thus the machine, perform as we wish.

In the first kind of machine executive procedures are called upon to control individually each mechanical variable, whereas in the second they are called upon to control the autonomous systems, with each system regulating internally a subset of the mechanical variables. Let us elaborate: Where the total number of mechanical variables is quite small, the first design is obviously felicitous; but where that number is large, then it is roughly apparent that the first design will prove to be overly cumbersome and costly. By contrast, a machine of the second kind would be inelegant where few mechanical variables were concerned, since any arbitrary configuration of those variables could not be achieved directly, but only indirectly through the modulation and interaction of the autonomous subsystems. The advantage of the second kind of machine, however, is that, given the right organization, it may achieve an approximation to a desired configuration of a very large number of mechanical variables through the regulation of the relatively few variables of the autonomous subsystems.

B. The Keyboard Metaphor: Address-Specific (Individualized) Control

One nineteenth century view of the coordination of voluntary movement (see Luria, 1966) was that it was exclusively a function of efference, specifically of the giant pyramidal cells of the cortical motor strip. The corticospinal projection was thought to contain all the details relating to the spatial and temporal patterning of commands to muscles. The motor area of the cortex was regarded as a kind of keyboard on which an executor's "hand" could play out the score for a movement. "Pressing" (exciting) one key brought about a determinate degree of extension at a given joint, pressing another brought about a determinate degree of flexion, and so on. This classical view promotes a machine of the first kind since it assumes that a coordination of movements can result from a single stage of exact computation. It allows that a motor program can be written as machine language instructions, that is, in terms of the innervational states of individual muscles. Insofar as this view assumes that control proceeds from the top down to a specific address – the individual muscle or the individual motoneuron – we will refer to the style of control as address-specific (individualized) control. Let us examine the keyboard metaphor and its associated style of control from the orientation of two major problems that any theory of coordinated movement – of action – must resolve.

The joints and the permissible motions of the complex biokinematic chains that compose the skeletomuscular hardware of animals, comprise a large number of degrees of freedom [Bernstein (1967) suggests that it may reach three figures.] For each coordinated act, therefore, the values for each individual degree of freedom would have to be prescribed. To be more precise, the executive at the keyboard would have to specify from the set of all possible combinations of muscle contractions the particular combination that would achieve the desired objective. The executive problem in this case is analogous to that of finding the optimum of a function of many variables. Algorithms that theoretically allow the solution to such problems prove to be infeasible in practice. Thus one could differentiate the function with respect to each of its variables and equate the derivatives thereby obtained. But the solution of the resultant system of equations is no simpler a problem than the direct search for an optimum. Gel'fand and Tsetlin (1962) remark that even where the number of variables (degrees of freedom) is relatively small, say, four or five, the computation of an extremum or optimum is exacting and often impossible for contemporary computational procedures (see also Sivazlian & Stanfel, 1975). An algorithm may exist, but only with respect to an abstract machine. A physically realizable machine brings with it extralogical or extra-algorithmic principles (cf. Shaw & McIntyre, 1974) best understood as cost variables, such as how much heat the components of the machine can tolerate, and how much time the machine can spare for computation. An animal that must react adaptively to the contingencies of its environment does not have infinite time to select the right combination of biokinematic variables; its best policy, argue Gel'fand and Tsetlin (1962) and Greene (1972, in press, a), is to aim for a rough approximation through the use of "quick and dirty" procedures. At all events, we will refer to the problem described above as the problem of *degrees of freedom.* ①

The second major but related problem is posed by *context-conditioned variability.* - ② The keyboard variables – the individual muscles and their innervational states – do not have fixed movement consequences. On the contrary, these variables relate to movement in a way that is dependent on contextual contingencies. If the nineteenth century executive were truly indifferent to context (as the notion of control solely through efference suggests) and wrote his movement score as a function, with the states of individual muscles (m) as its arguments, that is, as $f(m_1 m_2 . . . m_n)$, then the function would be indeterminate in that for given *ms* and for given values of these *ms* the resultant act would not be fixed. In short, the specified function and the resultant behavior would relate equivocally. The reason, in part, is that other variables that affect the biokinematic chains are not specified, are not taken into account, in the function. To develop this point let us consider various forms of context-conditioned variability as delineated by Bernstein (1967).

C. Sources of Context-Condition Variability

① In the first place, there is the anatomical source of context conditioned variability: The muscles can vary in their roles with regard to joint movement. Consider, by way of example, the upper pectoralis major, which inserts proximally in the clavicle and distally in the upper shaft of the humerus. With the arm in an approximately horizontal position, in which the axis of the humerus is just below the horizontal axis of the shoulder joint, contraction of the pectoralis will adduct the arm in the horizontal plane. But from an approximately horizontal position, in which the axis of the humerus is slightly *above* the horizontal axis of the shoulder joint, contraction will adduct the arm in the *vertical* plane (Wells, 1961). The moral (for a brain as well as for a student of kinesiology) is that a muscle's role cannot be taken for granted; at each phase of a movement, an individual muscle's action is contingent on (among other things) the muscle's line of pull to the joint's axis of motion.

Cognate with this class of equivocalities is the realization that the role a muscle plays depends not only on the disposition of limb segments but also on the external force contingencies. Lowering the arm from a horizontal side position against a resistance requires the use of the adductors of the arm, notably, the lattisimus dorsi; but in lowering slowly (that is, against gravity) the adductors are palpably soft, for the responsibility of the movement befalls the adductors, the deltoids, which perform their task by lengthening or, as Hubbard (1960) prefers to call it, pliometric contraction

Ideally, agonists and antagonists are paired in the keyboard arrangement so that an executive might simply excite alternatively one and then the other. If the machine being instructed was composed of hinge joints, each of one degree of freedom, then a fixed mechanical arrangement could be assumed between an agonist-antagonist pair, and the ideal keyboard arrangement would be realizable. But where ball joints are concerned, such as at a hip or a shoulder, the collection of muscles that is instrumental in moving the limb and the collection opposing the movement vary with each trajectory. Rather than there being a fixed anatomical relation between agonist and antagonist, there is, on the contrary, a problem of having to choose the muscles to be employed in a given movement (Weiss, 1941).

② In the second place, there is the context-conditional variability resulting from mechanical sources. Most notable among these is the fact that, depending on the dynamic and static conditions of the limb segment, the same innervational state of a muscle may give rise to a variety of motions of the segment differing in displacement and velocity, and different innervational states may produce identical motions. The lesson here is a simple one: The innervational states of an individual muscle and the movements they entail relate equivocally.

A closely related source of variability is indigenous to multilink kinematic chains of which a whole arm, a whole leg, or the whole body are examples. Quite simply, the movement of any one link will result in a displacement of the links attached. The consequence of this is that the attached, or "light" links, passively carried by the agonist, or "heavy" link (in the terminology of Eshkol and Wachman, 1958), will induce forces and moments as reactions to the "heavy" links trajectory and thereby complicate its control. From the perspective of the keyboard metaphor and exclusively efferent control, multilink biokinematic chains look capricious. Closer examination, from a different perspective, reveals, however, that facility with a gross body skill is synonymous with exploiting these reactive consequences to the fullest (Bernstein, 1967). One characterization of the skilled performer is that, of the changes in forces at the joints necessary for a given movement, the performer provides (economically) only those changes in forces that are not provided reactively.

Most evidently, the two mechanical sources of context-conditioned variability go hand-in-hand. Thus, because the links have mass, once impelled, they gather momentum and develop kinetic energy. A given degree of muscle activity acting against a movement may stop it, simply retard it, or even reverse it; the same degree of activity, in concert with the movement, may induce marked acceleration. What follows from a given degree of muscle activity depends on the kinetic conditions of the links. While the significance and ubiquity of this principle was ignored by proponents of the keyboard metaphor (see Bernstein, 1967), it is also given short shrift in contemporary theories of coordinated movement (see Stelmach, 1976). In part, this negligence seems to be due to the assumption that the innervational states of muscles are in phase or concurrent with the movements of biokinematic links (Hubbard, 1960), an assumption that deserves our attention, if only briefly.

The "in-phase" assumption is a most convenient one because, as Hubbard (1960) elegantly points out, it permits the luxury of inferring muscle events from movement events. For example, as the elbow flexes, the biceps shortens and the triceps lengthens, from which we might infer that there was continuous graded stimulation contracting the one muscle and relaxing the other.

Fast movements, often referred to as ballistic, are anomalous from the perspective of the "in-phase" assumption. Their control is characteristically "bang-bang" (Arbib, 1972): an initial burst of acceleration as the agonist contracts, an intervening period of inactivity and then a burst of deceleration as the antagonist acts to degenerate the kinetic energy of the link. In fast movements muscle activity is simultaneous with only a small portion of the movement. But perhaps the "in-phase" assumption does hold for movements conducted at a more leisurely pace, movements that we might refer to as nonballistic. Hubbard has argued and demonstrated that even here, the "in-phase" assumption is found wanting (Hubbard, 1960); as far as he can discern, the basis of slow movements is the same as that of fast movements – that is, discrete bursts of muscle activity that alternately act to accelerate and decelerate the link. The control of slow movements, in this perspective, is characteristically "bang-bang-bang," and so on. There is some support for this characterization (e.g., Aizerman & Andreeva, 1968; Chernov, 1968; Litvintsev, 1968). It appears that the slow movement of a link, say, wrist extension or elbow flexion, is the result of pulls by both opposing muscles, where each muscle pulls ten times per second (Aizerman & Andreeva, 1968; Hubbard, 1960), first one and then the other.

③ In the third and final place we may recognize the context-conditioned variability that arises by virtue of the physiology. We can relate here only a small part of what is, most obviously, a very lengthy story.

The motor unit, conventionally defined as an alphamotoneuron together with the bundle of extrafusal muscle fibers that it innervates, may be considered as the functional final common path. Alphamotoneurons sometimes have monosynaptic connections with several descending systems. However, these monosynaptic projections to alphamotoneurons represent but a small part of the total neural projection to these cells, and of themselves probably do not bring about motoneuron firing (Evarts, Bizzi, Burke, DeLong, & Thach, 1971). In very large part, the major influences ultimately exerted on motor units occur via the segmental interneurons so that their modulation of motoneuron activity is highly flexible. Significantly, the same descending "instruction" might at different times encounter quite different "states" in the segmental interneurons; its affect, therefore, on a target motoneuron, is open to considerable variation.

The point is that the segmental apparatus of the spinal cord is an active apparatus that does not passively reproduce supraspinal instructions (Paillard, 1960). On the contrary, it appears that supraspinal and spinal influences relate in a free-dominance fashion (see below) in the coordination of acts. There is evidence that the state of the segmental apparatus can (among other things) convert a flexion reflex into one of extension (Lisin, Frankstein, & Rechtmann, 1973), enhance or inhibit contractile states evoked by cortical stimulation (e.g., Gellhorn, 1948), and affect the latency of voluntary movment (Gurfinkel' & Pal'tsev, 1965).

The implication of context-conditioned variability is simple enough: The motor apparatus cannot be regulated solely by efferent impulses arising in the cells of the motor strip. The contemporary perspective on coordinated activity contrasts with the nineteenth century perspective in recognizing that a continuous afferent flow of exteroceptive and proprioceptive information is the backdrop against which acts are constructed. Indeed, for Bernstein (1967) the decisive factor in coordinated activity is not the efferent impulses but the complex system of afferentation that tailors the components of the activity to the prevailing contingencies.

Patently, the problem of context-conditioned variability could be solved for the first kind of machine we described above by making available to it detailed information about the current states of the muscles and joints. It is commonly understood that signals to the sensorimotor cortex refer to muscle tension, muscle length, joint angle changes, and their time derivatives (Granit, 1970). It follows, therefore, that the flavor of the nineteenth century keyboard metaphor might be preserved by allowing that the specifications of individual muscle variables be guided by detailed information on the muscles and the positions and motions of the individual biokinematic links. A fine-grain description of afference might be coupled to the fine-grain description of efference, and all the details of the act – all the individual degrees of freedom – computed in a single step. It is our impression that a hybrid perspective much like this characterizes a number of current accounts of coordinated movements and motor programs (e.g., Keele & Summers, 1976; Schmidt, 1976). Unfortunately, this perspective, while proposing a resolution to the problem of context-conditioned variability, does not address the problem of degrees of freedom. We might even claim that it compounds the latter problem by requiring that the values of the large number of degrees of freedom on the input side be mapped by the executive onto values for the large number of degrees of freedom on the output side.

II. THE CONCEPT OF COORDINATIVE STRUCTURE

There are two closely intertwined and popular approaches to solving the related action problems of degrees of freedom and context-condition variability. Both approaches are evidenced to a degree in the second kind of machine described above. One approach seeks an optimal grain-size for describing the skeletomuscular units that serve as the vocabulary for acts. The other approach seeks to describe and understand an organizational format in which the responsibility for planning and executing an act is optimally distributed across the various computational components of the acting system.

In this section we will consider the first approach. In particular, we will examine and elaborate on the point of view that the free-variables of biokinematic chains are able to be partitioned into collectives (Gel'fand, Gurfinkel', Tsetlin, & Shik, 1971), where the variables within a collective change relatedly; and that the action vocabulary is these collectives, rather than the individual degrees of freedom.

In the literature, collectives of biokinematic variables take a variety of forms and are given a variety of labels. Our choice is *coordinative structure* (Easton, 1972; Turvey, 1977), which we will define generally as *a group of muscles, often spanning several joints that is constrained to act as a unit.* A coordinative structure is a relatively autonomous system: It regulates internally a number of biokinematic degrees of freedom, but is itself to be regarded as a single degree of freedom. There are grounds for distinguishing between marshalling of such automatisms and their modulation or tuning (see Boylls, 1975; Greene, 1972, in press, a; Turvey, 1977). Marshalling a coordinative structure may be characterized as defining the "ballpark" of a component activity, whereas tuning a coordinative structure may be characterized as tailoring the component activity to current contingencies, that is, making appropriate adjustments within the "ballpark" (Greene, 1972, in press, a).

Notable sources of tuning are movements of parts of the body, in particular the head and the eyes (Easton, 1972; Turvey, 1977), and memory (Boylls, 1975). But the primary source of tuning, in most natural circumstances, must be the detection of information about the relations among the body parts (propriospecific information), information about the properties and layout of the environment (exterospecific information), and information about the relation of the body to the layout of the environment (expropriospecific information) (see Lee, 1978). The informational support for tuning (and for coordinated movement in general) is amodal. The three kinds of information described above are secured through the partially overlapping sensitivities of the various perceptual systems (Gibson, 1966), although vision is the most bountiful and oftentimes the most reliable supplier (Lee, 1978).

A brief examination of the guidance system of an airplane follows. It illustrates the conception of a coordinative structure and provides a framework for understanding tuning.

A. The Guidance System of an Airplane as an Illustration

The kinematic state-set of an airplane consists of all the configurations into which its movable parts can enter; for purposes of illustration, only those movable parts that must be coordinated in order to guide the airplane in flight are considered.

An airplane typically has a control system with a minimum of five hinged parts: two ailerons on the rear edge of the wings that can be moved up or down to control roll; two elevators on the horizontal portion of the tail section that can also be moved up or down to control pitch; and finally, the rudder on the vertical tail fin that can be moved left or right to control the yaw of the aircraft. Construed as a *freely-linked* kinematic chain, the airplanes guidance control system has five degrees of freedom, one degree of freedom for each hinged part. Each degree of freedom provides a coordinate dimension for the state-space comprising all possible configurations of the five movable components. Or, put differently, each possible configuration of the five control components can be represented by a point in a hyperdimensional kinematic space of five dimensions. Such a space of possible configurations is clearly too complex to be imagined, much less mastered, for manually controlled flight. It is instructive to compute the information load placed on a pilot in the above situation, where each aileron, elevator, or rudder has to be independently controlled; that is, where the style of control is address-specific with each kinematic degree of freedom individually controlled.

Assume that although each hinged part can move continuously, only k number of discrete positions are truly effective because of the inertia of the aircraft in flight. Thus, the total number of possible discrete configurations of a system with five free kinematic links is k^5. However, for the sake of simplicity, let k take on only eight values – a conservative estimate, since most airplanes have considerably more sensitive guidance control systems. This system's total kinematic state-space would consist of $k^5 = 8^5 = 32,768$ independent states. Thus, the information load on a pilot who must manually select a particular configuration of the guidance control system in order to select a desired flight pattern is $\log_2 n$, where n is the total number of independent choices. For our particular example, $n = 8^5 = (2^3)^5 = 2^{15} = 32,768$ and $\log_2 n = 15$ bits.

To grasp how complicated this task would be, consider a more familiar case: If we assume a comparably sensitive manual control system for an automobile (i.e., $k = 8$) with three degrees of freedom (accelerating, steering, and braking), then only 9 bits of information must be processed by the driver for each guidance decision [$\log_2 8^3 = \log_2 (2^3)^3 = 9$ bits]. However, even here we are stretching the limits of human information processing (Miller, 1957).

In the case of the 15 bit information load demanded of the pilot of the airplane with the crude guidance system described above, we are dangerously in excess of the average load believed permissible. Clearly, a safe aircraft requires a guidance system with more manageable degrees of freedom, so as to reduce the information processing load on the pilot. This can be accomplished by imposing constraints on the guidance control system.

The kinematic links of the airplane can be constrained in the following way: Let the ailerons of each wing be inversely yoked so that one moves up as the other moves down, and vice-versa; yoke the rudder on the vertical fin of the tail section to the ailerons, so that it moves left when the right aileron is depressed, and vice-versa; and yoke the elevators on the horizontal portions of the tail section so that they move in unison, going up and down together. With its movable guidance components so constrained, the airplane has a greatly reduced kinematic state-set and can be controlled more simply. The guidance system now has but two degrees of freedom, namely, the aileron-rudder sybsystem, where the inverse up-down movement of the ailerons is mechanically coordinated with the left-right movement of the rudder (a *macro* with one degree of freedom). Assuming that this coordinated guidance system is at least as sensitive as the cruder version, then we may note that its state-set is $k^2 = 8^2 = 64$. The latter number represents a dramatic reduction of 32,704 from the original state-set, consisting of 32,768 possible unconstrained configurations. More importantly, however, these constraints on the design of the guidance control system achieve a 500 fold reduction in the information processing load placed on the pilot (that is, $2^{15} - 2^6 = 2^9$ or 512).

The joystick, through which the pilot exerts his control of the airplane, links the two subsystems so that they are mutually constrained to act as a unit. It follows from the immediately preceding paragraph that the significant feature of the guidance system so produced is that the joystick represents for the pilot but two degrees of freedom, each defined over eight values – an information processing load of six bits (that is, where $n = 8^2 = (2^3)^2 = 2^6 = 64$; and $\log_2 n = \log_2 64 = 6$ bits). By moving the joystick left or right (one degree of freedom), the plane banks or turns in either direction; by moving the joystick forward or backward (a second degree of freedom), the plane's angle of ascent or descent is controlled.

Patently, the two-dimensional space of the joystick is the control space for the coordinative structure formed by constraining the freely moving parts of the airplane in the manner described. (With the free-variables left unconstrained, the control space was five-dimensional.) Furthermore, we may note that the two-dimensional control space represented by the joystick's movements is synonymous with the coordinative structure defined over the freely moving parts; the two concepts – control space and coordinate structure – are formally equivalent.

Finally, we recognize that the airplane's guidance system provides a good illustration of the radical difference in efficiency that exists between a system in which the free-variables are controlled individually, and a system in which it is collectives of free-variables that are controlled.

B. Examples of Coordinative Structures

Let us now take a look at some of the biokinematic events that have been (or may be) promoted as instances of the concept of coordinative structure.

The activity of a single limb during locomotion consists of two broadly defined phases: support and transfer. The support phase, during which the fact is in contact with the ground, is composed of extensor activity over the limb joints; the transfer phase, carrying the foot from one support to the next, is composed essentially of flexion. In an ingenious experiment (Orlovskii & Shik, 1965) conducted with dogs locomoting freely on a treadmill, a very brief impedance was applied at the elbow during transfer-flexion. In consequence, the movement at the elbow was slowed but so was the movement at the shoulder and wrist. However, a similar impedance delivered during support-extension did not retard the movement at the other joints. It is arguable that the link motions during flexion are constrained to act as a unit by means of spinally-mediated afferentation (cf. Boylls, 1975). But what of the extensors? They, apparently, are not linked by shared afference, but they do appear to be linked – that is, they do behave as a unit during locomotion. Witness to this claim is the observation that, across various gaits, the timing of limb extensor EMGs is nearly invariant with respect to step cycle and, further, that the activity periods of extensor muscles relative to each other change little as speed of locomotion changes (Engberg & Lundberg, 1969). The implication, perhaps, is that in locomotion the limb extensors are constrained to act as a unit by means of common *efference* (Boylls, 1975).

A unitary arrangement of joint changes that has been investigated quite thoroughly and that, therefore, provides an exemplary case, is that which preserves the stability of the head during respiration (Gurfinkel', Kots, Pal'tsev & Fel'dman, 1971). With inspiration and expiration, the torso (in both its upper and middle parts) deflects backwards and forwards, respectively. The displacement is of sufficient magnitude such that, if left unchecked, marked excursion would occur in the overall center of gravity. However, the respiratory-induced oscillations in the torso are balanced by antiphasic oscillations at the hip and at the cervix. Changes in the angle of the hip and of the cervix are simultaneous with changes in the angle of the torso, and the relation among these changes in invariant with frequency of respiration. This constraint on the biokinematic chain is wrought neither by means of mechanical conspiracy nor by spinally-mediated afferentation (Gurfinkel', Kots, Pal'tsev & Fel'dman, 1971); as with the extensors during locomotion, the coupling source is probably efferent.

Controlling two joints of the arm provides a further case in point. When a person is requested to simultaneously flex or extend his wrist and flex or extend his elbow, the joints are moved mainly in a coupled fashion (Kots & Syrovegin, 1966), although this synchrony is achieved with less practice in the case of changes of the same type (e.g., flexion-flexion) than in changes of the opposite type (e.g., extension-flexion) (Kots, Krinskiy, Naydin, & Shik, 1971). Significantly, the two rates of change of joint angle preserve one or another invariant ratio that is not attributable to mechanical coupling. Individuals differ in the ratios they use and they tend to have three to seven such ratios. Furthermore, they use a different subset of these ratios (usually three or four of them) for each of the four combinations of flexion and extension (Kots & Syrovegin, 1966).

Finally, let us take note of observations on the production of speech that suggest that oftentimes movements of the tongue, lips, velum, and jaw may be constrained as a unit (Kent, Carney, & Severeid, 1974).[2] To illustrate, in uttering the word *contract* lowering the velum is initiated with the release of oral closure for /k/, and elevating the velum begins with the tongue tip movement for alveolar closure (Kent et al., 1974). In uttering the word *we,* the transitions from the glide /w/ to the vowel /i/ is mediated by the contemporaneity of a forward gesture of the tongue body, and a release of lip protrusion. With increase in emphatic stress, there is an increase in the displacement and velocity of the tongue body and in the displacement and velocity of the upper lip. However, the relation

[1] We thank Carol Fowler for bringing these particular observations to our attention.

between the lingual and labial displacements and velocities remains invariant over variations in stress (Kent & Netsell, 1971). Apparently, for utterances like /wi/, the stress contrast modulates *both articulators or neither articulator.*

It is dimly apparent from these examples that where several muscles are constrained as a unit, whatever the mechanism, the activities of the individual muscles covary in terms of a ratio that is relatively fixed and indifferent to overall magnitude changes in these activities. In reaching this tentative conclusion we are somewhat guilty of the "in-phase" assumption, for our examples have crossed the muscle state-link movement boundary and, glibly, we are treating the two as isomorphic. Nevertheless, we believe the conclusion has heuristic merit and, following Boylls (1975), we proceed to identify two prescriptions for a ccordinative structure.

C. Structural and Metrical Prescriptions

A "structural prescription" refers to the ratios of activities in the muscles composing a coordinative structure that are invariant with respect to absolute activity level. As Boylls (1975) remarks, a metrical prescription is like a "scalar" quanitity, which multiplies by the same amount the activities of each muscle in a coordinative structure. In the example just given, of uttering /wi/, emphatic stress is analogous to a metrical prescription, for it magnifies the lingual and labial activities to the same degree; the ratio between the two activities that is preserved over stress is the structural prescription.

The specification of structural and metrical prescriptions for coordinative structures is, in part, what is meant by tuning. A change in structural prescription changes the dynamic topography of a biokinematic chain whose links have been constrained to act as a unit; a change in the metrical prescription changes, among other things, the speed with which the dynamic topography is realized. Our guess is that metrical prescriptions can be modulated more rapidly and with greater facility than structural, and there are a few experiments in favor of this view.

From the work of Asatryan and Fel'dman (1965) and Fel'dman (1966) we learn that where the muscles at a joint have been constrained to act as a unit – either for the preservation of a particular posture against opposing moments of force, or for the purpose of moving, on signal, to a new prescribed position, again against opposing moments of force – the muscle complex is describable as a nonlinear spring with definite stiffness and damping parameters. In the case where a posture is to be maintained, if the opposing moments of force are changed unexpectedly, the limb segment moves initially to a posture that is in accord with the original parameters, and only then does it move to a posture that is in accord with the new parameters relevant to the new moments. In the case of moving to a prescribed position, if the moments are changed subsequent to the signal to move but prior to movement, the limb will move initially, but erroneously, to a position that would be predicted for the "spring" parameters present at the time of the signal. One might interpret these observations to mean that once a coordinative structure has been activated, the parameters of that structure cannot be modified until the task, for which it was set, is complete. But a more prudent interpretation is that the temporal scale over which changes wrought through tuning can occur does not always overlap the temporal scale over which can occur changes wrought through generated kinetic energy. Where the scales do overlap, the personality of a coordinative structure can change – in flight, as it were.

It can be shown by experiment (Vince & Welford, 1967) that a movement by a hand begun slowly in response to a signal for a "slow movement" can be accelerated in response to a further signal, one that is for a "fast movement," in very much the same time that it would take to initiate a fast movement from rest. And this is so even if the second signal arrives during the latent period of the first. On this experiment and another (Megaw, 1970), in which the second signal called for a slightly different movement from the first signal, it appears that the form of an "initiated" movement is less rapidly altered than the vigor with which it is conducted. In our terms, structural prescriptions are less rapidly alterable than metrical prescriptions.

An especially interesting illustration of metrical modulation is to be found in the activity of the baseball batter (Hubbard & Seng, 1954). In this illustration we can point to the derivative properties of the optical flow field at the eyes as the information for metrical prescription. As with all batting skills, it is mechanically advisable to move in the direction of the ball. The right-handed baseball batter does so by lifting his left and leading foot, moving it forward and parallel to the ball's line of flight, to finally place the foot some distance in front of, and probably slightly to the side of, the foot's initial position. The start of this step is synchronized with the release of the ball from the pitcher's hand. The duration of the step, however, and the start of the swing (which more often than not coincides with the completion of the step) are inversely related to the speed of the ball, to which the speed and duration of the swing remain relatively indifferent (Hubbard & Seng, 1954).

We may consider the act of batting as supported by a function defined over a small number of coordinative structures, which for present purposes suffices as our definition of an action plan (Turvey, 1977). It can be hypothesized that the batter's stepping pattern arises primarily from the activities of knee extensors and hip abductors and flexors constrained to act as a unit. A structural prescription on this coordinative structure defines the dynamic topography of the stepping movement. On release of the ball the batting plan is initiated; during its unfolding, the plan is tailored to the current contingencies by the optically specified metrical prescription: The duration of the step (and hence the initiation of the swing) is functionally related to the speed of the ball.

In this last example we catch a glimpse of a central problem for the theory of how acting and perceiving conflate: In the performance of acts exterospecific, propriospecific, and expropriospecific information must be selectively percolated through the action structures at the right time. Conventional theories of selective attention do not address the question of how the selection of information is temporally constrained so as to be compatible with the dynamical requirements of the system it serves.

An interesting observation of Orlovskii's (1972) may have some bearing on this problem. Given supraspinal stimulation of the spinal cord known to enhance flexor and extensor contraction in the inert animal, it was shown that when this stimulation was continuous with locomotion, the effects of the stimulation were manifest only at select points in the locomotory cycle. One might interpret this result as saying that the interaction of coordinative structures created "holes" or "slots" through which the continuously present supraspinal influences could "flow" (cf. Boylls, 1975. Is this an instance of a general principle? In that the visual information that supports activity is not characterizable as momentary signals or stimuli but as continuous optical flow fields (Gibson, 1958; Lee, 1974), may we conjecture that the "introjection" of information into an act is constrained by the interaction among coordinative structures mediating the act? That selective percolation at the right time is defined, in very large part, by the act itself?

D. The Concepts of a Control Space and a Moving Point of Control

Consider once again the airplane guidance system described earlier. The concept of a control space is instanced by the two-dimensional space describing the joystick's movements, and it provides an elegant way of characterizing the two tuning functions of metrical and structural prescriptions. The two subsystems of the airplane's guidance system and the control space are depicted in Fig. 1. In

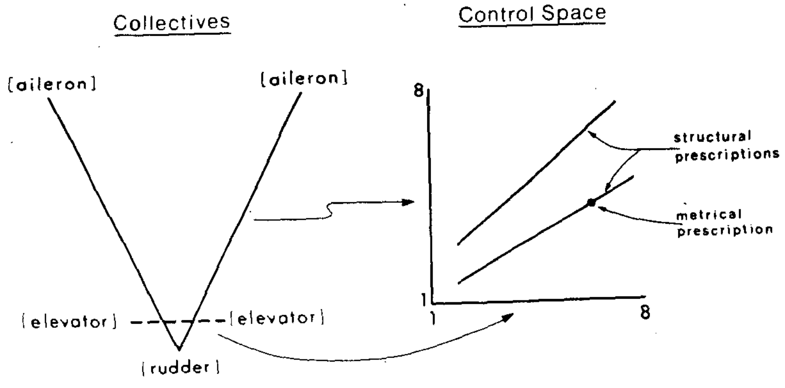


FIGURE 1.

reference to a task such as that of Kots and Syrovegin (1966), one of the subsystems can be likened to the collective of muscles regulating the change at the elbow, and the other can be likened to the collective of muscles regulating the change at the wrist. As Fig. 1 shows, any point in the control space is a metrical prescription, since it defined values for the variables of the collectives; a line in the control space is a structural prescription, since the slope of a line in the space defines a particular relation holding between the kinematic values of the two subsystems, that is, a ratio that is preserved invariant over metrical change.

The potential usefulness of the concept of a control space suggests that we should seek a way of determining such spaces for the biokinematic variables of natural systems. Consider a kinematic analog to animals and humans: A robot with the capability of locomoting across a room and opening a door. We restrict ourselves to considering simply the control of a single limb of the robot. Let the robot's limb consist of two rigid segments of unequal length, connected together by a joint (an "elbow") that permits 360° rotation in the plane, with this articulated limb connected by a similar circular joint to the main body of the robot (a "shoulder"). What is the structure of the control space of this two-joint system, and what is the minimal number of degrees of freedom needed to describe it?

See math article in Sci. Am May or June 1984

The robot arm described above is, essentially, a compound planar pendulum with two kinematic degrees of freedom. In general, a precise representation of the total kinematic state-set of a mechanical system of l degrees of freedom can be provided by a control space defined over l parameters. The sturcture of the control space of the limb can be determined by taking the (topological) product over the unrestricted motions of the multiple linkages of its freely jointed kinematic chains. In the case of our *idealized* robot, the segment extending from the shoulder can rotate through a 360° planar angle around the shoulder joint; a similar rotation is permissible for the lower segment connected to the elbow joint. Furthermore, since these two segments are independent, for each angle that one of them assumes, the other is free to assume any one of its continuous angular positions. Thus, all possible positions of this articulated limb may be represented as the topological product of the angular positions of two circles. This topological product determines a manifold with two degrees of freedom known as a *torus* – a closed surface of two dimensions resembling the outside of a bagel.

Intuitively, one can conceive of the topological product of two circles as that surface (a torus) generated by stringing a small circle on a large circle and then moving the small circle in such a way that the larger circle consistently penetrates its center. All the points on the surface of this object specify possible kinematic states, or configurations, of the articulated limb of the robot. To locate a specific kinematic state on this surface requires imposing a coordinate system on the surface. Such a coordinate system is readily provided by dividing the circumference of the torus into degrees and, similarly, by dividing the circumference of the small circle into degrees. These two coordinate dimensions – two degrees of freedom – can then be used to locate every possible combination of circular joint values permitted by the kinematic linkages of the robot's limbs. If the values assumed by the two joints are truly independent of one another, then the torus provides a natural model of the control space of the robot's arm because it represents every possible kinematic state and no impossible ones. A person might have thought that a surface like a sphere or an ellipsoid would have done just as well, but neither of these is the product of variable-sized circles. The sphere is a product of equal circles and the ellipsoid is generated by a circle and an ellipse. The torus is the only closed surface that can be generated as the topological product of two *variable* circles.

Of course, the control space described is too idealized even for our robot. Neither of its joints will really allow a full 360° of rotation because we restricted the joints to the same plane. Thus, like the human arm, the robot's segmented limb will be restricted in its freedom of rotation. In general, the individual kinematic links of an animal or a human are anatomically constrained as to restrict free variation – a fact that must be expressed in their control spaces. In illustration, suppose that the shoulder joint of the robot permits free variation of its kinematic link through an angle limited to 180°, whereas the elbow joint permits free variation of its kinematic link through only 90°. Then the restricted control space that represents these natural anatomical constraints is but one quarter of the surface area of a half torus (cut in the way one would halve a bagel). In general, the natural, anatomical constraint placed on the degrees of freedom of ideal kinematic systems can be represented as bounded portions of the ideal topological manifold, corresponding to the control space of that system.

It is worth noting that the method of taking topological products over kinematic chains with more liberal joints will produce higher dimensional manifolds that represent control spaces of such systems. For instance, an articulated limb consisting of a ball joint and a 360° hinge-joint yields a topological product specifying a closed manifold of points – a control space – of three degrees of freedom, namely, the part of the space lying between two concentric spheres.

In the preceding we have demonstrated how the kinematic state set of a collection of biokinematic free-variables can be reduced to a minimal representation as a control space. This reduction has the virtue of providing an ideal solution to the problem of determining the minimal degrees of freedom required for the control of a coordinative structure by tuning.

To summarize: The concept of a control space is formally equivalent to that of a coordinative structure; and tuning a coordinative structure is formally equivalent to *a moving point of control* in the control space. The position of the point of control defines the metrical prescription and the direction of movement (a line, a plane, or a hyperplane) through the space identifies the structural prescription.

Let us now return to the airplane guidance system. By linking together the parts of the guidance system in the manner described, a system of five degrees of freedom is reduced to a system of two degrees of freedom. Put in a slightly different way, the linkage constrains three degrees of freedom and leaves unconstrained two degrees of freedom. Consequently, to complement the constraint supplied by the linkage, the source of control of the joystick must possess at least two degrees of freedom. In the course of piloting the plane, the pilot supplies the requisite constraint not supplied by the linkage. If for some reason the pilot could move only his right arm, and then only in a plane parallel to the sagittal plane – that is, forward or backward – then the pilot would not be able to control the plane. The single degree of freedom at the pilot's disposal is not the complement of the guidance system.

What does the moving point in the control space represent in natural systems such as animals and humans? After all, unlike the case of the airplane control space, there is neither joystick nor pilot to guide the moving point in a literal sense.

As suggested earlier, it is reasonable to assume that in animals and humans the constraints that complement the available degrees of freedom of the coordinative structures and tune them to their precise parametric values, originate primarily in the environment, and are picked up through perceptual activities. If this is so, then it strongly suggests that the unit of analysis for action must be of a grain sufficiently coarse to include, in addition to the anatomical and physiological aspects of the actor, certain relevant portions of the perceptual environment. In other words, the theoretical analysis must be at a grain that is truly ecological and not just psychological or physiological on the one hand, nor just physical or informational on the other (see Fowler & Turvey, in press; Shaw & McIntyre, 1974).

III. TOWARD A DEFINITION OF COALITIONS

Later in this section we will present a quasi-formal analysis of different organizational styles as models of control systems. This analysis carries us toward a formal definition of a coalition and a potential resolution to the problems of context-conditioned variability and degrees of freedom.

To set the stage for this analysis, we must make some prefatory comments on major principles of organization in complex systems; we do this through a brief comparison of a hierarchy and a heterarchy. In addition, we will present an example of one principle – that of free dominance – as manifest in the relation between preserving balance and performing acts that create imbalance.

A. Preliminary Remarks on Organizational Principles

A familiar form of organization is the hierarchy: The burden of computation on higher stages is alleviated by apportioning minor computation and processing to lower stages. We (Shaw, 1971; Shaw & McIntyre, 1974; Turvey, 1977) and others (e.g., Minsky & Papert, 1972) have claimed, however, that the hierarchy does not do justice to the organizational style manifest by complex dynamical (biological) systems. A hierarchy does not appear to provide the necessary computational power (see Sutherland, 1973); and significant biological events, such as remodeling or aging, are not easily defined over a hierarchy (cf. Shaw & McIntyre, 1974). Let us consider, therefore, in these preliminary remarks, the distinction between a hierarchy and a style of organization that, in theory, approximates more closely the biological form, namely, a heterarchy.

A pure hierarchy is characterized by the unidirectionality of commands or information flow. Given two structures or systems that are at different levels of the hierarchy then A, the higher system, always commands B; B never commands A. In a pure hierarchy, therefore, the relationship between any two levels of the hierarchy is immutable. In less pure forms, this unidirectionality feature is relaxed slightly to allow some flow of command from the lower to the higher level; in this case we say, more properly, that there is a fixed asymmetry in the command flow between structures of different levels. Significantly, where A and B are systems at the same level of the hierarchy, there is no conversation between them at all.

A second hierarchical feature is that for any given system the role that it plays is singular and immutable. The corollary of this feature is no less important; namely, that for every function to be computed (or role to be played) there is a specific system in which that function is invested. Collecting these two features, it is evident that a principle characteristic of a hierarchy is the centralization of control, the investment of decision making in one executive system.

A heterarchy is distinguished from a hierarchy in that "free-dominance" defines the relation between any two systems. Partitioning the systems into agents and instruments is arbitrary in that the extant relations depend on context, on the task being performed. The hierarchical unidirectionality of command is replaced by reciprocity in the flow of information. Further, the hierarchical fixedness of roles is replaced in a heterarchy by "functional pluripotentialism" (Filimonov, 1954 cited by Luria, 1966) meaning that no system is responsible for solely a single function, and that any system can assume a (limited) variety of roles as situation and task demand. It follows from this redundancy of function that any inventory of basic constituent elements will be equivocal. Relatedly, management of a heterarchy is not the prerogative of any one system. Many systems would function cooperatively in decision making, although not all systems need participate in all decisions. Collecting these features, it is evident that a principal characteristic of a heterarchy is decentralization of control.

B. Reciprocity of Control Between Transport and Postural Activities

We may distinguish two classes of activity in gross motor tasks: transport and postural (cf. Smith & Smith, 1962). Both classes may be regarded as transformations of posture; that is, configurations of trajectories as the limbs move from

one relatively stable arrangement to the next, although transport transformations are oftentimes more intricate and sometimes more arbitrary than postural. The principal distinction between the two is that transport activities are oriented to the local conditions of stimulation, for example, the flight of the ball, or the motions of an opponent, whereas postural activities are oriented to the global conditions or terrestrial stimulation, the global physical invariants (Shaw & McIntyre, 1974), such as the horizon, gravity, and the ground plane. It goes without saying that most gross motor acts – as manifest in tennis, soccer, etc. – involve a tight confluence between the two classes. Our question is: How is this confluence realized?

1. Constraining the selection of transport activities. Fomin and Shtil'kind (1972) have introduced the term "pedate system" for any system with legs such that the system's normal contact with the surface of support is by means of the plantar parts of the feet. For nature's pedate systems, surface contact through the feet, intertial contact through the vestibular system, and optical contact through the ocular appraratus are the three sources of information about the system's orientation and movement relative to the environment (Lee, 1978). Of the three, vision is the more informative and influential; the vestibular system is not sensitive enough for fine balance control (Lee & Lishman, 1975), and surface contact through the feet is ambiguous about the body's relation to the environment when the feet move relative to the environment, such as when the surface is compliant, unsteady, or narrow (Lee, 1978).

From Gibson (1966) and others (Lee, 1974; Warren, 1976) we have learned that the optical flow patterns at the eye are specific to one's movements with respect to the layout of environmental surfaces. To illustrate, a person attempting to maintain an upright steady stance is perturbed by transformations of the total optic array: a form of inclusive optical expansion induces backward body sway, and a form of inclusive optical contraction induces forward body sway (Lee & Aaronson, 1974). Witness to the human pedate system's sensitivity to this visual source of expropriospecific information is the observation that body sway can be driven phasically by extremely small oscillations in optical expansion and contraction (Lee, 1978).

We can claim, therefore, that while standing or locomoting, the maintenance of an upright posture is an active process (cf. Aggashyan, Gurfinkel', & Fomin, 1973) oriented principally (but not solely) to preserving the *absence of certain kinds of inclusive optical change.* Patently, any transport activity is, in the final analysis, a disturbance of the body's relation to the global invariants which, on the above, is specified primarily by the optical flow pattern. But for a great many transport activities, the activity is possible only if, during the movement, a relatively stable relation is preserved between the body as a unit and the global invariants. Could this be achieved by a simple feedback system, that is, by a process in which the perturbation is corrected subsequent to the activity or, better still, subsequent to phases of the activity? The problem with any feedback solution is that oftentimes the specified compensatory changes are for states that are no longer current. It is evidently the case that while some form of feedback (e.g., velocity or acceleration feedback) is necessary to the integrity of the transport activity postural activity relations, it is not sufficient. Let us consider in this regard the concept of "region of reversibility" as it relates to the concept of pedate system.

The set of all transformations of the biokinematic chains defines a phase space of which a subset is the region of *controllable* transformations. Within the latter there is defined a particular subset such that for any two points in the subset, there is a control process by which either point can be attained from the other; in short, for any movement defined within the subset there is an inverse. This subset is the "region of reversibility," and by the use of the term "equilibrium" for a pedate system, Fomin and Shtil'kind (1972) mean that the kinematic state of the system is within this region. Significantly, the region of reversibility for a particular pedate system is not constant, and among possible sources of variation, we may recognize the conditions of the support surface (compare ice skating to running on the road) and the speed at which the body is moving relative to the surface supporting its locomotion. Now it follows that a major constraint on the planning and executing of many transport activities is that they conserve the pedate system within the region of reversibility. More precisely, and more practically, the constraint is that transport activities do not carry the system *too closely to the boundaries* of the region. Proximity to the boundary is costly in that coordinative effort would have to be disproportionately allocated to postural activities at the expense of transport activities.

By way of summary, it is proposed that preserving a relatively invariant relation to the global invariants in the course of transport activities is partially achieved by an *equilibrium-oriented constraint on the selection of transport activities.* Let us consider a further possible factor.

2. *Anticipatory postural activities are transport-specific.* As alluded to above, preserving balance through feedback alone would often be too late and too slow. This tardiness, however, can be circumvented. When a cat detects an incipient stumble, approximations to the proper muscular response are rapidly generated to preserve the upright posture of the cat long enough for relatively low-level feedback mechanisms to take charge (Roberts, 1967). A particularly sophisticated version of this style of control is suggested by the observations of Belen'kii, Gurfinkel' and Pal'tsev (1967).

On receipt of an auditory signal, a participant is requested to raise his arm rapidly forward to the horizontal position. In the interval prior to the first signs of activity in the deltoid muscles of the shoulder, the muscles most responsible for the movement, there is evidence for considerable modification in the muscle states of the trunk and lower limbs. If it is the right arm that is raised, activity in the biceps femoris of the right leg and the sacrolumbar muscles of the left side precede activity in the deltoids. In addition, a definite anticipatory relaxation

occurs in the left biceps femoris. We see, in short, an orderly pattern of change – of fixing and relaxing links in the kinematic chain of the body – preceding the transport activity of raising the arm. This pattern is both stable and *specific to the transport activity:* The pattern is constant over repetition and the pattern anticipatory to lowering the arm is distinctively different from that anticipatory to raising the arm (Belen'kii et al., 1967; Pal'tsev & El'ner, 1967).

We may interpret these anticipatory changes as intended to minimize the perturbations of the pedate system that would result from the movement of the arm. But insofar as these changes do occur *prior* to the movement and are specific to the moment, we may recognize the larger implication that, at least for this limiting case, the specification of a particular transformation of a kinematic chain, which is a particular transport activity, is concurrently the specification of a particular transformation of other kinematic chains, which is the cognate, postural activity.

If these anticipatory postural adjustments are absent or impaired (owing to brain injury), then pronounced excursions in the center of gravity accompany the arm movement (Pal'tsev & El'ner, 1967). Nevertheless, the anticipatory adjustments are not the whole story, for in the normal case other postural adjustments, presumably of a more precise nature, accompany and follow the movement of the arm (Pal'tsev & El'ner, 1967). It seems as if the anticipatory adjustments put the pedate system into the *ballpark* (see Greene, 1972, in press, b) of postural arrangements appropos the dynamics of moving the arm and appropos the disposition of the limb subsequent to the movement. We may state, therefore, the larger implication, noted above, more simply and somewhat differently: The plan for a transport activity, such as an arm motion, specifies the ballpark of necessary postural activity or, relatedly, a transport plan "pied-pipes"[2] an approximate, postural plan.

The preceding statement, in both its simple and more complicated forms, must be qualified on two counts. First, the relations between transport and postural activity is not a fixed-dominance relation, as pied-piping would seem to imply. The weight of the evidence (Belen'kii et al., 1967; El'ner, 1973; Pal'tsev & El'ner, 1967) favors a free-dominance relation. Second, there is the question of the generality of this form of control. There is the possibility of course, that the balance-oriented fixing and relaxing of biokinematic links, preparatory to and *specific to* a transport activity, is manifest only in simple motor tasks such as studied by Belen'kii et al. (1967). It can be argued, however, on rational grounds, that the form of control described above would be apt for many forms of transport activity; as a general principle, approximating a desired state through feedforward makes the task of feedback regulation considerably more simple and more efficient (cf. Greene, 1972, in press, a).

---

[2] This term was suggested to us by Robert Remez, with all due respect to John Robert Ross.

In the general case, then, the intended transport activity can be the basis for specifying an anticipatory but approximate feedforward adjustment of postural control structures. In the acquisition of a skill (say, a gymnastics routine) it would be beneficial for the performer to become sensitive to the postural activity implications of intended transport activities (see Belen'kii et al., 1967, for a modest demonstration). An advantage of this sensitivity is that by approximating postural controls ahead of time the performer can devote more coordinative effort to the intricacies of the skill.

In sum, we recognize that the confluence between the transport and postural activities is mediated by a reciprocal, free-dominance relation. Activities oriented to local conditions of stimulation are constrained by information about the actor in relation to the global invariants; and, in turn, activities oriented to the global invariants are attuned to the emerging transport activities. On this limited analysis it would appear that the subsystems supporting transport and postural activities are components of a heterarchical organization.

C. Coalitions: Systems with Functional Integrity

We turn now to distinguishing among control systems as they are modeled by the placing or relaxing of certain restrictions on nets of states. In very large part, our intention is to show that a solution to the problem of degrees of freedom must emerge *pari passu* with a solution to the problem of degrees of constraint – a problem that was introduced in the concluding remarks of Section II. When taken together, the overarching problem becomes that of how a (biological) system and its environment mutually constrain one another. The reciprocity of animal and environment is captured by the term ecosystem; a special system that exhibits *functional integrity* (defined below) and that will be said by us to be a coalition. Our strategy, in a nutshell, will be to develop the concept of a coalition by contrasting it with systems whose control principles are both simpler and less abstract – systems such as aggregates, chains, complex-chains (e.g., hierarchies), and heterarchies.[3]

In order to delineate clearly how one type of control system necessarily differs from another, only *pure* cases of each type will be considered. The reader, therefore, should be cautioned against confusing the labels for control systems as used here with the way these labels are often used in the literature. As remarked earlier, pure hierarchies are defined as possessing no mutability among levels of control, that is, no free-dominance relations. Many theorists, however, are primarily interested in hierarchical organizations where the principle of immutable, superordinate control is not sacrosanct, and where inversion of dominance relations is allowed. By out taxonomy, the latter would define a mixed case more aptly described as heterarchical organization of a hierarchy.

---

[3] Elsewhere we have treated the concepts of heterarchy and coalition as synonymous (e.g. Turvey, 1977); here they are distinguished.

It is our belief (although it should be proven) that any system of a mixed-type can be defined as a logical product of a proper subset of pure systems. Hence, the analysis to be given should introduce the set of minimal contrasts required to characterize exhaustively all the subcategories of *pure* control systems.

The above assumption that only pure cases are needed for complete logical characterization of control systems allows the following conjecture to be made (again, a proof would be desirable): *All control systems are properly included as special cases in the category of structures that we shall call coalitions.* In other words, by placing appropriate restrictions on the properties of coalitions, each of the other type of control system can be defined. This inclusion relationship can be represented as follows:

aggregates $\supset$ chains $\supset$ hierarchies $\supset$ chain-complexes $\supset$ heterarchies $\supset$ coalitions

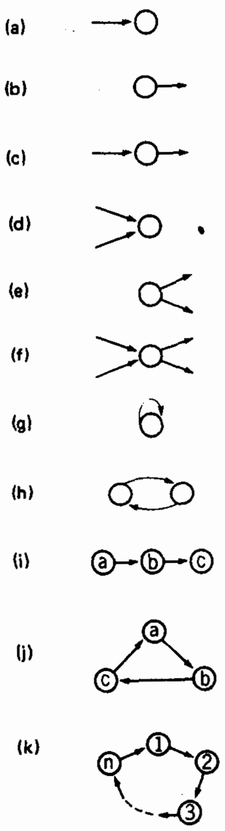
Each structure on the left of the inclusion sign is a special case of that structure on the right under which it is included. There are many different ways to distinguish these structures from one another; for our purposes we will consider only how they differ with respect to their inherent principles of organizational control – what we shall call their dominance of control principles. It will prove to be the case that such an analysis has important implications for *what* such systems can do as well as *how* they might do it. The goal is to show how the concept of a coalition provides a more adequate control system model for living systems than any of the lesser models because it alone is sufficiently well-structured to offer a potential solution to the degrees of freedom problem and the problem of context-conditioned variability.

1. *Nets.* Nets of states provide the medium or structural support for defining control principles. Four primitives comprise a control net: first, a set S of elements called "states"; second, a set R of elements called "dominance relations"; third, a function DOM whose domain is S and whose range is contained in R; and fourth, a function $\overline{\text{DOM}}$ whose domain is R and whose range is contained in S. The first and second primitives are self-explanatory. The third primitive asserts that a function DOM exists such that $\text{DOM}(a, b) = a \text{ DOM } b = a \rightarrow b$ (read as "a dominates b"). The fourth primitive asserts the inverse, namely, $\overline{\text{DOM}}(a, b) = a \ \overline{\text{DOM}} \ b = b \rightarrow a$ (read as "a is dominated by b" or "b dominates a").

All possible control nets may be constructed from the intuitive elements given in Fig. 2.

From inspection of Fig. 2 it follows that nets may or may not possess feedback control loops. One also sees that dominance loops may be defined on a single state, a pair of states, a triplet of states, etc. Hence, the circuit of the dominance relation may be monadic, dyadic, triadic, tetradic or, in general, n-adic, where n is the number of states in the loop.

The properties depicted in Fig. 2 can be used to distinguish several classes of control systems. A strongly connected system is one that has the maximum number of n-adic loops; that is, where there are k states, n (the number of loops)



(a) a dominated state (*input* state)

(b) a dominating state (*output* state)

(c) a transitive state that is both dominated and dominating (*mediating* state)

(d) a state with more than one dominating state (*convergent* state)

(e) a state that dominates more than one state (*divergent* state)

(f) a mixed-dominance state (*convergent-divergent* state)

(g) a state with a feedback loop (*recursive* state)

(h) a feedback loop between two states (*strong* dyad)

(i) a chain with linear transitivity i.e., aDOMb, bDOMc therefore aDOMc

(j) a triadic loop (*weak* heterarchy) exhibiting circular transitivity, i.e., aDOMb, bDOMc therefore aDOMc, bDOMa and cDOMa

(k) an n-adic loop for a system of n states

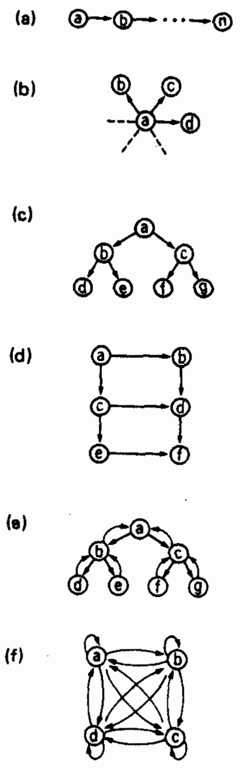
FIGURE 2.

equals *k*. Control systems, therefore, are potentially distinguishable on a connectedness dimension.

It is also possible to distinguish control nets in terms of their "spread" of control, that is, the extent to which every state participates equally in the functions computed by the net. The greater the number of divergent states, or divergent branches from single states, the more centralized the control of the system. Conversely, the greater the number of convergent states, or convergent branches onto single states, the more focused the effect of the control. Moreover, the more strongly connected the net, then the greater the decentralization or spread of control. These two properties, *strength* of connectivity among states, and *spread* of control (or of the effect of control) provide useful dimensions by which the diverse varieties of control nets may be distinguished. Indeed, the

logical product of the fundamental net elements [Fig. 2 (*a*) – (*k*)] yields all structural models needed to support the descriptions of all possible control systems differing along these two dimensions. Some of the types of control systems that have enjoyed popularity among systems theorists are presented in Fig. 3. Let us now survey the range of control systems typically discussed. For the sake of completeness, we start with the most trivial of "systems" – the lowly aggregate.

2. *Aggregates.* Aggregates consist of collections of isolated states that exhibit free-variation. In *free-variation,* each state behaves in a manner unconstrained by the behavior of any other state in the collection. *By control of a given state, we mean a relationship among states of a collection that reduces the degrees of freedom that the given state posses in free-variation.* Hence, the control principle for an aggregate is trivial, since such a collection of *k*-states exhibits a maximum of *k*-degrees of freedom and requires a separate source of constraint



(a) non divergent, non convergent system without loops (a *chain*)

(b) radial net with complete centralization of control (divergent *links*)

(c) divergent net without loops (a pure *output hierarchy*). Reversing the arrows gives a convergent net without loops (a pure *input hierarchy*)

(d) divergent-convergent net without loops (a *chain-complex*)

(e) convergent-divergent net with dyadic loops (a weak *hierarchically* organized heterarchy)

(f) convergent-divergent net with full complement of n-adic loops (a *strong heterarchy*)

FIGURE 3.

for each state, that is, *k*-degrees of constraint. Clearly, a perfectly controlled system is one in which for each degree of freedom exhibited there exists a degree of constraint; or put differently, a perfectly controllable system is one in which for each dominating state, there exists a reciprocal state that dominates it. Such reciprocal states will be said to be *dominance-duals.*

An example of an aggregate is a gas consisting of randomly excited particles. Although such kinetic structures may be stochastically described (the gas laws), they do not consist of states that are jointly controllable. (Only an ideal gas would qualify as a perfect aggregate.) Therefore, let us reserve the concept of a controllable system for those collections of states that can be constrained as a whole in such a way as to exhibit fewer degrees of freedom than permitted under free-variation. Collections of states exhibiting such constraints will be said to be *structured.* Put differently, a *structure* is a system of states possessing fewer degrees of freedom than a collection of free-varying, unconstrained states.

The concept of *control* refers to either a *dynamic* process or a *syntactic* procedure that acts to bring about structuralization of a system. Typically, however, the concept of control is reserved for *transient,* rather than permanent, structuralization of a collection of states achieved by time-invariant rule, rather than time-variant law (Pattee, 1973). For instance, although it is appropriate to say that a population (an aggregate) of molecules is *dynamically* controlled in accord with the gas laws, we shall for the most part be concerned with systems (e.g., biological systems) that are *syntactically* controlled in accord with a program of constraints (e.g., DNA, perceptual information, or remembered rules).

Since, by definition, aggregates are syntactically uncontrollable systems, their study offers little help toward a solution to the problem of how coordinative structures (a syntactic concept!) may become attuned to the environment.

3. *Chains.* The simplest nets possessing nontrivial principles of control are chains. A chain is a concatenation of states that obeys the linear transivity principle of dominance, namely, if *a* DOM *b* and *b* DOM *c,* then *a* DOM *c* (Figs. 2i and 3a). To see why collections of chained states are controllable systems in the above sense, consider a collection of seven states [*a, b, c, d, e, f, g*] partitioned into the following pair of chains: $a \rightarrow b \rightarrow c; d \rightarrow e \rightarrow f \rightarrow g$. A control signal with but two degrees of freedom can provide impetus to activate each chain by stimulating states *a* and *d,* respectively. This represents a savings of five degrees of freedom over the seven degrees of freedom the states would exhibit as a free-varying aggregate of states.

But notice that to control the system, the signal must be specifically addressed to enter each chain at the start-state, the extreme left state. For this reason we call such structures address-specific systems. As weak as such structures are, they nevertheless have the virtue of allowing a partition of several states (a chain) to be controlled as if it possessed but a single degree of freedom (form the perspective of some executive). Traditional examples of such systems are "stimulus-response" chains or reflex-arcs. As argued earlier, the shortcomings of such models for action control systems is that (a) they require a system with total centralization of control (an executor or homunculus) to orchestrate the chains; this means that the degrees of freedom problem is left unresolved; and (b) chains are rigid, undirectional structures, and therefore offer no way in which the action system of an animal or human might become appropriately tuned to environmental exigencies.

4. *The concept of pluripotentiality.* A network of chains can be constructed by concatenating chains by means of divergent and convergent net modules (see Fig. 2). For instance, a planar net of five states comprising two chains can be constructed as depicted in Fig. 4. With the use of divergent or convergent states to build chain-complexes, an important new property is introduced into control systems that is not found in simple chains. Figure 4 represents two dominance functions: DOM(*a, b, c*) and DOM(*a', b, c'*). Notice that these two chain-functions have a state in common, *b.* This means that *b,* unlike the other four states in the complex, plays a *functionally equivocal* role in two distinct functions. Furthermore, as more complex nets are built from divergent and/or convergent states, a greater number of states like *b* will lie at the intersections of chains. The number of such functionally equivocal states in a system will be said to provide a measure of the system's *pluripotentiality.*

Thus, a system with great pluripotentiality is one whose states are likely to be engaged in a large variety of functions – a measure of the richness of the system's multipurposiveness. Thus, unlike aggregates and isolated chains, chain-complexes (networks) are necessarily systems with some degree of pluripotentiality, where certain states assume a degree of universality, or nonspecificity, of function. By generalizing this property, it is reasonable to ask whether there might not be a continuum of controllable systems ranging from those with functionally univocal (unique) states through those with states that have lesser or greater functional equivocality to those remarkable systems (like the human brain?) that appear to have essentially functionally universal states. (A universal Turing machine being such a system.) Indeed, as will become evident, such a continuum of functional plasticity does seem to exist – ranging from simple chain-complexes to coalitions.

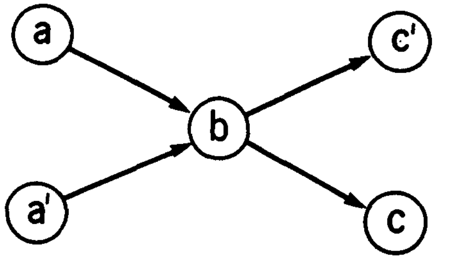


FIGURE 4.

5. ***Hierarchies: Nets without control loops.*** Many interesting properties of control systems, in addition to pluripotentiality, can be modeled by chain-complexes. A very special and extremely popular property is that of being hierarchically organized. As we shall see, such systems allow for the introduction of a degree of generality into control systems not permitted in systems organized at but a single level of control

Control nets may be constructed from two kinds of elements distinguished by what they dominate. A *terminal-state* takes some dynamic aspects of the environment as its value, such as being "on" or "off," conducting or not conducting. Thus, we say a terminal state dominates values in the *execution* mode, i.e. *a*DOM [1, 0]. On the other hand, a *state-variable* is a state that takes another state, or a partition of states, as its value. Thus, we say a state variable dominates values (other states) in the *control* mode, that is, *A* DOM [*B, C,. . .,N*] or *A* DOM [*b, c,. . .n*] (where lowercase letters represent terminal-states). By *levels* of control, we mean an ordering of dominance relations among state-variables, states, and values. Every physically realizable control net by definition must have a minimum of two levels – a control level and an execution level. Hence, a control system necessarily exists simultaneously in two modes: the control (syntactic) mode, and the execution (dynamic) mode. We can now make explicit the intuitive notion of a hierarchically organized control system.

A *hierarchy* is a net with two or more levels of control constructed solely from divergent elements (see Figs. 2 and 3). Notice that a hierarchy can only be composed from divergent (or convergent) state modules (see Fig. 3c). A chain of state-variables constitutes a degenerate hierarchy but can still have levels of control (see Fig. 3a). Significantly, the notion of the levels of control in a net should not be confused with that of *pure* hierarchy. Not all structures organized at multiple levels of control are hierarchies in the strict sense defined above. Consider Fig. 5. It depicts a chain-complex with two levels of control – a primary level that consists of links (*a, d*), (*b, e*), and (*c, f*) and a secondary level that consists of state-variables *A, B, C.* Despite the fact of more than one level of control, the system depicted in Fig. 5 is *not* a hierarchy; it is constructed from components other than divergent (or convergent) state modules.

Recall that the dominance of control principle for chains (which are based on linear transitivity) is too inflexible to account for context-conditioned variability. A linearly transitive structure can not be tuned to the many contingencies of the environment in which it operates. A similar argument can be levied against chain complexes such as hierarchies, in spite of the introduction of higher levels of control. On the other hand, hierarchies or other nets with more abstract levels of organization have the important advantage over chains and chain-complexes with but a single level of control. The advantage is that state-variables may be used to represent collections of nonspatially contiguous substructures (e.g., states, chains, or subcomplexes of chains) that can be treated as if they possess but a single controllable degree of freedom. In this way, a complex collection of biokinematic chains might be optimally organized under relatively simple control principles. Thus, the advantage of hierarchical organization is a certain degree of economy of control.

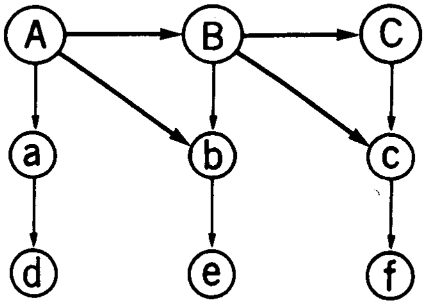


FIGURE 5.

Unfortunately, such economy of control is not gained without a loss of a desirable property. The loss is in the lessening of the "spread" of control by which the functional plasticity of natural systems might be modeled. Because of its divergent organization, a hierarchy necessarily has a high degree of centralization of control centered around a single state-variable – sometimes called a "pontifical" state or "executive" state. Notice also that the spread of control is inversely related to the degree of pluripotentiality in the system: The farther the states are from the topmost level of the executive state-variable, the less functional equivocality enjoyed by the states. In Fig. 3c, *A* has the greatest functional equivocality, *B* and *C* next, with *d, e, f, g* bringing up the rear. This follows from the fact that *A* DOM (*B, C*) while *B* DOM (*d, e*) and *C* DOM (*f, g*); hence *A* DOM (*d, e, f, g*). Therefore, *A* plays a role in four functions (chains), *B* and *C* in two chains each, while *d, e, f, g* are terminal (and in this sense functionally univocal). Thus, we might say that the spread of control (or, inversely, of pluripotentiality) is *anisotropic* in hierarchies.

While systems with such gradients of control may exhibit a certain economy, they suffer from the defect of being too easily infirmed by localized insults. If a superordinate node of control is lost (say, by injury or destruction to neural tissue in a living system), then all subordinate portions of the structure dominated by it will be left syntactically uncontrolled (e.g., paralysis or spasms). This shortcoming is a direct result of the fact that hierarchically organized control nets, no less than chains, are governed by a linear principle of transitivity of dominance. Removal of a mediating state (or state-variable) necessarily disconnects portions of the control net.

To summarize: pure hierarchies therefore provide no means for explaining the functional plasticity observed in natural systems whose functioning often remains intact under a variety of insults. Similarly, because the gradient of the flow of control is ever descending in a pure hierarchy, such control systems are not responsive to environmental fluctuations to which natural systems are observed to adapt. But what type of control principle, and what type of structural organization, is required to model the adaptive, pluripotential, functionally resilient systems observed in nature?

6. ***Heterarchies: Nets with control loops.*** There is but one way in which a system might maintain a stable level of functioning under insults that destroy connectivities among subordinate states or state-variables, and that is to have in reserve redundant or backup structures that may be conscripted in case of an emergency. But clearly this requires that dominating nodes in a control net be able to receive information from below regarding the nature and extent of the insult to the system; for how else might the executive nodes know which reserve states to recruit? However, since this is but a species of the context-conditioned varibility problem, no pure hierarchy without control loops could know, even in principle, which reserve structures to conscript. Indeed, strictly speaking, without feedback from below, the higher nodes would not even know there was an insult to the net.

McCulloch (1945), recognizing the need for a solution to this problem, suggested that natural control systems function as reliably as they do primarily because they seem to violate the principle of linear transitivity of dominance at will – exhibiting what we earlier called "free dominance." For this reason, he argued, if for no other, natural systems must be organized as "heterarchies" rather than as pure chain-complexes, hierarchies, or other networks without control loops.

A *heterarchy* is a net characterized by reciprocity in dominance of control relationships, namely, if *A* DOM *B*, *B* DOM *C*, then not only does *A* DOM *C* hold, but *C* DOM *A* and *B* DOM *A* as well. (Recall Fig. 3e, f). Therefore, heterarchies exhibit control governed by a principle of *circular* dominance which is manifested as loops that follow dyadic, triadic, or *n*-adic circuits through the system.

This control principle of circular transitivity offers a potential solution to the problem of context-conditioned variability. Circular transitivity of dominance relations makes it possible to introduce feedback into the system so that subordinate nodes (terminal states) that experience insult or constraint, due to environmental vicissitudes, can inform the superordinate nodes of the prevailing state of affairs. In this way, the virtue of heterarchies over hierarchies, or other nets limited to unidirectional flows of control, is the capacity to be sensitive to perturbations in contextual constraints. Such "context-sensitivity" is a necessary condition for any system that is capable of adaptive behavior in a changing environment – an indispensable property for all living systems. A minimal heterarchy with this adaptive property can be modelled as shown in Fig. 6.

Let us assume that at some time t_1 *A* DOM *B*, *B* DOM (*C*, *D*) and hence *A* DOM (*C*, *D*) (the solid arrows). In addition, let us also assume the existence of feedback loops *C* DOM *A* and *D* DOM *A*, such that if state *B* is compromised by insult, such that the linear transitivity of control from *A* cannot be received by *C* or *D*, then *C* and *D* can so inform *A* of the quiescence of *B*. In other words, if at some later time t_2 *A* does not receive continuous confirmation that *B* has conveyed the appropriate control signal to *C* and *D*, then *A* will know to disconnect *B* and recruit the backup structure *B'* in its place (Fig. 6 at time t_2).

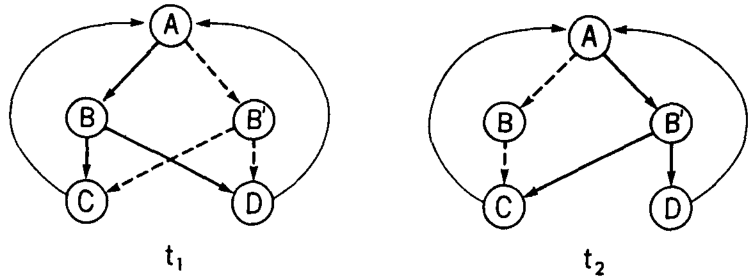
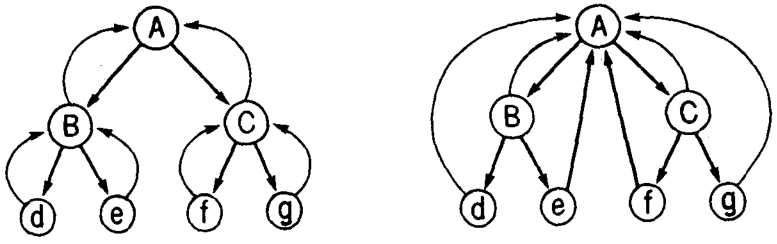


FIGURE 6.

Unfortunately, the above solution to the context-conditioned variability problem is far from optimal, since it increases the anisotropic pluripotentiality of the system. This means that the highest node not only is burdened with the planning and execution of the downward-flow of control, but must now take on the added burden of processing information fed back to it from every subordinate level of control. Recall that the fundamental fallacy of too much central control is that insult to the top-most node renders the system completely inoperable.

For extremely complex hierarchically organized heterarchical systems, where the "node-to-terminal ratio" (number of levels of control) is quite high, an attempt is often made to reduce the dependence of the system on central control by introducing a greater spread of control in the form of an optimization assumption. An example of such an optimization assumption is the *Province of Ignorance Principle* (cf. Turvey, 1977). This principle asserts the condition that no level in a hierarchical-heterarchy is permitted to dominate, nor to be dominated, by any level not immediately adjacent to it. Thus, in a system such as the one depicted above (Fig. 6), it is permitted that *A* DOM (*B*) and *B* DOM (*C*, *D*), but not *A* DOM (*C*, *D*). Similarly, no feedback would be permitted to "leapfrog" over intermediate levels of control; for instance (*C*, *D*) DOM *A* would not be allowed. This means that superordinate nodes can activate, but cannot tune, nodes immediately inferior to it. Moreover, whatever tuning takes place from the environment must be passed upward to the higher nodes following a "domino" principle.

We can now illustrate how the Province of Ignorance Principle can be used to reduce the unequal spread (anisotropy) of pluripotentiality in a hierarchically organized net. This is accomplished by allowing only dyadic feedback loops. This system with optimized control loops can be compared with a system possessing both dyadic and triadic control loops. Consider Fig. 7a depicting a system in which the Province of Ignorance condition holds: Here we see that no state-variable either dominates or received feedback from more than two other states, regardless of how high a level it may occupy. Hence, although node *A* is



(a)

System satisfying Province of Ignorance assumption

(b)

System violating Province of Ignorance assumption

FIGURE 7.

structurally central, it is not really functionally central. Moreover, it is also conceivable that once node *B* or *C* has been activated by *A*, they may continue to function independently of *A*, since their control for tuning emanates from below rather than from above.

By contrast, Fig. 7b depicts a system in which the most superordinate node is both structurally and functionally central. Here subordinate nodes can only be tuned to fluctuations in the lower level context by control from above. Consequently, without information from central control they can not operate at all. Thus, we see how the Province of Ignorance Principle can optimize control by introducing more homogeneous distribution of pluripotentiality. Furthermore, this property has an important byproduct, namely, the creation of relative autonomy among the functional components of the system.

The system represented in Fig. 7, however, is still a mixed case consisting of heterarchical control principles superimposed over a hierarchical structure. Let us consider a pure heterarchy, in which pluripotentiality is perfectly homogeneous, and in which every state or state-variable is functionally central.

Notice in Fig. 3f that each state has the same number of input and output arrows, indicating that no state is functionally more central than any other. Such a pure heterarchy is a strongly connected feedback control system. In principle, such a system is formally capable of computing any well-defined function whatsoever. This follows from the fact that if we allow the heterarchy to consist of an arbitrary number of appropriately selected states, each able to compute a certain primitive function, then the system can be programmed to simulate the computational power of a universal Turing machine – a device for which it has been shown there exists no computational limitations.

Does this mean that heterarchies necessarily provide an ideal model for any natural system? We think not, for natural systems, unlike purely mathematical ones, do much more than just compute functions. Rather as von Neumann (1966) observed, the truly remarkable thing about living systems is that they exhibit actions or goal-directed behaviors – what he called their *effectivities*. Effectivities

are manifested by animals in two fundamental ways: as *appetitive drives*, such as seeking food, mates, and shelter, and as nonconsumatory, or *conative* activities involving planning, such as sorting food, building nests, and shelters before they are actually required, seeking tools, pursuing an education, etc.

Even more remarkable is the fact that animals or other living systems do not express such effectivities in an unsystematic way, but seem to have them integrated into an organized whole. For this reason it is accurate to say that the goal-directed activities of living systems possess a *functional integrity*.

It is precisely this property of functional integrity, that is, of having an organized system of effectivities, that distinguishes pure heterarchies from what we shall call pure coalitions (again we admit the possibility of nonpure or mixed cases). In the last section we explore the concept of control systems that exhibit functional integrity and attempt to demonstrate why such structures are functionally distinct, at least in terms of their dominance of control principles, from heterarchies.

7. *From heterarchies to coalitions.* To motivate the next class of control systems (coalitions), we would do well to pause and consider carefully the fundamental limitations of heterarchically organized control systems. Let us examine a simple control system – a room with a thermostatically controlled heater.

The system has three sources of *internal* control and three sources of *external* control. Let us consider the three internal sources of constraint first: Assume that the thermostat has been set to a criterion of 72°F. As the room cools down below this set-point, the thermocouple in the thermostat expands and closes the circuit, thereby activating the heater. However, when the temperature of the room reaches 72°, the thermocouple in the thermostat contracts and breaks the electrical circuit, thereby deactivating the heater. We can schematize these control functions as follows: *Rm* DOM *th*, *th* DOM *ht*, *ht* DOM *rm* and therefore, *th* DOM *rm* – a clear case of circular transitivity of control.

The beauty of this simple heterarchical control system is that under *normal* circumstances (i.e., those for which the system was designed), it possesses functional integrity: It is capable of achieving a goal-directed function, or effectivity. The effectivity achieved by the "room-thermostat-heater" system is obviously that of keeping the temperature of the room at 72°F.

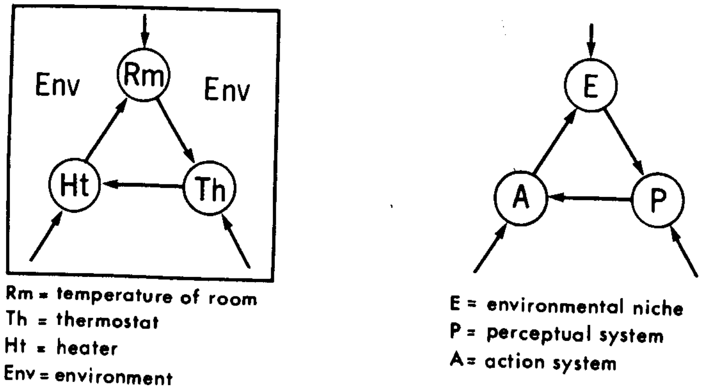
The system also has three potential sources of external control, only one of which it was designed to handle. These potential sources of external control are represented by the arrows on each of the three states from the environment. Inputs to a system from the outside (that is, from states not within the closed feedback system) are called *feedforward* relations. Every *simple* feedback system is designed to control but one dimension of feedforward relations with its environment, although such relations may be indeterminately rich. Our simple system has the effectivity of being able to control only the thermal dimension of its environment. This is represented by the input arrow to the room, construed as a thermal niche in the broader environment.

This feedforward relation to the room indicates that, as a compartment, it is not completely insulated form the broader thermal environment. Indeed, it is this fact that allows perturbations in temperature and that, therefore, makes necessary the effectivity of the control system in question. By contrast to the controllable thermal dimension of variability, the other two arrows represent feedforward relations to dimensions of the broader environmental context that are uncontrollable by the system as it now stands. To control the potentially perturbing effects on the system from nonthermal aspects of the environment, this thermal control system with but one effectivity (i.e., to control heat of room) would have to be expanded to include at least two other effectivities. For instance, the feedforward relation to the heater might represent the way in which the quality or quantity of fuel delivered to the heater might vary below standard, thereby rendering the system incapable of achieving its goal of maintaining the room at 72°F. A control system that is unable to satisfy its effectivity thereby experiences a loss of functional integrity.

Similarly, the functional integrity of the thermal control system might be violated by a feedforward relation from the environment to the thermostat, say by sunlight (a nonthermal dimension) falling upon it. Since the air of the room is relatively transparent, and the thermostat is opaque, light energy falling upon it would be converted to heat, thereby constraining the thermostat to give a spuriously high reading of the actual temperature of the room. Again, such a feedforward relation between the system and its environment would be uncontrollable by the feedback dimension of the system and would mitigate its effectivity, thus compromising the functional integrity of the system.

One is tempted to ask, however, whether simple control systems, such as the one above, might not simply be enlarged so as to subsume the offending feedforward dimensions of the extended environment? Unfortunately, this strategy of subsumption leads to a potentially infinite regress, for any new states brought in to enlarge the system are "Trojan horse" states, in which new uncontrollable feedforward relations are necessarily hidden. Indeed, it is inevitable that the more one tires to close a heterarchical system by subsuming more and more dimensions of variability from its environment under its state-set, then the greater will be the number of feedforward relations smuggled in by this process. Therefore, the technique of creating ever more encompassing heterarchical systems has the unfortunate consequence of increasing, rather than decreasing, the distance between the proposed system and a viable solution to the problem of context-conditioned variability for that system.

The above argument can be generalized to heterarchically organized, "psychological" control systems in the following way: Every control system has a logically irreducible triad of components – a component in which the dimension of variability to be controlled is defined (an *environmental* "niche"); a component capable of measuring, or sensing, values along the dimension of information in question (a *perceptual* system); and a component capable of modulating the information dimension (an *action* system). (Notice the similiarity of Fig. 8b, a "psychological" control system, to the thermal control system discussed above.)



(a)

(b)

FIGURE 8.

The argument regarding the regress of feedforward-control applies *a fortiori* to complex psychological control systems that must retain their functional integrity. In complex environments, the sources of uncontrollable variability are greatly multiplied. Indeed, the above argument can be generalized to show that any naturally instantiated control system will be victimized by uncontrollable feedforward inputs from the environment, unless certain very strong assumptions regarding the "fit" of the system to the extended environment are made.

In other words, to avoid this problem of context-conditioned variability, and to achieve functional integrity, the control system must somehow be initially designed with the *whole* of the environment in mind. It is not adequate merely to design a system to operate adaptively with respect to a narrow environmental niche, unless a guarantee can be given *a priori* that all the variables of the system, taken within the context of the larger environment, are somehow logically *closed*. This is tantamount to demanding that the system must qualify as an *ecosystem*. Consequently, the concept of functional integrity will have to apply to the *whole* ecosystem and not just to the operational component (e.g., the animal or human) alone.

It is our contention that an ecosystem is not merely a very large heterarchy composed of an organism coupled to an environment of potentially perturbing influences. An environment that is sufficient to support a functioning organism must itself be thoroughly organized. If one were to think of the organism as a heterarchy, one could also think of the environment as a heterarchy. This step gives more credit to the environment and acknowledges its order, but is not sufficient to capture the idea of an ecosystem. The interaction of two complicated heterarchies otherwise indifferent to one another would result in an immensely complicated heterarchy. Suppose that one heterarchy had K degrees of freedom and the other had L degrees of freedom. The new system would have KL degrees of freedom, and we would have magnified, not solved, the degrees of freedom problem.

What must be added to the heterarchical property of circular transitivity is a symmetry relation (cf. Shaw, McIntyre, & Mace, 1974), which binds the organism to its environment. If we hold to the image of an organism-environment relation as one of two interacting heterarchies, we would add the requirement that each heterarchy be tailored to the other. They cannot be indifferent to one another; rather, one must complement the other. Recall the example of the airplane guidance system in which sufficient degrees of constraint had to be provided to correspond to the degrees of freedom. This was an example of mutual tailoring of systems. Our candidate for the symmetry relation binding the heterarchies is *dual complementation.* To say that an organism is the dual complement of its environment is also to assert than an environment is the dual complement of its organism, i.e. that the environment is just as thoroughly organized as its organism *and is specific to it* (cf. Gibson, 1977). The relation of dual complementation also carries with it the idea that it is the overarching whole formed by the duals, that is the proper unit. Neither member of a dual pair is properly constrained without the other, or without the whole being defined by their closure. Although defined at a coarser grain than circular transitivity, dual complementation is no less a control principle. An organization that can be characterized as having this style of control is what we call a coalition; and we take a coalition to be the minimum sufficient organization to capture the intuitive notion of an ecosystem. A coalition is not a system-plus-context. It is the minimal system that carries its own context. Finally, a coalition is, for us, the minimal organization that can properly be said to have functional integrity. An organism, perhaps as a heterarchy, is not enough.

In conclusion, we doubt that any approach to the problem of degrees of freedom, or the problem of context-conditioned variability, will be adequate if treated as any style of organization less than a coalition.

ACKNOWLEDGMENTS

The writing of this paper was supported, in part, by Grant HD-01994 from the National Institute of Child Health and Human Development, and Grant RR-5596 from the National Institutes of Health, both awarded to the Haskins Laboratories. The authors wish to acknowledge Carol Fowler, Robert Remez, and James Todd for their contributions to the authors' appreciation of the problems approached by this paper.

REFERENCES

Aggashyan, R. V., Gurfinkel', V. S., & Fomin, S. V. Correlation and spectral analysis of fluctuations of the human body during standing. *Biophysics,* 1973, *18,* 1173–1177.

Aizerman, M. A., & Andreeva, F. A. Simple search mechanisms for control of skeletal muscles. *Automation and Remote Control,* 1968, *29,* 452–463.

Arbib, M. A. *The metaphorical brain: An introduction to cybernetics as artificial intelligence and brain theory.* New York: Wiley, 1972.

Asatryan, D. G., & Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture – 1. Mechanographic analysis of the work on the joint on execution of a postural task. *Biophysics,* 1965, *10,* 925–935.

Belen'kii, V. Ye., Gurfinkel', V. S., & Pal'tsev, Ye. I. Elements of control of voluntary movements. *Biophysics,* 1967, *12,* 154–161.

Bernstein, N. *The corrdination and regulation of movements.* Oxford: Pergamon Press, 1967.

Boylls, C. C. A theory of cerebellar function with applications to locomotion. II. The relation of anterior lobe climbing fiber function to locomotor behavior in the cat. *COINS Technical Report 76-1,* Department of Computer and Information Science, University of Massachusetts, 1975.

Chernov, V. I. Control over single muscles or a pair of muscle antagonists under conditions of precision search. *Automation and Remote Control,* 1968, *29,* 1090–1101.

Easton, T. A. On the normal use of reflexes. *American Scientist,* 1972, *60,* 591–599.

El'ner, A. N. Possibilities of correcting the urgent voluntary movements and the associated postural activity of human muscles. *Biophysics,* 1973, *18,* 966–971.

Engberg, I., & Lundberg, A. An electromyographic analysis of muscular activity in the hindlimb of the cat during unrestrained locomotion. *Acta Physiologica Scandinavia,* 1969, *75,* 614–630.

Eshkol, N., & Wachman, A. *Movement notation.* London: Weidenfeld and Nicholson, 1958.

Evarts, E. V., Bizzi, E., Burke, R. E., DeLong, M., & Thach, W. T. Central control of movement. *Neurosciences Research Program Bulletin,* 1971, *9,* No. 1.

Fel'dman, A. G. Functional tuning of the nervous system with control of movement or maintenance of a steady posture – III. Mechanographic analysis of the execution by man of the simplest motor tasks. *Biophysics,* 1966, *11,* 766–775.

Fomin, S. V., & Shtil'kind, T. I. The concept of equilibrium of systems having legs. *Biophysics,* 1972, *17,* 137–141.

Fowler, C. A., & Turvey, M. T. Skill acquisition: An event approach with special reference to searching for the optimum of a function of several variables. In G. Stelmach (Ed.), *Information processing in motor control and learning.* New York: Academic Press, in press.

Gel'fand, I. M., Gurfinkel', V. S., Tsetlin, M. L., & Shik, M. L. Some problems in the analysis of movements. In I. M. Gel'fand, V. S. Gurfinkel', S. V. Fomin, & M. L. Tsetlin (Eds.), *Models of the structural-functional organization of certain biological systems.* Cambridge, Mass.: MIT Press, 1971.

Gel'fand, I. M., & Tsetlin, M. L. Some methods of control for complex systems. *Russian Mathematical Surveys,* 1962, *17,* 95–116.

Gellhorn, E. The influence of alterations in posture of the limb on cortically induced movements. *Brain,* 1948, *71,* 26–33.

Gibson, J. J. Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology,* 1958, *44,* 182–194.

Gibson, J. J. *The senses considered as perceptual systems.* Boston: Houghton Mifflin, 1966.

Gibson, J. J. The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Granit, R. *The basis of motor control.* New York: Academic Press, 1970.

Greene, P. H. Problems of organization of motor systems. In R. Rosen & F. M. Snell (Eds.), *Progress in theoretical biology* (vol. 2). New York: Academic Press, 1972.

Greene, P. H. Strategies for heterarchical control – an essay. I. A style of controlling complex systems. *International Journal of Man-Machine Studies,* in press. (a)

Greene, P. H. Strategies for heterarchical control – an essay. II. Theoretical exploration of a style of control. *International Journal of Man-Machine Studies,* in press. (b)

Grillner, S. Locomotion in vertebrates: Central mechanisms and reflex interaction. *Physiological Review,* 1975, *55,* 247–304.

Gurfinkel', V. S., & Kots, Ya. M., Pal'tsev, Ye. I., & Fel'dman, A. G. The compensation of respiratory disturbances of the erect•posture of man as an example of the organization of interarticular interaction. In I. M. Gel'fand et al. (Eds.), *Models of the structural-functional organization of certain biological systems.* Cambridge, Mass.: MIT Press, 1971.

Gurfinkel', V. S., & Pal'tsev, Ye. I. Effect of the state of the segmental apparatus of the spinal cord on the execution of a simple motor reaction. *Biophysics,* 1965, *10,* 944–951.

Hubbard, A. W. Homokinetics: Muscular function in human movement. In W. R. Johnson (Ed.), *Science and medicine of exercise and sport.* New York: Harper, 1960.

Hubbard, A. W., & Seng, C. N. Visual movements of batters. *Research Quarterly,* 1954, *25,* 42–57.

Keele, S. W., & Summers, J. J. The structure of motor programs. In G. Stelmach (Ed.), *Motor Control: Issues and trends.* New York: Academic Press, 1976.

Kent, R. D., Carney, P. J., & Severeid, L. R. Velar movement and timing evaluation of a model for binary control. *Journal of Speech and Hearing,* 1974, *17,* 470–488.

Kent, R. D., & Netsell, R. Effects of stress contrasts on certain articulatory parameters. *Phonetica,* 1971, *24,* 23–44.

Kots, Ya. M., Krinskiy, V. I., Naydin, V. L., & Shik, M. L. The control of movements of the joints and kinesthetic afferentation. In I. M. Gel'fand, V. S. Gurfinkel', S. V. Fomin, & M. T. Tsetlin (Eds.), *Models of the structural-functional organization of certain biological systems.* Cambridge, Mass.: MIT Press, 1971.

Kots, Ya. M., & Syrovegin, A. V. Fixed set of variants of interaction of the muscles of two joints used in the execution of simple voluntary movements. *Biophysics,* 1966, *11,* 1212–1219.

Lee, D. N. Visual information during locomotion. In R. B. MacLeod & H. L. Pick, Jr. (Eds.), *Perception: Essays in honor of James J. Gibson.* Ithaca, N. Y.: Cornell University Press, 1974.

Lee, D. N. On the functions of vision. In H. L. Pick, Jr. & E. Saltzman (Eds.), *Modes of perceiving and processing of information.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

Lee, D. N., & Aronson, E. Visual proprioceptive control of standing in human infants. *Perception & Psychophysics,* 1974, *15,* 529–532.

Lee, D. N., & Lishman, J. R. Visual proprioceptive control of stance. *Journal of Human Movement Studies,* 1975, *1,* 87–95.

Lisin, V. V., Frankstein, S. I., & Rechtmann, M. B. The influence of locomotion on flexor reflex of the hind limb in cat and man. *Experimental Neurology,* 1973, *38,* 180–183.

Litvintsev, A. I. Search activity of muscles in the presence of an artificial feedback loop enclosing several muscles simultaneously. *Automation and Remote Control,* 1968, *29,* 464–472.

Luria, A. R. *Higher cortical functions in man.* New York: Basic Books, 1966.

McCulloch, W. S. A heterarchy of values determined by the topology of nervous nets. *Bulletin of Mathematical Biophysics,* 1945, *7,* 89–93.

Megaw, E. D. Response factors and the psychological refractory period. Unpublished thesis, University of Birmingham, England, 1970.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 1956, *63,* 81–97.

Minsky, M., & Papert, S. Artificial Intelligence. *Artificial Intelligence Memo, 252.* Artificial Intelligence Laboratory, MIT, Cambridge, Mass., 1972.

Orlovskii, G. N. The effect of different descending systems on flexor and extensor activity during locomotion. *Brain Research,* 1972, *40,* 359–371.

Orlovskii, G. N., & Shik, M. L. Standard elements of cyclic movement. *Biophysics,* 1965, *10,* 935–944.

Paillard, J. The patterning of skilled movements. In J. Field, H. W. Magoun, & V. E. Hall (Eds.), *Handbook of physiology: Neurophysiology,* (vol. 3). Washington, D. C.: American Physiological Society, 1960.

Pal'tsev, Ye. I., & El'ner, A. M. Preparatory and compensatory period during voluntary movement in patients with involvement of the brain of different localization. *Biophysics,* 1967, *12,* 161–168.

Pattee, H. H. The physical basis and origin of hierarchical control. In H. H. Pattee (Ed.), *Hierarchy theory: The challenge of complex systems.* New York: Braziller, 1973.

Roberts, T. D. M. *Neurophysiology of postural mechanisms.* New York: Plenum Press, 1967.

Schmidt, R. A. The schema as a solution to some persistent problems in motor learning theory. In G. Stelmach (Ed.), *Motor control: Issues and trends.* New York: Academic Press, 1976.

Shaw, R. E. Cognition, simulation and the problem of complexity. *Journal of Structural Learning.* 1971, *2,* 31–44.

Shaw, R. E., McIntyre, M., & Mace, W. The role of symmetry in event perception. In R. B. MacLeod and H. L. Pick, Jr. (Eds.), *Perception: Essays in Honor of James J. Gibson.* Ithaca, N. Y.: Cornell University Press, 1974.

Shaw, R. E., & McIntyre, M. Algoristic foundations to cognitive psychology. In W. Weimer and D. Palermo (Eds.), *Cognition and the symbolic processes.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.

Sivazlian, B. D., & Stanfel, L. E. *Optimization techniques in operations research.* Englewood Cliffs, N. J.: Prentice-Hall, 1975.

Smith, K. U., & Smith, W. H. *Perception and motion.* Philadelphia: W. G. Saunders, 1962.

Stelmach, G. (Ed.). *Motor control: Issues and trends.* New York: Academic Press, 1976.

Sutherland, N. S. Intelligent picture processing. Paper presented at Conference on the Evolution of the Nervous System and Behavior, Florida State University, Tallahassee, 1973.

Turvey, M. T. Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Vince, M. A., & Welford, A. T. Time taken to change the speed of a response. *Nature,* 1967, *213,* 532–533.

von Neumann, J. *Theory of self-reproducing automata.* A. W. Burks (Ed.), Urbana, Ill.: University of Illinois Press, 1966.

Warren, R. The perception of egomotion. *Journal of Experimental Psychology: Human perception and performance,* 1976, *2,* 448–456.

Weiss, P. Self-differentiation of the basic pattern of coordination. *Comparative Psychology Monograph,* 1941, *17,* 21–96.

Wells, K. F. *Kinesiology.* Philadelphia: Saunders, 1961.