MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ARTIFICIAL INTELLIGENCE LABORATORY

Working Paper #198 February 1980

# Suggestions for Genetic A. I.

by Gary L. Drescher

ABSTRACT. This paper presents suggestions for "Genetic A.I.": an attempt to model the genesis of intelligence in human infants, particularly as described by Piaget's theory of the Sensorimotor period. The paper includes a synopsis of Sensorimotor intelligence, followed by preliminary suggestions for a mechanism (the "Schema mechanism") for its development, and a hypothetical Scenario which partially reinterprets Sensorimotor development in terms of that mechanism.

The Schema mechanism focuses on Piaget's concept of the competition and evolution of mental "schemas." The schema is modelled here as an assertion that one partial state of the mechanism's world-representation is transformable to another via a given action, taken when the schema is "activated". A proposed process of "correlation" allows a schema's assertion to be extended or revised in response to empirically-observed effects of the schema's activation. Correlation uses the formation and activation of schemas to propose and test hypotheses, in contrast with the passive tabulation characteristic of associationist mechanisms. Further features are proposed to enable schemas to become coordinated into composite structures, "compound actions", which can be used by other schemas; and to synthesize new "items" (state-elements) when existing ones prove inadequate to model the world.

The Scenario outlines how the Schema mechanism might begin to make its way through the progression of Sensorimotor stages; development culminating in Piaget's third stage is discussed. This development includes learning about the visual and tactile effects of eye and hand motions-- eg, learning how to look directly at an object, or to move a hand into view; and the organization of that knowledge to designate the tactile properties of "visual objects", and vice versa-- eg knowing how to touch an object which is seen-- paving the way to a sensory-modality-invariant representation of objects and space.

The Schema mechanism attempts to "learn from scratch", without built-in expertise or built-in structure in its learning domains. In the past there has been little success among AI programs of this genre. But many such attempts have suffered from mechanisms which were trivial in that they placed the full burden of acquiring and structuring knowledge on one or two simple tricks, whereas, I claim, the present effort shows a willingness to incorporate a multiplicity of elements into a complicated mechanism. In addition, the Schema mechanism benefits from its orientation around a nontrivial *theory* of development. Piaget gives a comprehensive account of the infant's evolution of primitive problem-solving and domain-specific (chiefly object-manipulation) knowledge; this account is used here as a roadmap that describes the proper course for the mechanism to follow. Thus, there is a nontrivial (or at least nonarbitrary) sequence of target abilities to use as a framework for evaluating and revising the mechanism's performance.

## Acknowledgements

I'd like to thank Seymour Papert, Hal Abelson, Danny Hillis, Marvin Minsky, Brian Silverman, Lucia Vaina, Margaret Minsky, Henry Lieberman, Sylvia Weir, Jose Valente, Tom Trobaugh, Puff, Jim Davis, Greg Gargarian, and John Berlow for their helpful advice on the content and/or form of this paper, and for their time and energy in discussing these matters with me. Also, Bill Silver and Spencer Recht worked with me on some ideas for a very early version of this work, and Gerry Sussman offered advice on a somewhat later early version.

I'm especially grateful to Margaret Minsky and Danny Hillis for their encouragement during some of the more insecure days of this research.

I'd also like to thank the many old Yippie chemists and their colleagues for supplying added fuel for my sense of wonder at the workings of the mind. These people perform an important public service despite enormous risk to their own freedom and livelihood.

The research described here was supported in part by my parents, Gladys Drescher and Irving Drescher, who were generously funding most of my tuition and living expenses when I began this work as an undergraduate.

# Suggestions for Genetic A.I

# 0. Introduction

There is an approach to psychology, pioneered dramatically by Jean Piaget, which tries to gain insight into human intelligence by the study of its genesis in individuals. This paper proposes a complementary genetic approach to artificial intelligence-- one based on Piaget's theory of the initial, Sensorimotor period of intellectual development.

In a 3-volume study-- [O.I.], [C.R.], and [P.D.I.]-- Piaget presents a theory of intelligence derived from observations of his 3 children's first few years of life. Interpreting their behavior, Piaget reconstructs the evolution of their underlying representations of reality, and their ways of using that representation to set and achieve goals. Most significantly, for present purposes, Piaget outlines certain "functional invariants"-- I would call them the mechanism-- of intelligence. These describe ways in which, at a given time, existing knowledge/behaviors interact with the environment, and with one another; and ways in which this interaction influences subsequent knowledge/behavior. Piaget claims that these invariants account not only for the development of Sensorimotor intelligence, but can be discerned at the root of the later periods as well.

Piaget makes no attempt to say what's "inside" his functional invariants. He confines himself to a low-level, black-box sketch of *what* they do, without investigating *how.* This is the gap which I propose to address. The intent is to use Piaget's Sensorimotor study as a specification of what his "invariants" ought to do, and attempt to engineer something which does it. I call this approach "Genetic A.I.".

This paper presents a **Schema mechanism** as a first approximation to the functional invariants of Sensorimotor intelligence. Following this is a hypothetical Scenario which outlines the path of development which the mechanism is intended to travel. So far I have only an informal anticipation of what the Schema mechanism will do; this is preliminary to actual implementation and experimentation.

The Schema mechanism focuses on Piaget's concept of the competition and evolution of mental "schemas." The schema is modelled here as an assertion that one partial state of the mechanism's world-representation is transformable to another via a given action, that action to be taken when the schema is "activated." A proposed process of "correlation" allows a schema's assertion to be extended or revised in response to empirically-observed effects of the schema's activation. Correlation uses the formation and activation of schemas as a way to propose and test hypotheses, in contrast with the passive tabulation characteristic of associationist mechanisms. This, I argue, is done in such a way as to avoid the debilitating pitfalls of associationism.

Features are proposed to enable schemas to become coordinated into composite structures, "compound actions," which can be used by other schemas; and to synthesize new "items" (state-elements) when existing ones prove inadequate to model the world. This paper's Scenario sketches the role such features might play in the child's series of reconstructions of the world in progressively more objective, less egocentric terms. Correlation allows certain learning to take place on a given level of reconstruction, while compound actions and synthetic items promote ascension to the next level.

The Scenario outlines how the Schema mechanism might begin to make its way through the progression of Sensorimotor stages; development culminating in Piaget's third stage is discussed. This development includes learning about the visual and tactile effects of eye and hand motions-- eg, learning how to look directly at an object, or to move a hand into view; and the organization of that knowledge to designate the tactile properties of "visual objects", and vice versa-- eg knowing how to touch an object which is seen-- paving the way to a sensory-modality-invariant representation of objects and space.

The Schema mechanism attempts to "learn from scratch", without built-in expertise or built-in structure in its learning domains. In the past there has been little success among AI programs of this genre. But many such attempts have suffered from mechanisms which were trivial in that they placed the full burden of acquiring and structuring knowledge on one or two simple tricks, whereas, I claim, the present effort shows a willingness to incorporate a multiplicity of elements into a complicated mechanism. In addition, the Schema mechanism benefits from its orientation around a nontrivial *theory* of development. Piaget gives a comprehensive account of the infant's evolution of primitive problem-solving and domain-specific (chiefly object-manipulation) knowledge; this account is used here as a roadmap that describes the proper course for the mechanism to follow.[1] Thus, there is a nontrivial (or at least nonarbitrary) sequence of target abilities to use as a framework for evaluating and revising the mechanism's performance. Such a framework is vital to debugging, for early intelligence in humans develops slowly and manifests mistakes and limitations which at first seem bizarre. In the absence of something like the Piagetian roadmap, it would be hard to tell that an artificial mechanism was advancing successfully even if it mimicked human intelligence perfectly, let alone knowing how to correct it when it went astray.

1. It is not presumed here that Piaget's theory has been proven correct, but merely that it's a plausible hypothesis.

This paper includes the following sections:

* a synopsis of Sensorimotor intelligence;
* a sketch of my current, preliminary approximation to the underlying mechanism;
* a reinterpretation of some key Sensorimotor developments in terms of that mechanism. (This can be skimmed first for a quick glance at the proposed theory.)

# 1. Synopsis of Sensorimotor Intelligence

This section summarizes my understanding of the development of Sensorimotor intelligence as described in [O.I.] and [C.R.].

The point of departure of Piaget's theory is

* The Schema: a unit of behavior/knowledge which, by Piaget's biological metaphor, interacts and evloves with its physical environment, and its fellow schemas. The initial schemas are merely those of reflex responses. For quite some time, the infant's schemas are closely associated with the her actions. Later sophistications, involving the combination of schemas, abstraction above specific acts and perspectives, and the "interiorization" of schemas' activity, will allow the schema to transcend a literal dependence on physical action, while retaining its procedural flavor. Schemas of looking, grasping what's seen, swinging, dropping, hiding one object under another, pushing one object with another, are examples of post-reflex schemas.

Piaget identifies as the functional invariants of intelligence **assimilation** and **accomodation**: respectively

* a schema's use of things in the world (including other schemas) as part of its own funcitoning; and
* the modification of schemas in adjustment to novelties in the world.

Of course, Piaget doesn't try to present complete, explicit rules governing the activity and modification of schemas. But his theory does try to characterize such rules and to give an intricate chronicle of the low-level results of their functioning.

The Sensorimotor period (from birth until about age two) is the first of three broad periods of development in Piagetian theory. Sensorimotor intelligence is expressed solely in *actions* which affect the world. In the later phases-- of Concrete Operations and then Formal Operations-- the *truth* of assertions about the world becomes the focus of intelligence, first for assertions about the *real* world, and later in the realms of the hypothetical and the abstract. [P.C.]

Piaget distinguishes six stages within the Sensorimotor period. Each successive stage is characterized by schemas of a new elaboration of "problem-solving"[1] or "goal-pursuing" activity (which never implies the eradication of less sophisticated schemas, or even that such schemas stop being created). The elaborations characteristic of a given stage do not appear simultaneously; the "stage" is just the period during which such appearances first peak. A stage's uniformity is thus a descriptive invention, and doesn't imply rigid chronological partitioning.

1. The infant's earliest behavior is only a 0th-order exhibit of "problem-solving"; later stages do greater justice to the term.

The infant's representation of reality-- space, objects, causation, time-- exhibits corresponding stages of development. In fact, Piaget argues-- and this is among his most profound insights-- that progressively more sophisticated techniques of intelligence, and progressively more sophisticated representations of reality, are but two indissociable aspects of the same development-- much as we are now accustomed to viewing intellect and affect as but views of a single process. At the outset of intelligence, problem-solving is just the dynamic expression of the infant's representation of reality-- a natural enough idea, since the infant's schemas are procedural: a thing is understood in terms of what can be done to/with it. So, more advanced promblem solving results from the application of the same mechanism to more sophisticated representations of reality, and vice versa. Eventually, of course, the child acquires explicit knowledge *about* thinking which can be used to improve methods of thought; but substantial maturing of intelligence occurs even before such meta-knowledge is evident in the child.

One critical feature of the infant's intelligence, not well captured by this summary, is the *incremental* quality of its development. At least at the outset of intelligence, each new capability observed in the infant is only slightly different than what was previously exhibited; the infant shows only minor adjustments of activity, in apparent response to experience in prior activity. It should be kept in mind that the actual steps are of much finer grain than are presented here. As intelligence progresses and there come to be more powerful schemas for interpreting the world, the steps grow bolder, and, in ways that I'll discuss, less dependent upon specific experience. So, the change from trivial to powerful steps is a smooth one; the increments by which intelligence improves are, in effect, of size proportional to the power of existing schemas, so the development is of an exponential character.

## 1.1 First Stage

### Reflex Activity, Solipsist Images[1]

The infant's initial schemas are those of reflex activity: eg closing the hand in response to a touch on the palm, or sucking something which touches the lips. These schemas are exercised either in response to the appropriate stimuli, or else spontaneously, as if for "play" or "practice".

1. I'm using a slightly different border between first and second stage than Piaget defines. This is of no importance; I just mention it in case anyone notices and is confused.

Schemas from the outset admit of modifications in response to exprienced results of their activity. For example, after many instances of disorderly reflexive groping for a nipple touching the mouth, an infant's sucking schema appears to notice that when the nipple touches (say) the left cheek, turning to the left will be propitious. Groping in adjustment to the nipple thus assumes a gradually more coherent appearance, as clues such as cheek-contact are exploited.

The early development of schemas also shows generalization and differentiation. For example, the sucking schema adjusts itself not only to the nipple, but also to other objects frequently presented to it: eg a finger or a toy. Often, the infant will suck such an object as contentedly as if it were a nipple. But when hungry, the infant responds with enthusiasm to the nipple while crying instead if given a finger to suck. The appearance of this discrimination suggests that, despite the production-like character of schemas' early, stimulus-triggered activity, the "desired" *result* of a schema's activity also affects its course.

The first few months of life also see the first so-called **primary circular reactions**. These are patterns of action, derived by gradual differentiation of reflex schemas, which tend towards repetition. For example, the grasp-reflex schema gives rise to a alternately-hold-then-release-object schema, and to a scratch-object schema, etc. As with pure reflex schemas, these sometimes repeat "emptily", that is without any stimulus/object to interact with.

Visual schemas developing at this time include those of tracking a slowly moving object, of visually exploring a stationary object, and of alternate glances between one object and another.

A striking feature of these early schemas is that they haven't yet "intertwined". For example, tactile stimuli elicit no visual response; things seen inspire no effort at prehension. Moreover, when for example a watched object passes beyond the infant's field of view, the infant either loses all evident interest in it, as though it no longer existed; or else, with apparent expectation of seeing it again, either continues to look off in the same direction, or gazes back to where the object was first seen. Similarly, an object which is touched but not seen may be repeatedly grasped then released; but if, say, it falls to a new position, the infant will neither search for it visually, nor move her hand to search for the object in a different position than where just grasped.

These observations imply that the infant's model of the world-- in the sense of what aspects of the world the infant can react to/exploit-- is (metaphorically) solipsist in nature: the infant's universe contains not objects of substance and permanence viewable from different perspectives, but rather "images", some visual, some tactile, etc., which change state in response to personal actions (themselves "known" only by the transformations they produce). The infant's early schemas organize the world into various solipsist spaces, each giving a group (in the mathematical sense) of operations: the operations are primitive motor actions (or, sometimes, passive expectation), and the things operated on are sensory states.

Note the lower bound on the level of abstraction on which Piaget studies intelligence. While he observes the organization of schemas into networks of transformations, he makes no atempt to say, for example, what low-level visual feature detection is constituent of the "states" in an infant's visual space. In later stages, *higher* levels of abstraction are considered-- referring, for example, to the position of an object, independently of the particular sensory perspective from which it's viewed-- as the infant's schemas come to represent things on higher levels of abstraction. The claim is: whatever level of abstraction an infant's representation uses at a certain stage, the same *function* of that representation determines what the infant understands and does, how the representation is adjusted and extended, and (as we'll first encounter in the next stage) how the next higher level of abstraction is formed.

## 1.2 Second Stage

### The Coordination of Primary Schemas

As reflex schemas elaborate into primary circular reactions, they also begin to intercoordinate and thus to bridge the gap between sensory modalities. The primary circular reactions, and the intercoordinations, both appear to have the same character of development: a schema acquires differentiated responses to, and anticipations of, sensory signals with which it was previously unacquainted. If the new signals of one schema are already familiar to another, then a functional intercoordination results, as when schemas of hand movements combine with sucking to form in integrated thumb-sucking schema.

Initially, an infant will suck her finger (or other object) only if it comes in fortuitous contact with the infant's mouth (or, slightly later, cheeks etc.). (Even then, the infant doesn't know how to keep her hand in place, and the hand is quickly pulled away.) But random hand movements may accidently brush the hand against the vicinity of the mouth. Not only will this trigger attempts to suck, but also, future hand trajectories will converge to the mouth more and more directly. Eventually, the infant can smoothly and spontaneously move her hand to her mouth, and insert and suck on a finger. Later, a more profound development is seen: the infant is capable of carrying a grasped object to her mouth and sucking on it; thus, prehension is coordinated with sucking.

More striking still is the coordination which develops between vision and prehension. Piaget discerns a number of milestones in this development:

* The infant watches the movements of her hand, and gradually learns to bring her hand into her visual field, and keep it there while watching it.
* The infant watches while grasping and releasing objects.
* The infant subsequently will turn to look at an object when the object touches her hand, or will move the object into her visual field to look at it.
* At some point, the infant will reach for an object if the object and the infant's hand are seen together.
* Eventually, the sight of the object alone will suffice to trigger a successful attempt to grasp it.

Of course, each of these bits and pieces of eye/hand coordination develops not as a sudden leap, but by gradually improved groping.

The acquisition of visual/tactile coordination has an important consequence: hereafter, the infant's learning and attention become oriented around "objects", not just particular sensory impressions. The appearance of this more objective behavior marks the onset of the next Sensorimotor stage.

## 1.3 Third Stage

## Secondary Circular Reactions, Objects of Subjective Permanence

Secondary circular reactions, characteristic of third stage behavior, consist of the repetition of actions in order to reproduce fortuitously-discovered effects on objects. For example:

* The infant's hand hits a hanging toy; the infant sees it bob about, then repeats the gesture several times, later applying it to other objects as well, developing a "striking" schema.
* A strange sound is made by accidentally striking the crib wicker with a toy. The infant reproduces the motion involved, and after more occasional fortuitous contacts, will rub the toy deliberately against the wicker. However, spatial contact between the objects is not understood as such. If the infant's position is changed such that the customary gesture fails to achieve contact with the crib, she repeats the gesture anyway, doing nothing that adapts to the altered situation.
* The infant pulls a string hanging from the bassinet hood, and notices that a toy, also connected to the hood, shakes in response. The infant again grasps and pulls the string, already watching the toy rather than the string. Again, the spatial and causal nature of the connection between the objects is not understood; the infant will generalize the gesture to inappropriate situations.

In these reactions, the infant responds quickly to a novel result by using a familiar schema to reproduce the result, even though the schema had never previously been used for that purpose. However, the effect is discovered by accident, and only the particular schema involved in the accident is used to reproduce the effect.

Nonetheless, thanks to the intersensorial schemas of the previous stage, the current schemas transcend particular primitive motor actions and sensory images. This, together with the more complex chain of actions involved in, say, seeing/grasping/moving/rubbing an object, give secondary circular the reactions the appearance of being goal-directed (where the goal is to reproduce the surprise effect), in contrast with the stimulus-bound appearance of the primary circular reactions.

The sense in which the third stage initiates the representation of objects rather than images is perhaps best described as follows: if one were to write a program that did the sorts of things that a third stage infant does, the program would most naturally be written on a level of abstraction that designated objects; a program to mimic earlier stages would most naturally lack such a level, and would instead be oriented around sensory images.

To the extent that they deal with objects rather than images, the secondary circular reactions can designate primitive *interrelationships* between objects-- but with the limitation that the relationship is given only by a particular schema of action, implying both unnecessary restrictions, and inappropriate generalizations, of the relation.

Similar progress, and limitations, appear in the third stage representation of objects' permanence and position:

* **Deferred circular reactions** appear. An infant, playing with a toy (via a secondary circular reaction schema) is momentarily distracted but soon turns back to where the toy was left and resumes playing with it. This is similar to, but more complicated than, the earlier feat of looking again at one image after shifting gaze to another; here, a coordination of body and hand movements, guided by vision, is required to recapture the object.
* When the infant is watching an object that falls, moving too quickly to track so that she loses sight of it, she will look downwards for it. At first this happens primarily when it was the infant who held and dropped the object, and is also catalyzed by the sound of the fallen object, or by tracking it momentarily when it starts to fall. Eventually, the reaction becomes reliable even in the absence of such clues.
* Similarly, if the infant holds (without looking at it) an object which falls, or is taken, from her hand, she learns at this stage to extend her hand and reclaim the object.

Thus, the third stage infant apparently conceives of objects as occupying particular positions at which they can be reclaimed if they vanish from view. Moreover, in contrast with the previous stage, the object can be sought in a *new* position, rather than the first or last place that it was recently perceived. However, closer observation shows that this reclamation is only understood with respect to a particular schema of action. The infant confronted with an object's sudden disappearance tries to recapture it either by extending the activity of a schema already invoked to keep sight of the thing-- eg for the falling object-- or by reusing a schema just used to secure the thing in the first place-- eg reaching to regrasp an unseen object removed from the hand. In this latter case, if that particular gesture fails to rediscover the object, the infant will *not* (until the next, fourth, stage) employ perpendicular motions in a systematic search for the thing, but may instead revert to looking for it in its original position.[1]

That the position of vanished objects is first conceived only in terms of particular action schemas is further attested to by the reaction of an infant to the intervention of an obstacle. If an infant of this stage is presented with a toy which, as she watches, is covered with a cloth, the infant will not attempt to raise the cloth to recapture the object-- despite the fact that the infant is quite capable of picking up a cloth when that itself is of interest. When the toy disappears, the infant either loses interest, stares at where it was, or looks back at where it was first seen (if that was a different place), but does not reach for it-- or, if already reaching for it when sight of it is blocked, will immediately give up!

In fact, even if the infant's attempt to grasp a toy is thwarted by a barrier which doesn't block sight of the toy, the infant appears to be oblivious to the barrier, making no attempt to displace it or move around it. The infant does, however, learn during this stage to grasp and extricate the hidden toy if *part* of it is visible.

The need to *rotate* an object presents intellectual difficulties similar to those posed by the need to move an obstacle. Suppose a third stage infant is presented with a bottle, but the bottle is held with the nipple facing away from the child, so that the nipple cannot be seen. Thus the important part of the bottle is obscured, not by a foreign object but by the rest of the bottle itself. The infant exhibits problems similar to those produced by a separate obstacle, giving up on the nipple when it is no longer perceived. The difficulty is not a lack of the motor skill required to rotate an object, since while the nipple is visible, the infant will turn the bottle to make the nipple accessible; this is done quite unsystematically, but persistantly until fortuitous success is achieved. So the difficulty is again a representational one, characteristic of this stage: the "potential nipple" (as opposed to the nipple when actually perceived) is understood only in connection with certain schemas known to actualize it. There is not yet a schema of rotation; the successes in orienting a visible nipple appear to be due to a series of separate movements, each guided crudely by the current perception of the nipple, and not organized into a coherent activity of reorientation. When, in the next stage, these attempts

1. This reversion to cruder techniques when more advanced ones fail tends to occur through all stages of Sensorimotor intelligence, and later intelligence as well.

are arranged in a coordinated structure, there will indeed be a schema of rotation, with respect to which the potential nipple can be represented.

Finally, it should be noted that during the third stage, a "potential-X-with-respect-to-prehension" is not strongly coordinated with a "potential-X-with-respect-to-vision". For example, an infant of this stage who has looked at, but not touched, an object which falls below her gaze, may look downward for it, but will not make any tactile search for it.

## 1.4 Fourth Stage

## Coordination of Secondary Schemas

The fourth stage brings a coordination of secondary schemas analogous to the second stages's intertwining of primary schemas. Just as the second stage allowed the infant's representation of the world to transcend specific primitive motor sequences and sensory impressions, and abstract these to *acts* upon *objects* (the subject of third stage learning), so the fourth stage coordinations will allow the infant's understanding to become independent of particular acts, preparing for fifth stage elaboration of the activity of objects themselves, and their interrelationships.

The fourth stage infant is capable of using a familiar schema for a new purpose in a new situation. This contrasts with the previous stage, whose secondary circular reactions did allow familiar schemas to be used for new effects, but only if these effects had previously been empirically (and fortuitously) produced.

A classic example of this is the removal of an object blocking the prehension of a desired toy. This may be catalyzed by the accidental displacement of the intervening object when the infant initially ignores it. But at some point, the infant's attention is focused specifically on moving the obstacle (at first clumsily, but successive efforts develop a well coordinated schema of **displacement** by picking up and moving, or by striking). The infant's behavior makes clear that she is not interested in the obstacle itself, since it is discarded and the desired toy is then grasped. The obstacle displacement was thus subordinated to that goal. (Interestingly, it isn't until shortly after this displacement coordination that Piaget observes the advent of the infant's ability to release one toy being held in order to pick up another.)

An important variation of the above displacement coordination is the removal of an object which blocks the view of a desired toy. In transition between the third and fourth stages, an infant might continue to reach for and grasp a toy whose view was blocked, provided that the infant had already started to reach when the object disappeared from sight. This, along with the extrication of partially hidden objects (from the previous stage), and the displacement of non-hiding obstacles, leads to the ability to react to the complete covering of an object by removing the cover and claiming the rediscovered object. This is quickly generalized into a game of repeatedly hiding and recovering an object.

Recall the third stage inability to, say, respond with prehension to a "potential visual object". During the fourth stage, "potential" (in contrast with actually perceived) objects with respect to different schemas are united in a way reminiscent of the second stage's marriage of visual and tactile perceptions. The ability to uncover a hidden object extends this unity: not only is there a prehensile remedy to a visual disappearance, but the remedy is complicated, involving a pair of secondary schemas which deal with two distinct objects. Thus, both the permanence and spatial localization of vanished objects are now understood, not just with respect to a given secondary schema, but with respect to coordinated pairs of such schemas. This begins to put objects in spatial relationship to one another. Similarly, the infant of this stage becomes capable of:

* systematic search. Eg, when the infant drops an object, her hand will not only be moved down to find it, but will also be moved perpendicularly in exploration of the immediate vicinity.

* systematic rotation. The infant can recover the obscured reverse side of an object.

* exploitation of perspective. The infant can shift her head to look around an obstacle.

* imitation of familiar but invisible movements. During the third stage, only visible actions, producible by existing schemas, are imitated; eg grasping a toy. (Interestingly, there is no imitation of a sequence, such as opening and closing a hand, which is exercised as a part of various familiar schemas, but not yet differentiated in its own right.) In the fourth stage, the infant will imitate an action (such as sticking out the tongue) which she has taken many times, but without having seen its effects. (Prior visual/tactile exploration of faces, in conjunction with sounds sometimes accompanying the gesture, provide clues that assist that identification.)

* systematic exploration of novelty. When presented with a new object, the infant applies in succession many familiar schemas to the object: shaking, striking, rotating, etc. During the third stage, a new object would tend to excite some schema or other, but the current emphasis is different: the schemas now seem focused on the object, while previously, understanding of the object seemed focused on a particular schema. (An unexpected effect of some exploratory action-- say, the production of an unusual sound-- may give rise to a secondary circular reaction repeating that effect. Piaget calls such a reaction **derived** to denote that it arose in the context of more structured activity, namely the exploration.)

Despite these advances, the fourth stage representations of reality still exhibit many limitations of subjecivity. The most striking of these is shown by the following experiment. The infant plays with a toy which is taken away and hidden under a pillow at the left. The infant raises the pillow and reclaims the object. Once again, the toy is taken and hidden, this time under a blanket at the right. The infant promptly raises, not the blanket, but the pillow again, and appears surprised and puzzled not to find the toy.

This sort of confusion is observed repeatedly during the fourth stage. It's a remarkable analog to the earlier reaction to disapperance by searching in the first or last place that the thing was recently perceived, or in a new position by extending a reclaiming schema. Then, hidden position was represented only with respect to the comparatively simple schemas that existed. Now, hidded position is understood in terms of combinations of such schemas, which relate pairs of objects. Although more complex, the representation is still procedural, and the procedures involved have only developed to the point of saying something like: "when this toy disappears, displacement of the pillow will rediscover it."

So the relationships among objects are yet understood only in terms of pairwise transitions, as in the cycle of hiding and uncovering a toy. The intervention of a third object is not properly taken into account. Moreover, the infant still comprehends the displacement of an object relative to herself rather that to another object. For instance, an infant who can easily turn a block around does not yet learn to orient it relative to a box so as to fit inside. Similarly, there is no comprehension of the need to put a stick in contact with a semi-distant toy in order to move the toy. These feats will be possible in the following stage.

## 1.5 Fifth Stage

### Experiments on Objects

During the fifth stage appear so-called **tertiary circular reactions**. These are little "experiments" which the infant conducts to see what an object will do. For example, an infant may repeatedly drop a toy, paying evident attention not to the act of dropping, but to the behavior of the *object* as it falls. Similarly, the infant experiments with varying ways of placing an object on an inclined surface to watch it roll, or perching it at the edge of a table so that it tumbles to the ground, etc.

These experiments extend the focus on an object's behavior, rather than personal action, noted during the last stage. But where fourth stage explorations merely use the object in existing schemas, the present experiments vary the exploratory schemas-- not just in *response* to surprise results (as with the derived secondary reactions noted in the last section) but in *provocation* of unexpected behavior. (Indeed, the specific autonomous activity of an object is yet unexpected by the infant, as evidenced by systematic inability to account for it when necessary. For example, an infant trying to dispose of an obtrusive cushion repeatedly pushes it back against a wall, but in such a position that it must fall back in the way again.)

Tertiary (like secondary) circular reactions can be coordinated with other schemas in a means-end relationship. For instance, an infant reaches through the bars of a playpen to grasp a long toy. The infant doesn't anticipate the solidity of the bars, which block the toy from being drawn closer. (The fourth stage infant learned about the soldity of an obstacle to prehension, but that was only with respect to movement of the hand itself! Here, the infant must learn that one object also blocks the motion of another object.) Although the infant already knows how to rotate an object (say to find its reverse side), there is not yet a schema for rotating one object relative to another, as is called for here so the toy can be oriented to allow passage through the bars. But, lacking such a schema, the infant nonetheless appears to identify the collision as the source of difficulty, and for a long while gropes for different ways of placing the object against the bars. Eventually, a successful orientation is found. On subsequent attempts, the infant's gropings converge more and more quickly to the solution, and a reliable schema of object-relative rotation is evolved.

The gropings of this example are tertiary circular reactions, as they involve deliberate *variations* of a repeated action, and with interest in the effect on the object (ie whether it is making progress through the bars), rather than in the action itself. Now there is an additional feature: the experiment is directed toward the *goal* of bringing the toy closer. Thus, many schemas influence the activity:

* the grasping schema which specifies the goal.
* the schema of turning an object, relative to one's self, which gives a point of departure for the new means needed to fulfill the goal.
* importantly, the many schemas which by now exist to describe objects and space; these are needed to interpret meaningfully the results of the experimental variations, to direct refinements of the evolving rotation schema.
* the intermediate approximations to the objective schema which so evolve.

From the observer's point of view, the coordination of these schemas results in an important amplification of the infant's intellectual capabilities: for the first time, the infant responds to an unexpected obstacle by "inventing" a way to overcome it, rather than just relying on an already-existing schema. Piaget concludes that this capability essentially falls out of

* quantitatively, the myriad schemas which can be brought to bear on a situation; and
* qualitatively, the higher level of abstraction on which the schemas now represent things, focusing on objects as such;

thus allowing the same principles of interaction of schemas to yield more sophisticated results.

Similar examples of the invention of new means are found when the infant learns to use a stick, an underlying support, or an attatched string, to move a given object. You may recall that some secondary circular reactions involved influencing one object by pulling another connected to the first by a string. But that effect was discovered entirely by accident, and with no appreciation of the physical connection. During the present stage, the infant wishing to influence a remote object learns to search for an attatched string, visually tracing the path of connection. As with the objective rotation schema, a great deal of intermediate groping is required to develop schemas for using a string, support, or stick. One interesting intermediate situation that Piaget observes regarding the use of a stick is that an infant who is trying to grasp an object just out of reach, and who has previously succeeded in using a stick to draw the object closer, will not think of doing that unless she is already holding the stick, or unless the stick is presented to her. This is somewhat like the state of a second stage infant who is learning to grasp what is seen, but only when the hand is *seen* next to the object.

These developments add to the infant's conceptions of objects and space. Through the tertiary circular reactions, objects are endowed with autonomous behavior; and the direction of such reactions towards goals involving a second object teaches the infant about the solditiy of objects, and relationships among objects themselves. This progress is also reflected in the multiple-screens problem of the fourth stage, described above. During that stage, some improvement is made in selecting the right place to look for a vanished object, but the accomplishment has an empirical character and the selection is often wrong, as though the infant had learned that "looking under the *blanket* sometimes works instead" but without really getting the point. On the other hand, the fifth stage infant learns reliably to search the place at which the object was seen to disappear.

## 1.6 Sixth Stage

### Simulation of Events

The fifth stage infant shows no sign of mentally "simulating" the activity of objects and learning from the simulation instead of from actual experimentation. But the sixth stage furnishes evidence of this ability. An infant who reaches the sixth stage without happening to have learned about (say) using a stick may invent that behavior (in response to a problem that requires it) quite suddenly, with dramatically less groping than similar inventions of the previous stage. Piaget argues that some "interiorization" of physical activity is responsible for this capability.

In addition, the infant now becomes capable of interpreting situations whose understanding requires representation of events not actually observed. For instance, consider yet another form of hidden object confusion, which the fifth stage infant exhibits: A toy is placed in a small box, without a lid, so that the infant still sees it. Before the infant has a chance to recover the toy from the box, the box is moved beneath a blanket where, hidden from the infant's view, toy is dumped out. The box is brought to view again, empty. The infant is surprised that the toy is no longer in the box, and does not attempt to search under the blanket. Analogously to fourth stage progress in the multiple-screen problem, the fifth stage infant does learn, empirically and unreliably, to search under the blanket. But when *two* screening objects are used in succession, a remarkably parallel confusion results: the infant does not understand the need to look specifically under that cover from which the box emerged. But now, during the sixth stage, the infant deals successfully with these situations, apparently able to represent the unobserved displacement of the toy under the screen.

The above developments are a miniscule sample of the explosion of intellect and knowledge which occurs during the sixth stage. The ability to represent one's own body in objective spatial terms, to understand personal orientation (as in being able to point back to a house that's no longer in sight), and the beginning of language all arise during this stage. The sixth stage thus forms a bridge between Sensorimotor intelligence and the subsequent periods.

## 2. The Schema Mechanism

NOTE: It is suggested that this section only be skimmed on a first reading. Then, as you read the next (Scenario) section, you can refer back to here to fill in required details. After reading the Scenario, you may wish to to reread this section in full. (Sections marked with an asterisk can be skipped entirely on a first reading.)

This section discusses ideas for what I'll call the **Schema mechanism,** intended to climb the ladder of Sensorimotor stages. By no means do I contend that I've understood how to achieve this. I claim only some preliminary ideas in that direction, which this section presents, more as an example of how such things might be thought about than as a theory expected to be close to working.

My intent is to regard Piaget's functional invariants of Sensorimotor intelligence as an approximate specification of how the mechanism should behave. Any proposed feature of the Schema mechanism should meet the following two criteria:

* empirical criterion. The feature should be roughly consistent with humans' capabilities and limitations according to Piagetian theory. A proposed feature is considered gratuitous if it would "preprogram" some ability which, by Piagetian theory, humans must learn instead.
* design criterion. The feature should make sense from an engineering standpoint. It should be possible to implement and should do things which there is reason to believe would be useful to intelligent behavior.
The design criterion includes attention to the *scale* of the problem: a feature is unacceptable which may work in trivial examples but would suffer an explosion of space or time requirements if faced with a situation with a realistically large number of elements.

The methodology here is to avoid incorporating just any clever trick that suggests itself, to focus instead on conforming to the Piagetian specification. I take this approach, not because I consider Piaget's theory certainly true, but because it seems like a powerful and plausible hypothesis. To try to implement it is a way of exploring the hypothesis. But conversely, the specification should not be followed blindly; it's only plausible that Piaget is *approximately* right. It certainly would not do to attempt to mimic some aspect of an infant's behavior without reason to believe that the mimicing feature actually accomplishes something. The above empirical and design criteria are intended to balance these concerns.

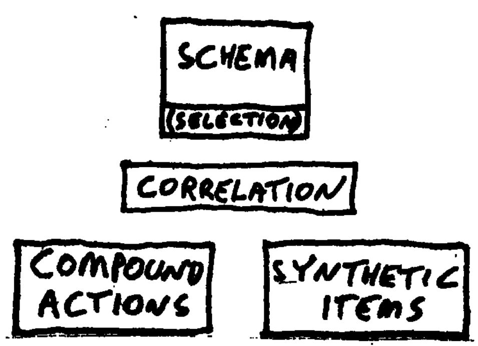



**Figure 2-1**

The above diagram suggests an organization of the central facets of the Schema mechanism.

* **The schema** is a unit of representation which asserts that if a particular action is taken when the mechanism's world-description is in a specified partial state, some new partial state will result; schemas state the "rules" by which the universe is believed to work. Schemas compete for selection to be **activated**; when a schema is activated, its action is taken. The competition and activation of schemas are the basis of the Schema mechanism's operation.

* **Correlation** is a process whereby spinoff schemas are formed to extend or revise a schema's assertion in response to the observed effects of the schema's activation.

The schemas created by correlation are constrained to be comprised only of already-existing actions and **items** (state-elements). The actions initially supplied to the Schema mechanism are primitive motor actions; the first items correspond to primitive sensory impressions. Arguably, the newborn infant conceives of things only in such terms; but the infant is able to build from experience more complicated actions and world-designators. An attempt is made to endow the Schema mechanism with such a capacity:

* **Compound actions** arise by the coordination of existing schemas into composite structures which can be used as actions by other schemas. A compound action can be much more complicated, and much more flexible, than the actions of its component schemas. Moreover, a compound action raises the level of abstraction on which things can be represented, in that a schema which uses a compound action is unconcerned with the details of the action's component schemas.

* **Synthetic items** are formed when existing items prove inadequate to express explanations of observed events.

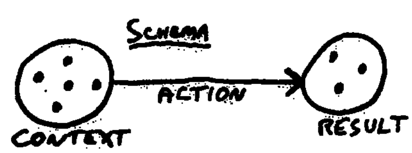
Then, section 3 (Sensorimotor Scenario) sketches the role such features might play in the child's series of reconstructions of the world in progressively more objective, less egocentric terms. A central theme of the Scenario is that correlation allows certain learning to take place on a given level of reconstruction, while compound actions and synthetic items promote ascension to the next level.

## 2.1 The Schema and its Selection

### 2.1.1 General Description

Let's start with a unit of representation which, after Piaget, I call the schema. It looks like this:

**Figure 2.1-1**




Its principal parts are the

* context packet: a collection of **items** (see below) which must all be **on** for the schema to be

applicable.

* result packet: items expected to be turned on (or off: **negative items**) if the context items are all on and the specified action is taken.

* action: elaborated below.

An item makes some assertion about the world; the item can be **on** or **off**. The first items are primitive sensory inputs (for example, something's-touching-my-right-index-finger-tip) and the first actions are simple motor actions (eg move-hand-forward). The first schemas correspond to simple reflex activity, with the "stimulus" in the context, and an empty result. It will be possible to form higher-level items (corresponding to objects and in general to arbitrary things/states) and actions (in general arbitrary transitions), to be discussed in the following sections.

In addition to the above, the schema includes a reliability index. This is discussed in connection with **correlation** in section 2.2.

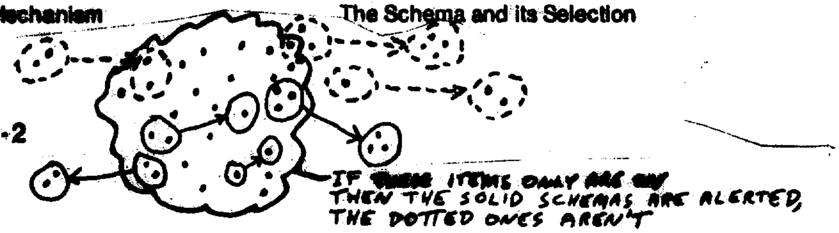
The basic idea of schema selection is that at every "clock tick" (let's use discrete time) the schemas respectively decide in parallel how important they are to activate. The importances can be compared in parallel (in time proportionate to the log of the number of values) and the winner **activated**. For now, "activated" just means effecting the schema's action if it's a motor primitive; but more interesting things to happen with activation will appear in following sections.

The philosophy behind the importance criteria I'm about to enumerate is this: the Schema mechanism has two mutually recursive toplevel purposes. One is to expand the frontiers of its knowledge by activating schemas which are new or which will lead to novel situations; this will cause, in ways to be discussed in subsequent sections, the creation and debugging of new schemas along those frontiers. The second purpose is to use existing, reliable schemas as means to achieving specified goals-- often, goals which will arise in the pursuit of the first purpose. Expressed more loosely, the mechanism tries both to create problems and to solve them, each effort nurtured by the other. The importance criteria presented throughout this paper reflect these two basic concerns.

Some importance criteria are **intrinsic** to a schema:

* A critical question is whether the schema is **alerted**, ie whether all the items in its context are on. This is requisite for most other importance criteria to be engaged.

Figure 2.1-2

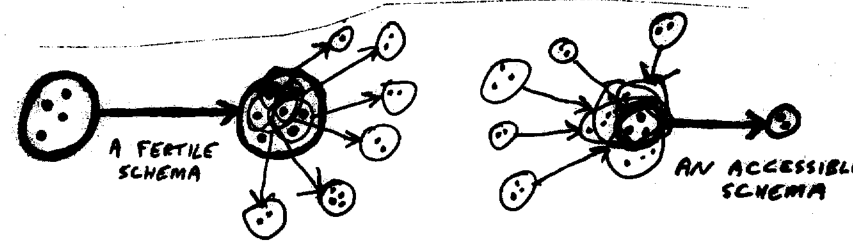



* A newly-formed or long-unused schema assumes extra importance, to encourage exploration of the unfamiliar. Some **novelty index** might exist (per schema) to reflect this.

* A schema is of increased importance to the extent that it "leads to" a complicated network of other schemas. This could be achieved via a **fertility index** per schema which, upon each activation of a schema, gets set to the number of schemas alerted in the following clock tick (or to some average of that and the index's previous value, or something). The number could be weighted by the importances of the alerted schemas.

* Conversely, a schema is important to the extent that it is "led to" by many other schemas: the development of such a schema contributes to the others' fertility. An **accessibility index** can measure this.

Figure 2.1-3




* It might be useful to have a **hysteresis effect** that maintains attention by temporarily increasing the importance of any schema which gets activated.

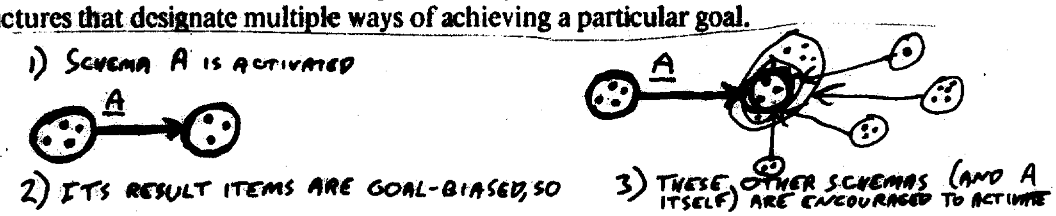
Other importance criteria depend upon the desirability of a schema's items:

* certain sensory items can be inherently "pleasurable" or "displeasurable", affecting the value of schemas containing them in their result packets.

* more important is a **goal biasing** feature-- essentially the above feature but with items of variable desirability: whenever a schema is activated, its result items are goal-biased. Then, other schemas whose result packets include those items assume greater value. This can facilitate:

  * the various circular reactions
  * the tendency towards imitation
  * best of all, it allows a given schema in effect to specify a "goal" for other schemas to pursue; if the given schema's result conditions don't actually result, other schemas predicting that result will assume greater value.
  * in section 2.3 it is shown how goal-biasing can contribute to the formation of structures that designate multiple ways of achieving a particular goal.

Figure 2.1-4




These goal-directed criteria affect a schema's importance proportionately to the reliabilty of the schema.

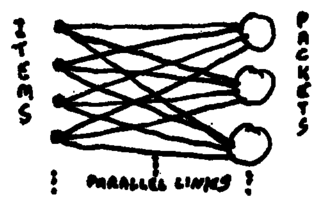
A schema can also ask to be selected for some purpose other than activation. For example, it might be stipulated that a schema which isn't alerted but which has a goal-biased result can compete for selection; if it's chosen, it causes its context items to be goal-biased, in the hope that some other schema can respond and cause the first schema to be alerted. Still other kinds of selection are presented in subsequent sections. Each schema decides the purpose for which it competes for selection at any given time; when the selection is made, the appropriate event (be it acitvation or whatever) takes place.

Additional criteria of schema selection will be made up as I go along.

### 2.1.2 Implementation *

A few words about the implementation of schemas. I'm assuming that all packets of items (ie, so far, the context and result packet of each schema) have parallel access to the state of all items:

Figure 2.1-5




(This assumption is, perhaps, extravagent, and it might be worthwhile eventually to try to get away with less. But it will do for now.) Each access line can be enabled or not, to denote inclusion or not of the associated item in the associated packet. It's assumed that the packet can compute in parallel a few very simple functions of its items, such as conjunction, disjunction, or count of items which are on; or, if we later have reason to give numerical values to items, summation or identifying the maximum value.

The computation of alertedness, and of goal-pursuit, is then straightforward. A schema can tell whether it's alerted by checking the conjunction of its context packet items. Goal-biasing of a schema's result is accomplished by sending a message to the result items (saying "you're a goal"), which can in turn be read by other schemas whose results contain the biased items. If we want to insist that it's important to contain *all* of the biased items, we can make available the current number of such items (ie place this number in some register to which all schemas have parallel access) so that each schema can check to see if its own result includes that number of goal-biased items.

When another schema comes along and goal-biases *its* items, the previously-biased items should lose their goal status. But their effect can persist, if we stipulate that schemas which were found to point to them retain their boosted importance for a while.

A similar method allows the computation of the accessibility index: bias the schema's context items so that other schemas can decide if their result packets include that context. The number of such schemas can then be tallied in parallel.

For the sake of simplicity, I'm assuming, somewhat arbitrarily, that the computations performed by various schemas communicating through the packets-to-items links is the only heavy parallelism which occurs in the Schema mechanism. In particular, any construction or modification of a schema requires the attention of some central resource, and can't be done by several schemas locally and autonomously-- rather, it can only come about through the serial event of schema selection.

As a point of reference, it might be wildly guessed that a "real" Schema-like mechanism would require millions of schemas and items, and that thousands of each would suffice to obtain some preliminary results.

## 2.2 Correlation

This section presents a way to modify schemas in response to empirical observations, to achieve a rudimentary form of learning. Two methods are presented, one for extending the prediction made by a schema, the other for restricting it. The extension mechanism watches for unexpected events which occur reliably when a given schema completes its activity; it is hypothesized that such events result from the schema's activity, and a new schema is created which reflects the hypothesis. The restriction mechanism is called into play when a schema is activated but its predicted result does not obtain. Then, a new schema is created with a more conservative prediction; and, an attempt is made to identify stricter context conditions under which the bolder prediction is accurate.

### 2.2.1 Extension

I need to define a **surprise**: a surprise is the set of items whose states have just changed unexpectedly. What's expected is that an item will maintain its state, unless the schema(s) whose action(s) just terminated predicted that that item would turn on (or off); then, that prediction is the overriding expectation.

Typically, when a schema is activated, many surprises will occur. Some of these may in reality be related to the schema's action. But the vast majority are likely to be coincidental. Nonetheless, all the surprise items are put in a **balloon** that's attatched to the schema's result packet. (Negative items are placed in the balloon to designate items which were *off* unexpectedly, rather than on; the state of a negative item is always the opposite of the corresponding positive item.) Balloon items don't count as part of the result packet for any purpose. But they are candidates for inclusion in the result packet.

On subsequent activations of the schema, those balloon items which don't match the actual state of the world are purged from the balloon.[1] (No new ones are added.) Eventually, either the balloon empties-- in which case it can be refilled to try again-- or else, its contents will be **confirmed** by an activation which doesn't produce any further purging. When that happens, a new **spinoff** schema is created, which copies the old schema but with the surviving balloon items included in the result.

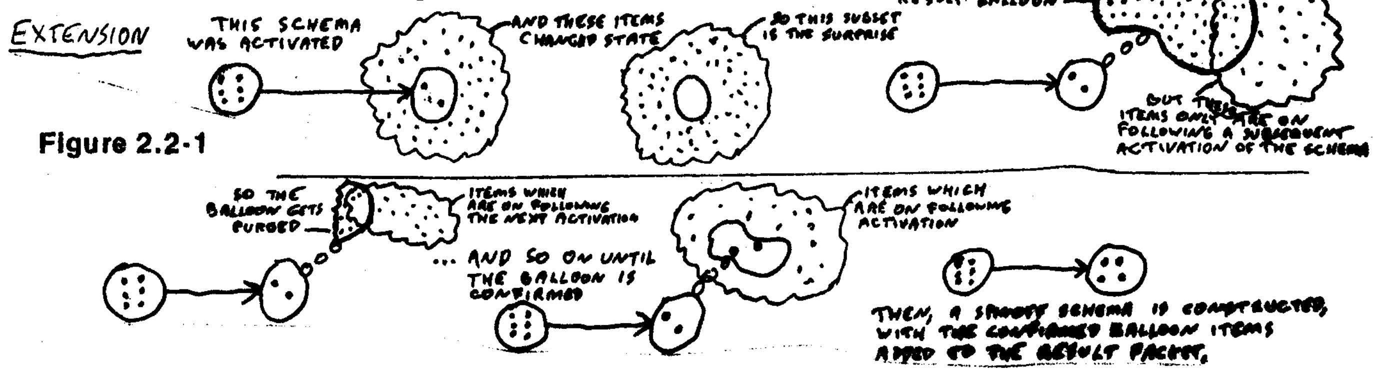



**Figure 2.2-1**

Regardless of how many spurious surprise items are initially stuck in the balloon, the above process ought to find quickly any "correct" ones, for the convergence is exponential: assume that unexpected item transitions unrelated in reality to a schema's activity have at most probability p of happening coincidentally when the schema is activated. Then, p or less of the remaining spurious items are expected to escape being purged on a given activation. So the number of extraneous items goes down as $p^n$, where n is the number of purging activations.

## 2.2.2 Restriction

Suppose a schema is activated unsuccessfully. That is, just after its action is completed, not all of its result items are on. The correlation feature tries to identify additional conditions for the schema's successful operation. Two things happen:

* A spinoff schema is created, copying the old one except for those result items which didn't behave as predicted.
* A balloon is attatched to the (original) schema's context. This is filled with all the items which were off when the schema was activated, and the negative items of all items which were on at that time. These are candidates for inclusion in the context packet, on the speculation that the absence of some of them was responsible for the schema's failure.

---

1. Only surprise items are initially placed in a schema's result balloon since, if a given event was expected anyway, there's no reason to attribute it to the schema. But an item escapes purging from a balloon if the item is on, even if not by surprise; for in that case, although the item's appropriateness to the result balloon wasn't really demonstrated, neither was it contradicted.

Following the *successful* activation of a schema, its context balloon is purged of any items which were off when the activation occurred. Such items are not necessary to the schema's successful activation, so they need not become part of the context packet. (Of course, some *disjunction* of such items may be necessary, but the correlation process will not discover them.) Like a result balloon, the context balloon is confirmed when a (successful) activation occurs which results in no further purging. Its contents are then added to the context packet of a spinoff schema which otherwise copies the original. If the balloon empties before being confirmed, it is refilled on the next unsuccessful activation and the process starts over. Context ballooons, like result balloons, don't actually count for anything until they're confirmed: in particular, a schema is considered to be alerted without regard for whether the context balloon items are satisfied.

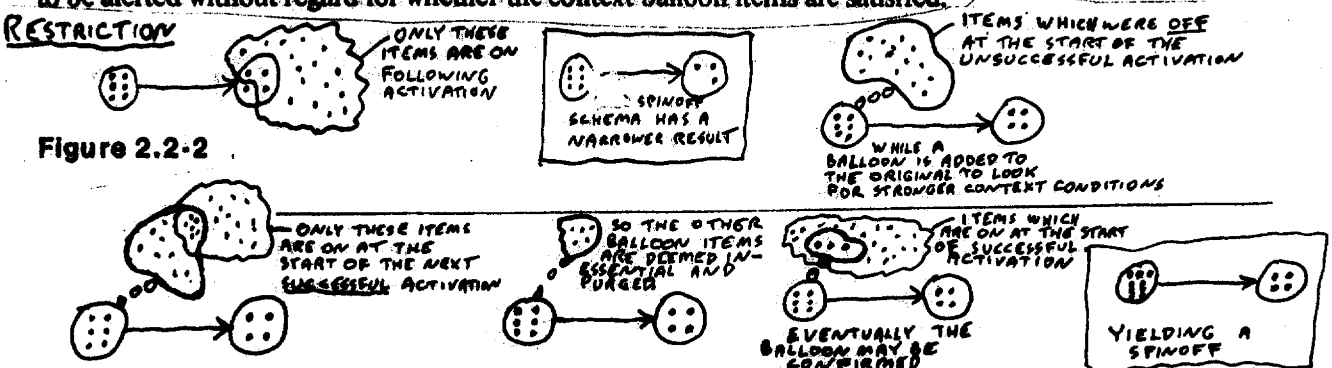


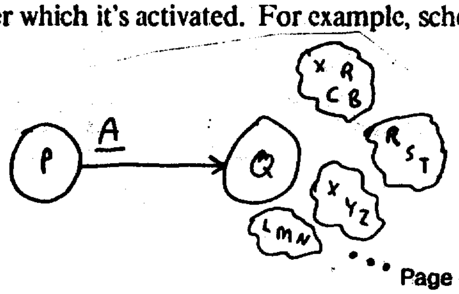
Figure 2.2-2

A schema also keeps track of its *reliability*, which ranges from 0 to 1. A successful activation upgrades a schema's reliability index, say by $x_n := (99x_{n-1}+1)/100$, while an unsuccessful activation degrades it to 99/100 of its old value. So the reliability index is altered only slightly by a given activation. But if a schema has a consistent level of reliability, the index will eventually converge to it.

There is a qualification on the creation of a spinoff schema: first, it's determined if there already exists a schema of the same identity (ie the same context items, action, and result items) as the proposed spinoff. If so, no new schema is formed, but the existing one's reliability is upgraded.

### 2.2.3 Investigation *

An unpredicted effect of a schema's activation will not come to be attributed to that schema unless the effect is reproduced in sufficiently many consecutive trials to confirm the effect. Often, this constraint poses an undue hardship: a schema might reliably have one of several different effects, depending on the conditions under which it's activated. For example, schema A might have any of these effects

Figure 2.2-3


depending on these items at the time of its activation

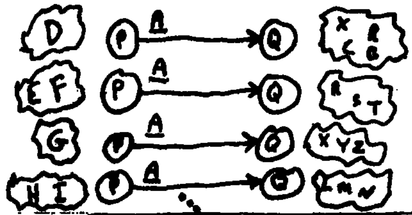


Figure 2.2-4

What we would like to see, of course, is the encoding of this state of affairs into the following schemas

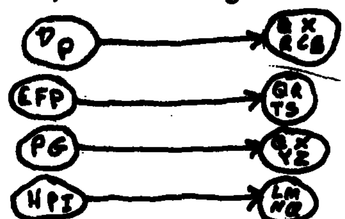


Figure 2.2-5

But there is a problem. The various possible effects would be unlikely to be confirmed in a result balloon, since each result combination is reliable only with respect to certain yet-unguessed constraints: schema A's context items alone do not insure any particular result combination. Conversely, there is nothing to foster the inclusion of any of the additional conditions in a context packet until *after* the corresponding result has been included in the result packet of some version of schema A. So neither the context nor the result packet can develop (much) in the desired direction until after the other has done so!

Fortunately, this impasse can be broken, at least in certain interesting cases. To say how, I first need the notion of a **recognized surprise**: this is a subset of a surprise, consisting of those surprise items which are included in the context of the next schema to be activated-- that is, those items which are recognized by being given a role in what happens next.

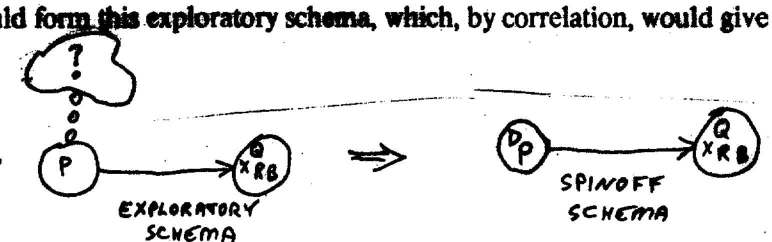
The proposed **investigation** process works like this: following the activation of any schema, an **exploratory** spinoff of that schema is created which replicates the original, but also contains in its result packet all the items comprising the recognized surprise which followed the original's activation. Now, the fact that some surprise items are recognized doesn't make it obviously likely that their state-change was actually caused by the previous activation; on the contrary, the relationship is most probably one of coincidence. But the exploratory schema is initialized with lowest-possible reliability, and with balloon attatched to its context as though the schema was already known to fail. The point is not ever to rely on the exploratory schema as it stands, but rather to use it to search for additional context conditions which might justify the new, rashly-appended result items. This search is passive: the exploratory schema is never activated in its own right. Instead, whenever a new exploratory schema is to be created, it is first determined whether the new schema matches an already-existing one. If so, no additional one is made, but correlation is done on the old exploratory schema to purge its context balloon. Soon the balloon converges, most likely to emptiness, but possibly to context conditions which make the exploratory schema well-founded and reliable.

Thus, in the example above, the desired schemas can form if their results tend to be used in the contexts of other schemas. For instance, if schemas A and B are activated in succession as shown here

Figure 2.2-6


then there would form this exploratory schema, which, by correlation, would give rise to the spinoff shown:

Figure 2.2-7




This resembles the first schema in figure 2.2-5, except that only the *recognized* part of the schema's effect is included in the result packet (C is omitted). The other schemas in figure 2.2-5 (or approximations thereto) could be of similar origin.

I should state clearly the motivation for the choice of the recognized surprise as the consequence to investigate. Any arbitrary subset of a surprise might be no less likely to have been caused by the previous activation, but there's a reason to prefer to go exploring for the conditions that led to a *recognized* surprise: simply, if this exploration is successful, its fruits are especially valuable, since the discovered results facilitate the activation of other desirable schemas. This coordination between schemas is particularly important in light of the mechanism presented in section 2.3 for the spawning of composite structures from coordinated components.

### 2.2.4 Discussion

The kind of empirical "learning" which the correlation feature might achieve is clearly rudimentary. The world-view implicit in the correlation process assumes that the items in a given packet relate to the schema "linearly", in that the correctness of the inclusion of each is considered independently of other items. (In fact, the purging of packets and balloons is like the tuning of a perceptron in which all the coefficients must be either 0 or 1 [Perceptrons].) Moreover, only already-formed items can be taken into account, and these are initially minimal. Clearly, very little of the universe is understandable as a linear combination of sensory impressions. Yet the correlation feature may contribute usefully to intelligence. The very earliest generalizations and differentiations of Sensorimotor schemas suggest that something like correlation may play a principal role. And the cumulative gropings of later stages may be achieved by correlation occuring within the framework of more complicated mechanisms, and in the presence of more advanced items and actions.

In short, correlation is (hopefully) a feasible approach to what "associationism" and its relatives attempted unfeasibly: acquiring knowledge from the empiricism of what happens in connection with what. Correlation has substantial empirical and design advantages over associationism:

* Piaget (among others) argues emphatically that infants show no sign of making arbitrary associations. Rather, such associations as are evidenced always seem to occur as incremental extensions or refinements of existing schemas. Correlation has this character.

* From an engineering standpoint, associationism is an evident disaster. Associationism tends to propose connections among any things that happen together in succession: if event Y follows event X, the prediction X->Y is proposed. Yet a great many sensory (and other) events occur together, most unrelated to one another. The number of ways that an arbitrary subset of one moment's events could be responsible for an arbitrary subset of the next moment's is absurdly large; a uselessly small fraction of such combinations are meaningful. In contrast, the correlation process should converge quickly to a meaningful relation, since either the antecedent or consequent conditions are "fixed" with respect to a search for the other. Moreover, expression of the "association" in a schema makes the association both *usable* (insofar as the schema is procedural and goal-oriented) and *testable/correctable* (through activation and correlation).

## 2.3 Compound Actions

Schemas expressed in terms of primitive motor actions may provide an adequate point of departure for Sensorimotor intelligence. But more complicated, more abstract actions are surely needed by all but the most rudimentary intellectual structures. This section addresses that need by the introduction of **compound actions.**

Before I present the details of compound actions, it is worthwhile to show the embryonic forms of two basic ideas from which those details arose. The first: since this is above all a *genetic* theory, there must be a reasonable way for compound actions to *develop* under the mechanism's own power. One straightforward way to do this is: when schemas A and B are activated in succession, A having assured the alerting of B by the inclusion of B's context items in A's result packet, a new action can be formed which consists of A's action followed by B's. A new schema can be created, with A's context, B's result, and the new, "abbreviated" action:

Figure 2.3-1


The result-context-inclusion qualification is to try to ensure that it will make sense to follow action X with action Y and expect the result predicted by schema B.

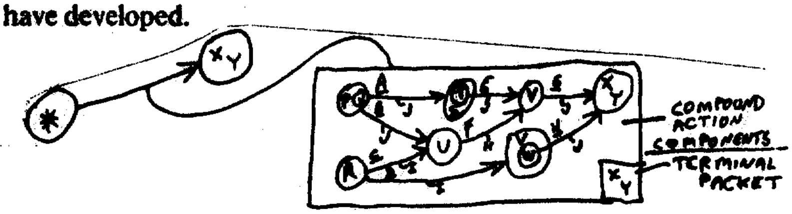
There are many problems with this idea, and it is especially deficient with respect to the second basic concern: a compound action ought to have advantages similar to those of a subroutine in computer programming. The abbreviated representation of a fixed sequence of steps is but a small part of what subroutines are for. More importantly, they allow matters to be represented on an elevated level of abstraction: the subroutine *internally* may decide to take different courses of action on the basis of many details, but externally it should perform a coherent function which, for purposes of building larger structures, can be thought of without (much) regard to those details.

The compound action feature is an attempt to synthesize these two concerns. Compound actions are to be formed in a manner which preserves the local, incremental flavor of the above abbreviation device, but in such a way as to allow for greater flexibility in the new action: in particular, to allow the compound action to embrace many potential paths to a common goal, to steer itself through such paths conditionally upon various details.

## 2.3.1 Execution

Let's start by examining a schema with an already-formed compound action; then, we can back-track and see how it might have developed.

Figure 2.3-2

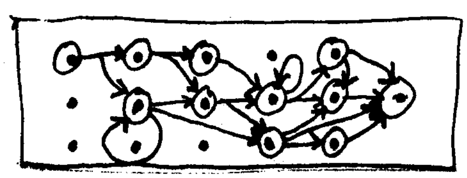



The schema at the left of this diagram has a compound action, which is comprised of the **component schemas** pictured in the rectangle. (For contrast with the component schemas, I'll refer to the schema using the compound action as the **user schema**.) The compound action is said to be **enabled** when any of its components is alerted. The asterisk in the user schema's context denotes the **action-enable item** of that schema's action-- an item which is on whenever the action is enabled. The overlapping of components' result/context packets denotes their being linked as predecessor/successor in the compound action, the significance of which is explained below.

Should the user schema be activated, its compound action is executed as follows: First, whichever component schema was responsible for enabling the action is activated. (If there are many such components, some sort of arbitration singles one out.) This happens as an automatic consequence of the user schema's activation; that is, the designated component need not compete independently for activation, as a schema ordinarily must do. Furthermore, once a component schema is activated as part of the compound aciton, it sends a **go-ahead** signal to the component(s) which it's linked to (for example, schema E above tells G to go ahead). Receipt of the go-ahead signal has the effect that the receiving schema activates automatically, without competing for selection, *provided* that that schema is alerted, and provided that the user schema is still active (see below). When two (or n) schemas have common contexts and actions-- for example, A and B above-- each, if activated, gives the go-ahead signal to *both* (all) of their successors in the compound aciton-- here, schemas E and F. Typically, the successors' contexts will be mutually exclusive possible results of the previous components' action. Passing the go-ahead to all the possible successors lets control flow conditionally to whichever successor finds itself alerted. But if several such components are alerted, some arbitration method should insure that only one is activated via a given go-ahead signal. The user schema becomes inactive after one of its action's **terminal schemas** (here, G or H) has been activated.

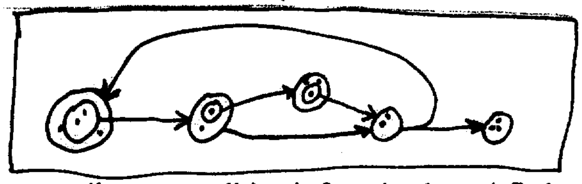
The paths of control in a compound action always converge to a common **terminal packet**, a set of items included in the result packets of each terminal schema (above, the terminal packet comprises items X and Y). The terminal packet is effectively the goal around which components are organized. As action components are activated, the flow of control can be conditional upon the result of the previous activations' results, always tending to converge to the terminal packet. This all happens, as it were, in the background: the schema selector has been free, since activating the user schema, to concentrate on other matters, unconcerned with the underlying decisions in the comound action's execution. The Scenario section (section 3) will give examples of compound actions where the divergence of activation paths (due to multiple possible results of components' activation) and reconvergence (by conditional activation of successor components) create a feedback system in which small errors of position and timing are continually corrected in order to maintain a desired path, as suggested by this figure:

**Figure 2.3-3**



Also possible is the expression of a *search* in a compound action, for example:

Figure 2.3-4



A loop of components repeats until some condition is found to be satisfied.

### 2.3.2 Annexation

Now we can consider how to augment a compound action with new components. One way to do this is not very far removed from the abbreviation method at the beginning of this section. When schema B has a compound acton, we say that A is **implicitly linked** to schema B when A's result packet includes: all of A's context items (execpt that the action-enable item may be omitted), and all the context items of *any* of B's action's component schemas. Thus, when A and B are implicitly linked and A is activated, it is assured that if A's predicted result is correct, B will be alerted. So now the **annexation** procedure: when implicitly linked schemas A and B are activated in succession, A becomes a component of B's action. It becomes the (a) predecessor of the component(s) whose context is included in its (A's) result. For example, if this schema

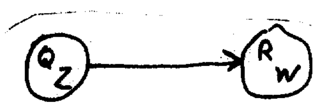


Figure 2.3-5

and the user schema from figure 2.3-2 were activated in succession, the compound action would acquire a new component:

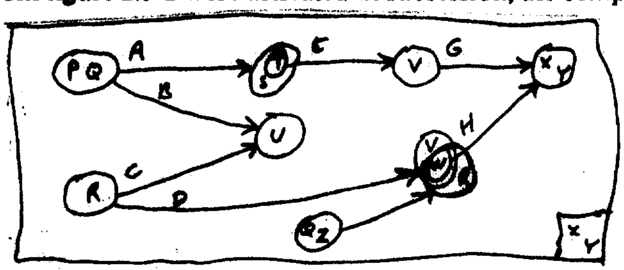


Figure 2.3-6

(Note that schema H now has *two* predecessors.) In a situation which was like the above except that schema B had a *primitive* action, B could be treated for purposes of annexation as though it had a compound action whose sole component was B itself; B's result would become the compound action's terminal packet.

If A and B are activated in succession, but are *not* implicitly linked, **extended annexation** can still take place. In this, a spinoff of A is created, with the result packet expanded to include whatever additional items are needed to implicitly link A to B. This is justified by adding the same items to the spinoff's context as well.

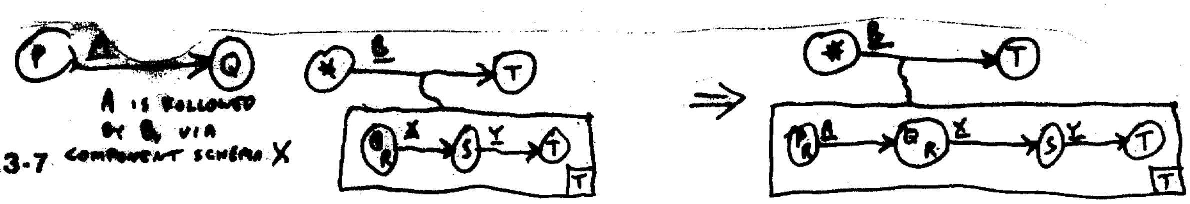



Figure 2.3-7

But extended annexation is contraindicated if the negative of a needed addtitional item is present in A's context or result.

The annexation process is really quite conservative: it merely collects information already implicit in existing schemas-- namely, the information that a given schema can serve as an entry point to a path leading to a particular goal. But this can be very useful: if schemas A and B are only *implicitly* linked and if B's result contains a set of goal-biased items, schema A has no way of knowing/asserting that it can be of assistance, even if it is alerted and B is not. Of course, it might be thought that a schema competing for selection could search for other schemas to which it's implicitly linked, checking their result packets for goal-biased items, as well as the results of schemas one more level of indirection away, and so on. But this seems like it would be a very hard computation to perform in parallel, and, in any case, the search could only reach finite depth in finite time-- remember, this would have to be done every clock tick to select a new schema. So instead, we can let annexation assist such a search: the schema selection process can, in effect, search in parallel, in *fixed* time through *arbitrarily* long linked paths-- but only among those paths whose linkage has been made explicit by the construction of a compound action. To see how this is so, consider, for example, what would happen if implicitly linked schemas A and B were to be activated in succession, A thus being annexed to B's compound action. Subsequently, the augmented B would be alerted when previously only A might have been, since now B's action-enable item is set by A's context too. So the path through schema A to B's result packet would now be assertable if B's result were to contain a goal-biased set of items. In general, the alerting of any compound action component lets it be asserted (via the action-enable item) that *some* path is known to exist from the current state of the world to the state denoted by the action's terminal packet. The path itself need not be traced through until (unless) a schema using that action is activated, at which time the path is elaborated incrementally, conditionally, and dynamically.

A new criterion of schema selection comes to mind here: whenever a schema has just **finished** activating, its result items are **continuity-biased**. This is like goal-biasing, except that continuity biasing promotes the selection of schemas whose *contexts* include the biased items. Continuity biasing promotes annexation by encouraging successively activated schemas to be related to one another via implicit chaining.

### 2.3.3 Repair

Consider next what happens if a path through a compound action is unsuccessful: say schema B below is autoactivated as part of X's compound action, but say B's prediciton fails and item P remains off.

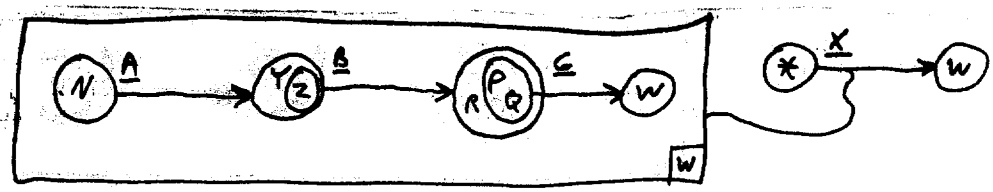


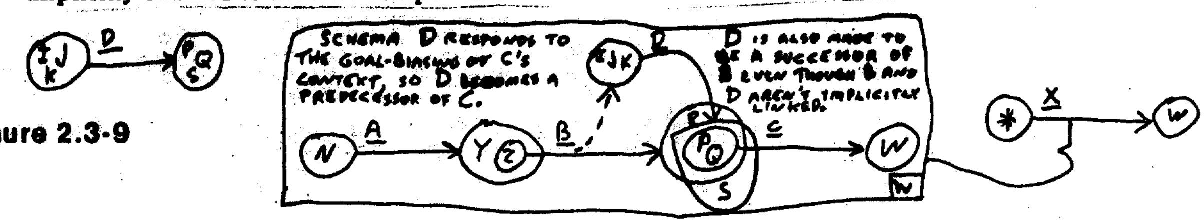

Figure 2.3-8

Schema C does not proceed obliviously on its way since, although in receipt of the go-ahead signal from B, C is not alerted. No other components have been told to go ahead, so the compound action proceeds no further.

I now propose a **repair** procedure by which the compound action's incapacity might be remedied. The idea is to try to enlist new component schemas-- or to use old components in new ways-- both to resolve the problem, and to incorporate the solution into the compound action in case the difficulty recurs. Thus:

* the context of the component schema which didn't get alerted can be goal-biased. A schema which responds successfully to this biasing will alert the stranded component and allow the compound action to proceed. Moreover, I'll stipulate that that schema be linked into the compound action, as predecessor of the stranded component and successor of the component whose prediction failed. Let's allow this successor relatic.ıship even if the failed component isn't implicitly chained to the new component.

Figure 2.3-9




True, in that case it may just be lucky that the new component was alerted when needed. But the action is no worse off with the new component than it was before. At worst, the new component won't be alerted in the future following the autoactivation of its predecessor; but if some other successor *is* alerted, the unavailability of the the new component won't even be noticed. At best, the new schema's predecessor may facilitate its being alerted, making it useful that they were linked together.

* more simply, when the compound action is interrupted, there might be a component of the action already alerted, though not given the go-ahead. The repair process can give such a schema a chance to resume the action, linking the schema as a new successor to the component that failed. (The go-ahead signal can't be waived like this in the ordinary, non-repair, operation of a compound action, since earlier components in the path may not cease to be alerted after they activate successfully; it would be chaotic to have components continue to autoactivate even after their turn in the path.)

* finally, if all else fails, the user schema's result items (which should include the compound action's terminal packet items) can be goal biased. Any schema which responds to the challenge can be linked as a new terminal schema in the compound action. In fact, this can be done even when the user schema has no failure at all. It was already proposed in section 2.1 that a schema's activation cause the goal-biasing of its result items. The linkage of a new terminal schema provides further motivation for doing this: it allows disparite paths to a goal to be linked together

into a composite structure which (externally) doesn't care which path is taken.

Each compound action has a number of reliability indexes, one per component schema. Each measures the likelihood of the compound action successfully reaching the terminal schema when the action was initiated via that component. If a user schema is activated and the terminal packet is reached, the reliability index associated both with that compound action, and with the component schema which enabled the action, is upgraded. This happens even if the schema was interrupted and had to be repaired, provided that the repair worked, ie the terminal schema was reached. Otherwise, the index in question is downgraded. The reliability asserted by an alerted user schema competing for activation is the product of the schema's own reliability index (measuring the likelihood of satisfying the schema's result if the compound action goes to completion), and the relevant action-reliability index.

Each compound action also has one "efficiency index" per component schema. This estimates how much time will elapse before the action completes if the action is initiated with that component. The efficiency index is updated according to the actual time elapsed in a given activation. A good efficiency index favors a schema in competition for selection.

The repair process, though similar to annexation, is less conservative. A schema whose compound action repeatedly fails to proceed as expected, but which is successfully repaired each time, is considered all the more reliable-- even though there is no necessary reason to think that repair components will be alerted again when needed (since their predecessors need not be implicitly linked to them). In effect, the repair process lets the Schema mechanism count on its own resourcefulness to get a schema out of a jam, to the extent that such resourcefulness has been shown to work before. Only if a compound action is interrupted irreparably is its reliability impugned.

### 2.3.4 Correlation and Customization

So far, the user schemas we've been considering have been closely bound to their compound actions: each user context has contained only an action-enabled item, and the result packet has been the same as the compound action's terminal packet. It's necessary to consider what happens to a user schema when correlation introduces additional items to the context and/or result. As far as possible, this should be no different than for a schema with a primitive action, but it turns out that some special considerations are needed when new items appear in the user context.

In order for a user schema to be alerted, some component of its compound action must be alerted, *and* any additional items in the user context must be satisfied. Thus, the additional items effectively superimpose extra constraints on *all* of the action's components. But suppose some of these additional items get cleared in mid-activation of the compound action. The action should then be interrupted, for nothing entitles us to believe that the user schema is still applicable. This interruption is easily enough arranged: simply require a schema to remain alerted if it is to remain active (recall that once a user schema ceases to be active, its components will not autoactivate). But a subtle problem arises here: perhaps a certain component of a compound action itself results in the contradiction of one of the items in the context of a user of that action. The action will abort as soon as the contradiction is detected, but if this mistake is not to be repeated perpetually, the contrary component must be disqualified from further participation in the incompatible user schema. Thus, the idea of **customization** arises: a counterproductive component can be removed from a particular user's copy of a compound action, without affecting other users' copies.

Aside from this customization, the revision of a user schema's packets has no effect on the components of its compound action. As for the effects of component schemas' correlation-revision on a user schema, I'm inclined towards a linear strategy, trying to minimize such influence. So I assume that a user schema will not be changed by the effects of correlation on its component schemas. If a spinoff version is made of a component schema, the original alone remains in the compound action. If the original proves highly unreliable and the spinoff is much better, let the repair process discover that and splice the spinoff in.

## 2.3.5 Implementation *

A compound action can acquire arbitrarily many component schemas. This suggests that it might be difficult to implement the copying of compound actions in bounded time. But something like Fahlman's **virtual copies** idea [Fahlman] can come to the rescue here: it's possible using virtual copies to have individually revisible versions of a compound action without actually copying any components, requiring only a small, local adjustment for the replication, augmentation, or customization of an action.

The central features of compound action implementation are these:

* There is an action enable packet which governs the state of the action enable item. This packet contains the alert item of each component schema of the compound action. The disjunction of the items in this packet sets the enable item.

* Each component schema has an **autoactivation packet**, the conjunction of whose items allows the schema to be activated without competing for selection. The autoactivation packet

contains the action's run item and some go-ahead items: these are just the run items of the component's predecessor schema(s) (slightly delayed so that a component can autoactivate when its predecessor has just finished running). If a component has no predecessor, its autoactivation packet includes a dummy go-ahead item which is on momentarily when the action's run item is first set.

* Each compound action has an **action run packet** containing the run items of all the schemas using the action. The disjunction of these items sets the action's run item, which in turn affects the autoactivation packets containing that item.

So, the addition of a new component to a compound action requires only that:

* the component's alert item is added to the action enable item;

* a new autoactivation packet, containing the appropriate action-run and go-ahead items, is attatched to the component; and

* the component's run item is placed in the appropriate autoactivation packet of the component's successor.

"Copying" an aciton when forming a spinoff schema is even easier than adding a component: simply add the spinoff schema's run item to the action's run packet. Here is an illustration of the ties among a user schema and its action's components:

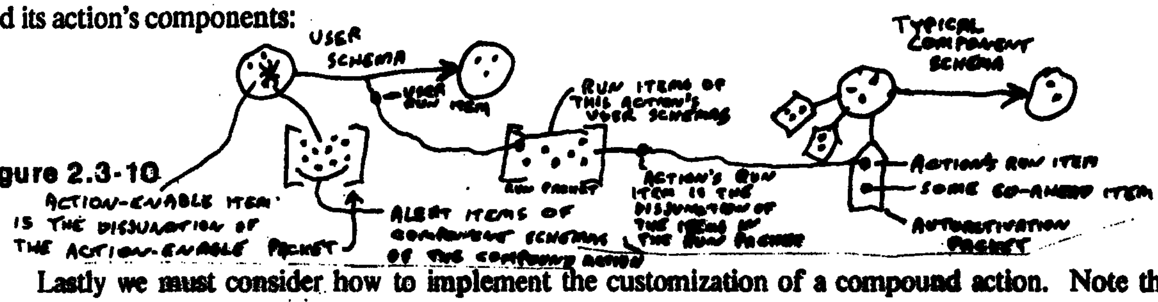



**Figure 2.3-10**

Lastly we must consider how to implement the customization of a compound action. Note that if a component schema causes some of the items in its user's context to clear, the action's run item will clear since the user run item which was keeping it set will clear. This will disable all the autoactivation packets associated with that action, and allow the user schema to compete for selection for doing repair. To exclude the offending component schema, we might add the negative of the user schema's run item to that component's autoactivation packet. However, should the current user schema give rise to a spinoff, the spinoff schema ought to inherit the restricitons of its parent; this wouldn't happen above, since the spinoff schema's run item wouldn't affect the previously excluded component. Thus, a modification: in place of the negated user run item, we include a negated **version ID** item in the autoactivation packet. This item is established, upon the creation of any spinoff schema, such that the item is set by the new schema's run item, or by any item which could set the old schema's version ID item. Thus, the exclusion of an action's component will be maintained through all replications of that action.

<u>Summary</u>

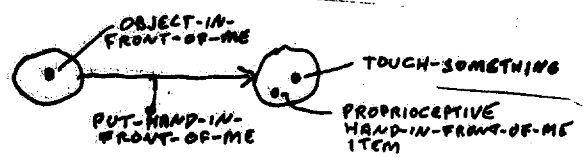
The compound action feature tries to do something which has no right to be doable: to build and extend a coherent structure by local, incremental processes which have no anticipation or understanding of the eventual coherence. The balance between the anarchy of local construction and the purposeful coordination of global functioning is a delicate one, and might prove unfeasible. But I think the ideas put forth in this section are promising enough to warrant further investigation.

## 2.4 Synthetic Items

The previous section proposes a kind of schema *action* more powerful, and on a higher level of abstraction, than primitive motor actions. In this section, I suggest a new kind of *item*, intended to represent higher-level things than primitive sensory inputs.

Consider a schema which remembers the position of an object, in the sense that it predicts that the object's rediscovery will result from moving the hand to a given place:

**Figure 2.4-1**

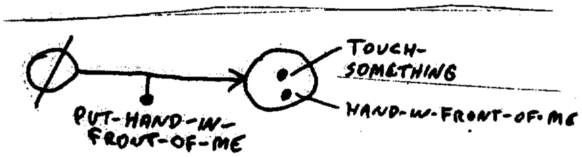



If this schema is to facilitate elementary Semsorimotor object-permanence, the object-in-front-of-me item must be able to remain set even when the object isn't perceived. This item therefore could not be any sensory-input item, or function of such items. Instead, it is paradigmatic of a new, **synthetic** item which is now introduced. This section talks about how a synthetic item is fonned, and how it is set and cleared.

### 2.4.1 Formation

The presence of an object at a given location has the property that, by default, it often remains constant. If not for this, it would be unreasonable to assume the continued presence of an object after it ceases to be perceived, and a schema such as the above would have no validity. This observation motivates the method I propose for creating a synthetic item. The following null-context schema

**Figure 2.4-2**

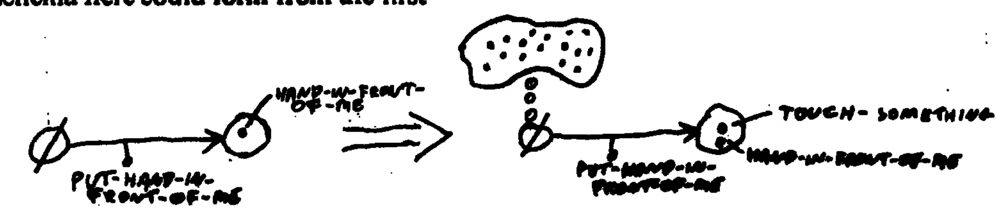



resembles figure 2.4-1, but the object-in-front condition isn't present. In this form, the schema would not be generally reliable, but (I'll assume for now) the correlation process wouldn't find any context items whose addition would rectify that. Significantly, the schema would be locally consistent as to whether it worked or not-- when it succeeded once, it would tend to continue to be successful for a while (since the object needed for its success would, I'm assuming, tend to remain in place for a while); similarly, its failure would foretell further failures in the immediate future. These factors-- the unreliability of a schema, the failure to discover additional context items by correlation, and the local consistency of the schema's success or failure-- combine to suggest that some variable, external thing or condition is requisite to the schema's successful activation. So, we can place in the schema's context a synthetic item which, at first, is only a dummy meant to signify *whatever unknown condition is required for the schema to work.* This item has the property that once set, it remains on (at least for a while) unless something happens to clear it. Mechanisms are suggested below to elaborate a synthetic item's significance by discovering conditions under which the item should be set or cleared. But first, more detail pertaining to the item's creation:

> In addition to the reliability index, each schema has a **consistency index.** This, like the reliability index, is upgraded or downgraded on each activation of the schema. But while the reliability index goes up or down according to whether the current activation succeeds or fails, the consistency index changes according to whether the curent activation did the same thing (ie succeeded or failed) as the previous activation (or last several activations, or some such).
>
> Now, recall the correlation process's pruning and confirmation of a context balloon in search of necessary additions to the context packet. In section 2.2.2, I said that if the context balloon empties, it gets refilled to try again. But I lied; that doesn't necessarily happen. Instead, *if* the schema has a high consistency but not-so-high reliability, then a new item is synthesized, and added to the context packet (and the correlation attempt terminates).

The second schema here could form from the first

**Figure 2.4-3**



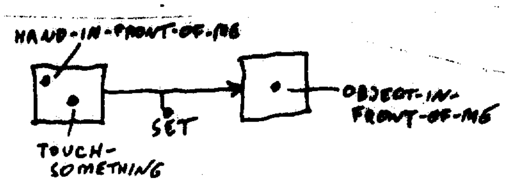

either by investigation (when touch-something occurred as a recognized surprise), or directly by correlation (if the touch-something result happened to occur on enough successive trials of the schema to be confirmed in a result balloon). In either case, a context balloon would appear (immediately in the case of investigation, or after the first unsuccessful activation in the second case) which would converge to nothingness, giving rise to the synthetic "object-in-front-of-me" item shown in figure 2.4-1.

## 2.4.2 Clue Schemas

It would be nice if touching-while-hand-is-in-front were to set the object-in-front item. More generally, there ought to be a way to recognize **clues** that a synthetic item should be set. Here's a possible way to do this, in two main steps: *proposing* a clue, and *verifying* a clue.

One simple way to propose a clue is: when a synthetic item is put in a given schema's context, the result packet of that schema is proposed as a clue that the item should be set. The rationale here is that the synthetic item designates some thing/condition in the world that can be viewed from some perspective; the given schema achieves that perspective, and with it a perception of the thing/condition. Above, for example, a certain "tactile perspective" is achieved by moving my hand in front of me; this enables me to "view" (by touch) an object in front of me. In the schema in figure 2.4-1, the proprioceptive hand-in-front-of-me item signifies that that particular tactile perspective has been achieved; the touch-something item signifies that something is indeed viewed from that perspective. We'd like to turn this schema around and propose that whenever that perspective is in effect, and a thing is in fact percieved (by touch), then there indeed exists an object in front of me. The proposal is implemented in a **clue schema**

**Figure 2.4-4**




The "action" taken by a clue schema is to set the item in the result packet.

Notice that if the proprioceptive hand-in-front-of-me item did not exist-- if touch-something were thus alone in the result packet in figure 2.4-1-- then an entirely inappropriate clue would be proposed: that whenever something was touched (regardless of where the hand was while doing the touching), object-in-front-of-me should be set. More generally, if a schema with a synthetic item in its context does not explicitly designate, in its result packet, the perspective which the schema achieves-- if it specifies only the perception which that perspective will produce-- then the proposed clue will be inappropriate. The clue proposal rests on the assumption that the perspective *is* explicitly designated.

Since the clue schema is thus formed on the basis of a weak heuristic, we'll make its initial reliability index low. (However, its novelty/curiosity values can be made high.) But this can be changed by the *clue verification* process: This consists of applying the reliability-update and correlation mechanisms to the clue schema, in much the same way as to an ordinary schema-- except, a clue schema's success or failure isn't detectable until arbitrarily later, when the synthetic item is "tested". To facilitate this deferred testing, an item which was set by a clue schema keeps a pointer back to that schema. At some point, the item may be in the context of some activated schema. If that activation is successful, the *clue* schema's reliability is upgraded; otherwise, the clue schema's reliability goes down, *and* the correlation mechanism in invoked, attatching a balloon to the clue schema's context in an effort to appropriately constrain it.

Some details:

* A synthetic item has a **confidence level** which is initialized high whenever the item is set, and which degrades with time. When this level reaches some minimum, the item is no longer counted on as being either on or off.

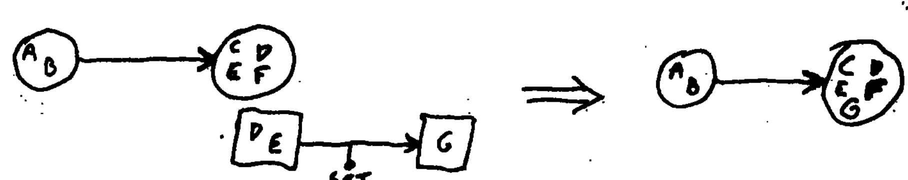
* Initially, a clue schema must compete for selection. But when it becomes sufficiently reliable, the clue schema autoactivates whenever alerted (the clue schema is then said to be **mature**). Thus, a synthetic item will eventually have its state maintained automatically by clue schemas activating invisibly in the background. These schemas will give substance to the original "dummy" item: the item's initial meaning is "whatever condition in the world makes the schema in whose context I appear work". But this serves only to seed the crystallization of clue schemas; once these congeal, they become the synthetic item's primary definition. This definition is analogous to whatever meaning a given sensory item has by virtue of its wiring to some input device; but the synthetic item's meaning has been programed. (Some subtle questions may arise concerning the meaning of an item when, pathologically, its clue schemas contradict one another; I haven't yet pursued this issue.)

* Following a successful test of a synthetic item, the item's back-pointer is erased, absolving the clue schema of responsibility for any subsequent refutation of the item. Or, corroboration of the item by another, more reliable, clue schema could be considered a successful test of the item, and shift responsibility (by changing the back-pointer) to the later clue schema.

* When a clue schema is immature, it's probably a good idea to **curiosity-bias** the item which it sets. This adds to the curiosity value (but *not* the goal-pursuing value-- recall section 2.1) asserted to the selector by any schema whose context contains that item. Thus, testing the item-- and refining the clue schema-- is encouraged.

* When a schema is activated whose result packet contains a clue to some synthetic item which is then set, a spinoff could be made which includes that synthetic item in the result:

**Figure 2.4-5**




So far, I've only discussed ways to *set* a synthetic item. Now here are some ways to clear one. First, a synthetic item is said to be **refuted** if a schema in whose context it appears is activated unsuccessfully; the item is then cleared and no correlation revision is done to the schema, since the *item* in question is (belatedly) regarded as faulty. More precisely, this happens only if there is but one unconfident item in the schema's context-- otherwise, there is ambiguity as to which item might be at fault (all suspects might be curiosity-biased to promote discovery of the culprit). More precisely still, an item can be refuted only if it had been set by an immature clue schema. This is in keeping with the philosophy that only until its clues are well established is an item defined solely by the schema in whose context it appears; afterwards, an item is on if its clue schemas say it's on, and a schema which is disappointed in its reliance on that item is itself regarded as incorrect, and dealt with by correlation.

An item can be cleared by being refuted; but it would also be nice to recognize a *clue* that an item should be cleared. For example, *not*-touching-something-while-hand-is-held-in-front is a good clue that the object-in-front item should not be on. So, we ought to have clear-clue schemas which function and are verified analogously to set-clue schemas, but are proposed as follows: if a schema with a single immature item in its context is activated and fails, then the conjunction of the *actual* (as opposed to predicted) states of its result items becomes a clear-clue. For example, if the schema

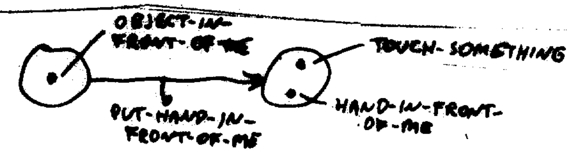



Figure 2.4-6

is activated without an object existing at the expected place, the schema will fail because the touch item won't go on. So there will appear a clue schema

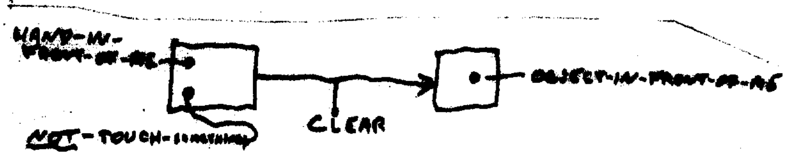



Figure 2.4-7

Here, as with the proposal of a clue to set an item, there is an assumption that the schema from which the clue schema is spawned designates both the perspective, and perception from that perspective, expected to result from the schema's activation. Here it is further guessed that when the schema is unsuccessful, the perspective has been achieved but the thing/condition designated by the synthetic item is not present to be perceived.

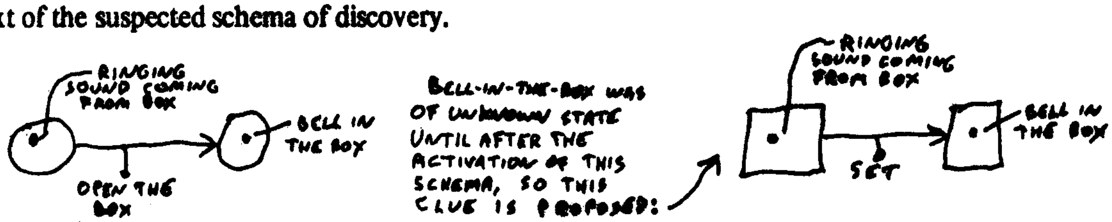
Of course, not all proposed clues will be as reasonable as those of the convenient examples of this section. But the correlation process can intervene to correct, or simply discredit, a clue schema which fails to perform well.

### 2.4.3 Discovery

When a synthetic item appears in a schema's result packet, there is a potential ambiguity in the schema's interpretation: the schema's action may *cause* the condition denoted by the item; or, it may merely cause the *discovery* of an already-existing condition (for example, a schema may say that when a ringing sound is heard in a box, the action of opening the box will result in a bell in the box; the action doesn't cause the object to exist there, but merely discovers it). In the latter case, the schema's context ought to be designated as a clue for the synthetic item: if the condition is merely discovered by the schema, the condition must already exist whenever the schema can be activated.

If, on a given activation of a schema, a result item was known to be off when the schema was activated, the ambiguity disappears: here, it is reasonable to assume causation by the schema. But if the state of the item was uncertain until just after the schema's activation, discovery can reasonably be proposed. This proposal takes the form of an immature clue schema which sets the synthetic item. The clue schema's context copies the context of the suspected schema of discovery.

Figure 2.4-8




Note that bell-in-box had already been determined to be a reliable result of opening the box in the specified situation; only the interpretation of discovery rather than causation is new in the above proposal.

<u>Summary</u>

The synthetic item feature is less well thought out than features of the previous sections. I'm fairly comfortable with the global aims for synthetic items, namely:

* the formation of synthetic items to represent things not immediately preceived;
* the "maintenance" of synthetic items' states in response to previously discovered "clues";
* the participation of synthetic items, so maintained, in ordinary schemas.

But the proposed implementations of these points seem somewhat unconvincing, and will probably be substantially revised.

# 3. Sensorimotor Scenario

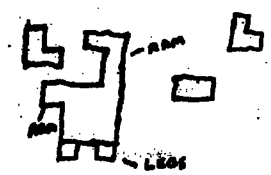
In this section I attempt to mutually reinterpret the Schema mechanism and Piagetian Sensorimotor theory. This is done by means of a Scenario suggesting how the Schema mechanism might accomplish some of the milestones of Sensorimotor development.

The Scenario has only been thought through to the beginning of the third Sensorimotor stage, and even that much is fragmentary. Many details are yet unconsidered. No doubt, this permits a twofold distortion in the Scenario: the mechanism's capabilities may be exaggerated, and the capabilities needed to account for various Piagetian phenomena may be underestimated. But that's all right. The Scenario is only intended as an *interpolation* between the current mechanism's power, and the requirements of the Sensorimotor progression. The difficulties encountered in filling in more details will guide the next round of revisions.

## 3.1 Microworld

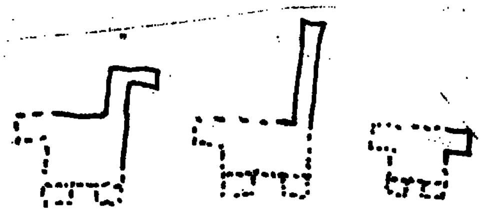
Let's imagine a two-dimensional universe populated by small, simply-shaped objects, along with a "body" to be controlled by the schema mechanism:

**Figure 3.1-1**



The body has two legs (for "walking") and two arms, each of which can extend itself in various directions, in discrete increments

**Figure 3.1-2**



up to a maximum length of, say, five increments. For each arm there exist *four* motor actions, for extending the arm one increment forward, back, left, or right (relative to the body). Also, for each of the possible (body-relative) positions of the tip ("hand") of the arm, there is an action which will take a sequence of sub-actions that will leave the hand in that position. I'll call these **positional** actions, and the first kind **incremental**.

Two kinds of primitive sensory items are associated with the arms: "touch sensor" items map onto the perimeters of the arms and go on when their regions come into contact with anything. Also, for each hand position, there is a **proprioceptive** item which is on whenever the hand is in that position.

The final hand feature is the ability to **grasp**. Let's posit a primitive motor action which, when activated, causes anything touching the hand to stick to it. This state continues for a short while, or until the **ungrasp** action is taken. There is also a **grasping** sensory item which is on when an object is being grasped.

Next, there is a "visual" system. This consists of a **retina**: a (say 5x5) array of items that maps onto a portion of the world in front of (ie above) the body. If some object takes up most of the area onto which a retina item is mapped, that item is on. So, for example, in the scene shown in figure 3.1-1, the retina might show:

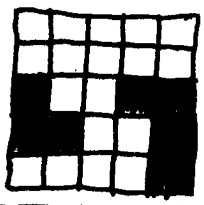


**Figure 3.1-3**

As with the arm, two kinds of motor actions exist for the "eyes": four incremental "glance" actions to shift the retina's mapping one unit forward, back, right, or left (with a range of say 5x5 possible orientations); and 25 positional actions to direct the glance to any one of those orientations. Twenty-five proprioceptive items exist to designate the current glance orientation.

I'm using considerable poetic license by calling this system "visual". Certainly I don't mean to imply a strong resemblance to natural vision. I don't, for example, mean to suggest that human intelligence, at the Piagetian level of abstraction, is (necessarily) at all concerned with the direct state of the retina, unprocessed by intervening levels of feature detectors. The intent here is to have a sufficiently constrained universe-- with each object mapping reliably to a certain simple pattern on the retina, with no rotations and no depth perspective-- that special feature processing becomes unneccessary.

The extent to which this scheme *is* intended to model human vision is that:

* it provides a perception of a field of space capable of containing 0, 1, or several objects; and
* by redirecting the glance, the mapping from objects to perception can be changed without any (obtrusive) physical displacement of anything.

Similarly, the hand/arm system is intended to resemble its human counterpart only qualitatively and abstractly.

Elaborations and extensions of the microworld will be mentioned as needed.

## 3.2 Original Schemas

Some schemas must be supplied at the outset. Let's assume the following "reflex" schemas:

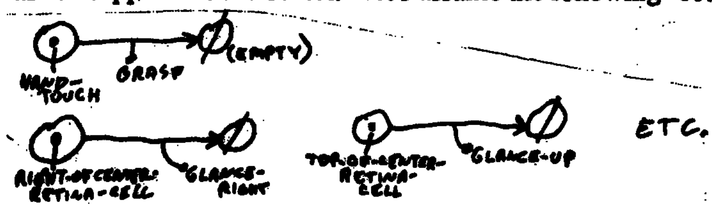



Figure 3.2-1

where "etc" denotes a collection of schemas, each of which has some single retina item in its context, and each of which has whichever incremental glance action will move the "image" towards the center of the retina. All of the reflex schemas have empty result packets.

In addition, for every action (incremental and positional hand and glance actions, plus grasp and ungrasp) there exists an initial schema with empty context and result:

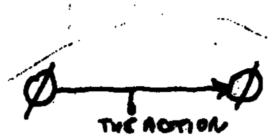



Figure 3.2-2

These are given less initial "novelty value" than the reflex schemas. That way, an alerted reflex schema will tend to be selected over these others: but when nothing else is happening, these action-only schemas can be played with.[1]

In order to encourage looking at things and touching them, all retina items and tactile items are intrinsically goal-biased. The central items are especially biased, to promote a canonical view of things; the touch-hand (-front, -right, and -left) are also especially biased.

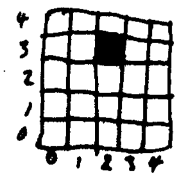
1. Remember, the smaller the context, the stronger the schema's assertion. A schema with an empty context is <u>always</u> alerted.

## 3.3 Development

### 3.3.1 First Stage

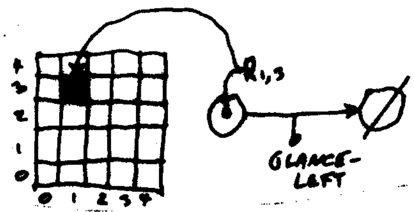
Suppose we provide an object whose image takes up a single retina cell. If the object appears here

**Figure 3.3-1**



and the (alerted) schema shown is activated, then the next state of the retina should be

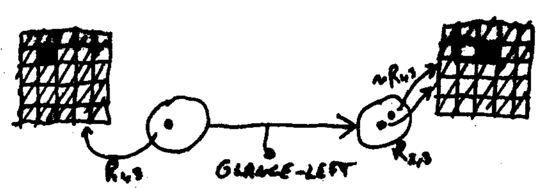
**Figure 3.3-2**




The item $R_{2,3}$ is a surprise, as is the negative of $R_{1,3}$, since no activated schema predicted the state changes. Thus, by correlation (section 2.2.1), these get placed in a result balloon on the above schema. Any coincidental surprises that may have occurred-- tactile, auditory (supposing for a moment that we have that modality), or even visual (in other parts of the retina)-- would also contribute to the result balloon. But these items would tend to be purged on subsequent activations of this schema. Only $R_{2,3}$ and $-R_{1,3}$ would be likely to survive long enough to be confirmed and placed in the result packet of a spinoff schema. If they, too, were purged-- no schema is *totally* reliable-- they would always have another chance in a later balloon. Conversely, if by unusual coincidence some extraneous items came to be included in the spinoff schema, all would not be lost: the reliability index of that schema would degrade until the schema was thoroughly discredited, and a revised schema would be spun off which would be free of the incorrect items.

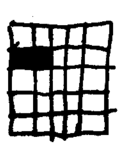
In any event, there would eventually come to be a schema:

**Figure 3.3-3**




Now, if that schema were activated here

**Figure 3.3-4**



then the $-R_{1,3}$ prediction would fail. The balloon attatched to the context of the schema (recall section 2.1.2) could be expected to discover the reliable schema

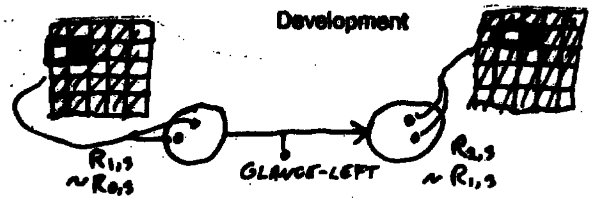



Figure 3.3-5

while purging of the result would yield (also reliably) the less bold schema

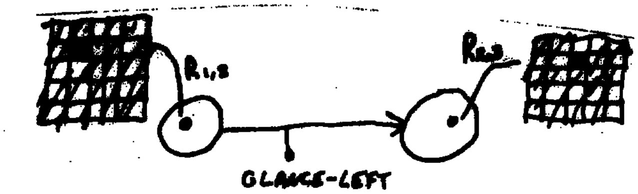



Figure 3.3-6

Note that the Schema mechanism must *learn* that when something moves to a new place, it disappears from the old one. Here we see this happening to first-stage images; but when items form to designate object-like entities in real spatial positions, this lesson will have to be relearned on a higher level.

Note too that a schema pertaining to one portion of the retina does not automatically generalize to other parts. For example, the schema

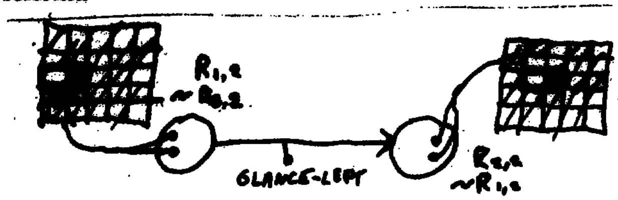



Figure 3.3-7

would develop separately from, but analogously to the schema in figure 3.3-5.

Similar schemas can evolve from each of the visual reflex schemas (except, those whose contexts have an item at the edge of the retina would have less reliable "bold" versions-- see why?). So there would come to exist a network of schemas

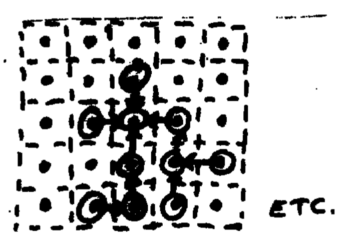



Figure 3.3-8

converging on the center of the retina. Successive activations of elements of this network would occur from time to time, clumping the schemas into compound actions (as per section 2.3.2). This would culminate in the creation of a schema with one large "map-to-center" compound action incorporating the whole network:

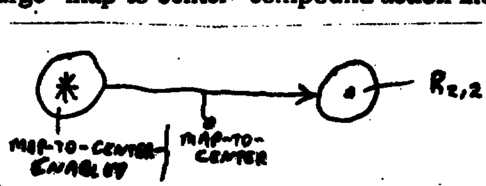


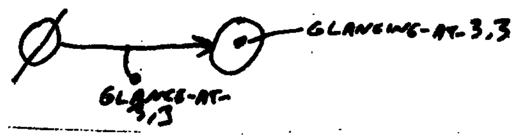

Figure 3.3-9

The action is enabled when any image appears on the retina, and has the effect of incrementally shifting the glance until the image is centered.[1]

Note that this map-to-center schema could also incorporate the "feedback" effect described in section 2.4.1. When a component of the map-to-center action is unsuccessful (say because the object moved slightly) another component is likely to be alerted instead and can be spliced in by way of repair.

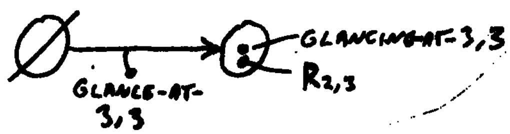
Let's consider the effects of activating the various "empty" schemas incorporating the positional glance actions. Ordinarily the only reliable surprise produced by such a schema would be turning on the proprioceptive "glancing-at-x,y"[2] item. We would thus expect schemas of this form:

**Figure 3.3-10**



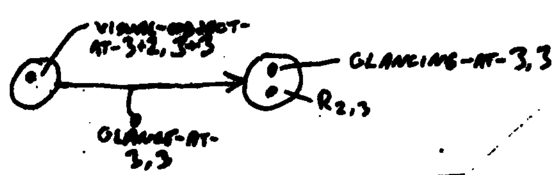
But suppose some object is situated such that glancing at 3,3 brings the object to view, say at $R_{2,3}$. If this activation is repeated many times with the same effect, we will have

**Figure 3.3-11**




This would lead to disappointment, however, when the schema was activated in the absence of the expected object. Balloons attatched to the context would try to find a stronger condition for the schema's prediction, but no known item would fit the bill. So a new item would be synthesized (section 2.4.1):

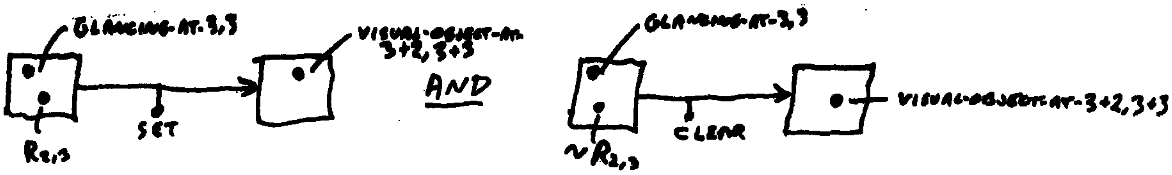
**Figure 3.3-12**




The formation of clue schemas

---

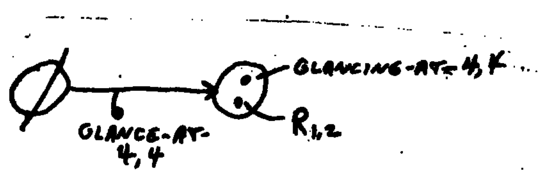
1. There may also develop map-to-off-center actions. The visual reflex schemas are set up to encourage mapping to the retina center, but such other mapping schemas as may evolve would do no harm.
2. The x,y here refers to the (body-relative) mapping of the retina, not to a cell in the retina.

Figure 3.3-13


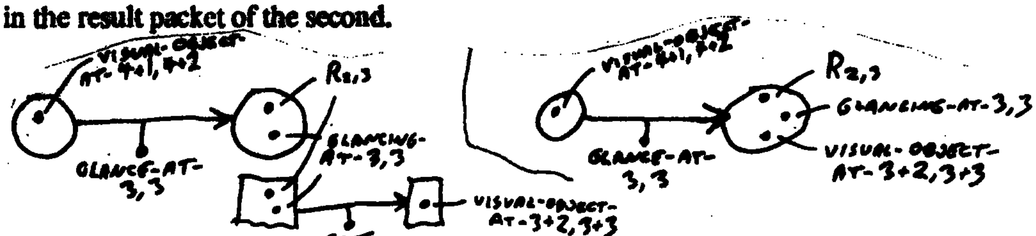

would occur as a straightforward analog of the clue-formation examples in section 2.4.2 and 2.4.3.

Similar synthetic-based schemas and clue schemas would form for each of the positional glance actions. Each such synthetic-based schema would know how to look back at a perceived object after the glance was redirected away from the object. Looking in the right direction and *not* seeing the object would clear the synthetic item, dashing any further expectation of the object.

Note that the visual positional items are not independent from one another. For example, the correlation revision of

Figure 3.3-14



would ideally come to include the previous object-at-3+2,3+3 item, rather than create a new one (figure out why this would be correct, if it's not already clear). If you step through the correlation sequence, you'll see that, depending on whether or not the object-at-3+2,3+3 item had been set before the above glance-at-4,4 schema was activated, that item may or may not come to be included in the schema's context balloon. If it is not, an unnecessarily distinct object-at-4+1,4+2 item will be synthesized. But object-at-3+2,3+3 can eventually merge with object-at-4+1,4+2 when each is found to be a clue for the other. The first schema below (which can arise by correlation) gives rise to the second one since the result packet of the first holds a clue for the item in the result packet of the second.

Figure 3.3-15



By the attribution of discovery (section 2.4.3), visual-object-at-4+1,4+2 could then be proposed as a clue for visual-object-at-3+2,3+3. Other schemas could propose the converse clue, establishing the equivalence of the two items.[1]

1. Perhaps there should be a provision for merging two items into one when each is established as a clue for the other.

Notice that n*m schemas of the form of fig. 3.3-13 and fig. 3.3-14 can develop, where n is the number of retina cells and m is the number of glance-mappings the retina can assume. Although this is not intolerably many when n=m=25, it is uncomfortable to contemplate an extension of this to larger retinas/glance ranges, or to other kinds of large networks of transitions. But this is mitigated by the following considerations:

* Only a subset of these schemas (proportionate to n or m alone) is needed to encompass the visual field. For instance, we could rely solely on those schemas whose actions map the image to the retina center. The development of this subset of the map-to-x,y schemas would be especially encouraged by the prejudicial arrangement of the visual reflex schemas (tending to center an image), the accessiblity and fertility indices, and the goal-biasing of the central retina items.

More generally, any system of transitions might concentrate on a preferred canonical mapping.

* A larger retina (or other transition field) could be divided into low and high resolution items (perhaps more than two levels of these). There could form a schema whose first action component maps something onto the coarse center, and whose next component fine-tunes to a higher-resolution center.

Further tricks for circumventing n-squared problems are likely to turn up with more thought.

To recapitulate: so far the Scenario has included the formation of schemas which will

* predict the "motion" of an image across the retina as the glance is shifted-- in particular, there is a schema to shift the glance so as to move an image from anywhere on the retina, to the retina's center.

* remember the "position" of an image, so that it is known where to look in order to see it again.

These have the character of some of the (solipsist) visual schemas of the first Sensorimotor stage. The development of similar "prehensile" schemas can also be anticipated. By the same processes as described above, schemas could form which:

* predict the migration of a touch sensation along the hand/arm as the hand is moved past an object-- in particular, a schema which, when an object touches any part of the arm, will incrementally move the arm until the object touches the hand; and

* synthetic items which delcare the presence of objects at various positions, with accompanying clue schemas; eg

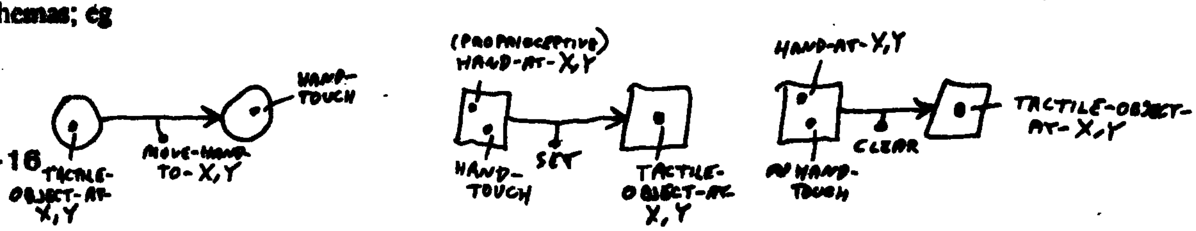



Figure 3.3-16

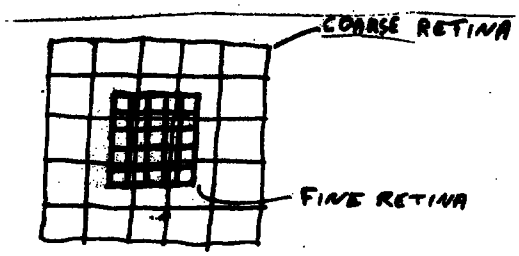
There would also be schemas dealing with grasping, and with moving the hand while the grasp is still in effect. These would be rather boring, though, since moving-while-grasping usually (in our microworld) entails no tactile change since the grasped object stays in the same hand-relative position. However, changes *are* producible with respect to other sensory modalities, and here the fun begins.

## 3.3.2 Second Stage

Recall the coordination of visual and tactile/prehensile systems which occurs during the second stage. Here's a suggestion of how the Schema mechanism might accomplish a like intertwining.

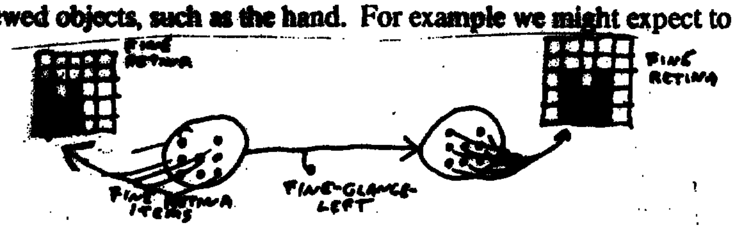
Schemas can develop which predict the visual consequences of moving the hand. For example, suppose a schema has been activated to look squarely at the hand, so that the hand appears in the center of the retina. In order to let this be distinguishable from other objects appearing there, let me retroactively postualte this addition to the microworld: besides the retina discussed so far, there is also a high resolution retina (I'll refer to the "coarse" and "fine" retinas). The fine retina has 5x5 items which map onto the same space as a 2x2 area at the center of the coarse retina.

Figure 3.3-17




There are also fine incremental glance actions of the same resolution as the fine retina. The "single-square" objects in the previous examples will now be assumed to ha e finer structure distinguishable via the fine retina; the hand, too, has such structure. All previous discussion of schema development should still hold, since the additional fine retina items ought not to interfere with the coarse items' inclusion in evolving schemas; this is assured by the linearity of the correlation process. However, specific visual anticipations can now form for often-viewed objects, such as the hand. For example we might expect to find

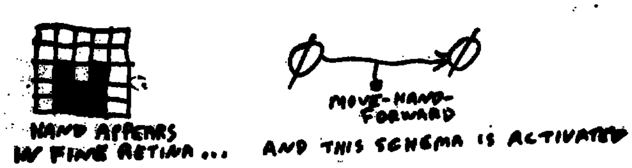
Figure 3.3-18




where the pattern shown is that of the hand. Similar schemas could anticipate the motion of the hand's image throughout the fine retina in response to fine glance actions.

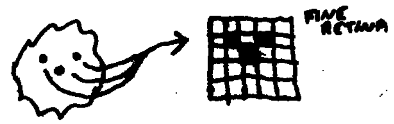
As I was saying, suppose the hand appears in the center of the retina. Suppose further that this schema

Figure 3.3-19



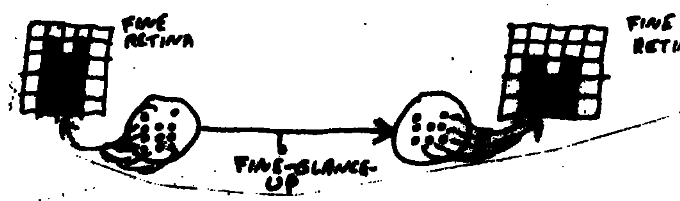

happens to be activated next. If we assume that incremental hand movements shift the hand by about one fine retina position, then we can expect these items

Figure 3.3-20



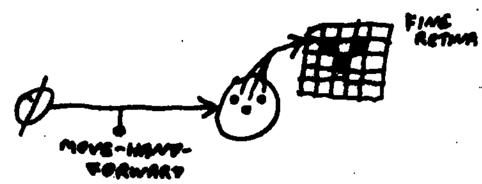

to turn on by surprise. If these attracted the schema mechanism's attention-- say by the activation of this schema[1]

Figure 3.3-21



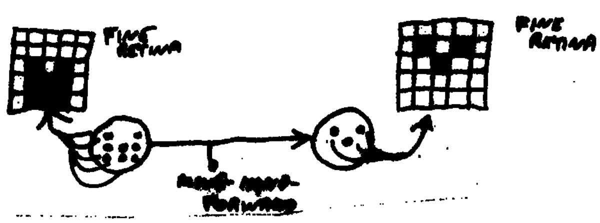

then the above items would be a *recognized* surprise.[2] By investigation (section 2.2.3), it would be included in the result packet of an exploratory spinoff of the move-hand schema

Figure 3.3-22




As an exploratory schema, this would never be activated in its own right. But after many activations of the parent move-hand schema-- a small fraction of which activations would take place when the hand appeared at the retina center, thus reproducing the surprise-- the context balloon on the spinoff schema would converge to yield

Figure 3.3-23




Similarly, there could come to be schemas that anticipate the visual consequences of incremental hand movements throughout the fine retina. Given such a network, the investigation process can spawn a schema that knows how to touch an object seen in front of the hand:

1. Should it appear distasteful to rely on this partly coincidental succession, keep in mind that this discussion doesn't assume that the coincidence happens always or even often, but merely that sooner or later it happens a few times (not necessarily consecutively).

2. It occurs to me here that a schema ought to be favored for competition for selection when the items in its context packet are surprises. This would work with continuity-biasing to promote investigation and annexation.

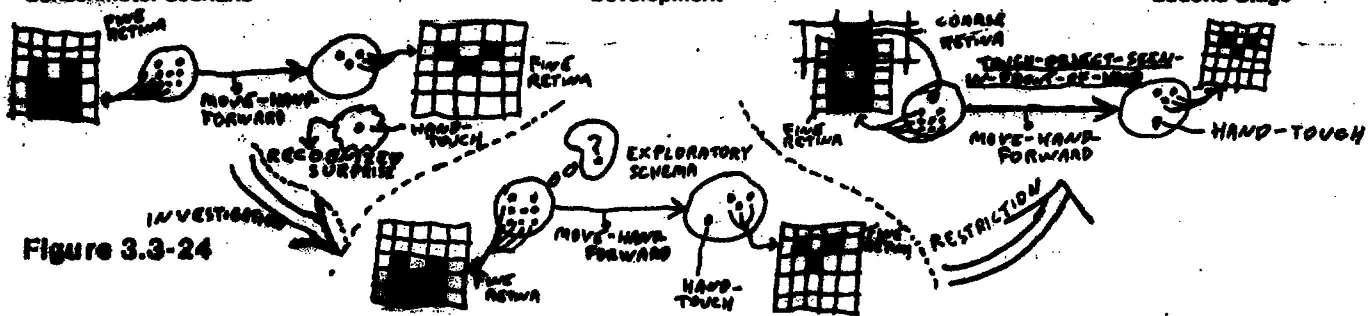



Figure 3.3-24

and similar local-touch schemas for objects nearby to the right and left. New entry points, of the form of figure 3.3-18, could be annexed to these schemas' actions, so that a slightly off-visual-center hand can first be moved to the position familiar to the last schema in figure 3.3-24, after which that schema can move the hand to touch the object. The local-touch schemas could be implicitly linked to grasp-what's-touched, and so be annexed to that schema's action to form a local-grasp schema, capable of grasping an object when both the object and the hand are in view. This schema, in turn, could be linked to a move-grasped-object-to-mouth schema (assuming a mouth for now); then, anything seen near the hand could be grasped and sucked.

If we had a move-hand-into-view schema, it could be coordinated with local-grasp to allow an object to be grasped when (initially) seen alone. The following sequence outlines the possible development of move-hand-into-view:

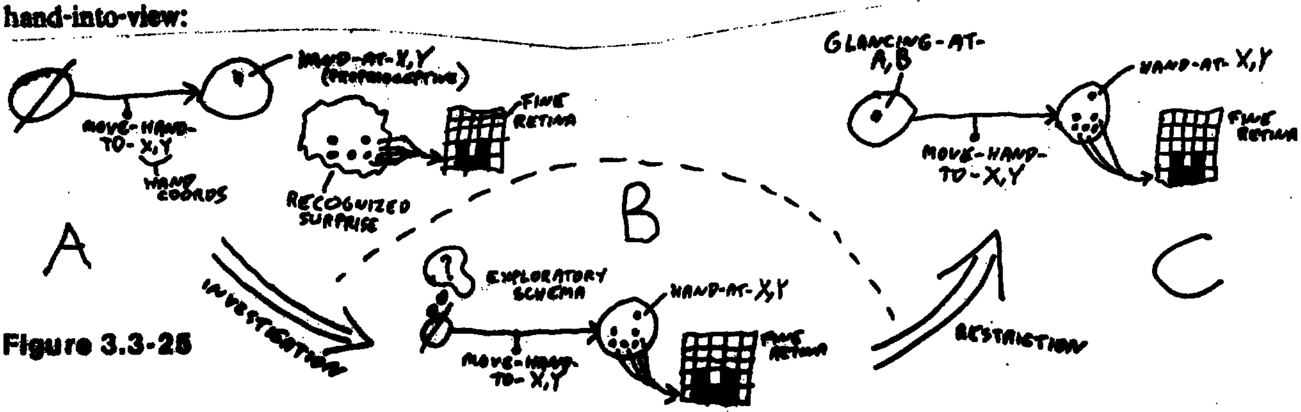
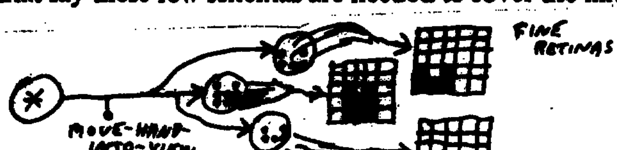



Figure 3.3-25

In A the recognized surprise occurs when a positional hand action happens to place the hand so that it appears in the center of the retina; then, some other schema (unspecified here) is activated which includes the see-hand items in its context, making the surprise a recognized one. So (B) an exploratory schema is spun off, the surprise items in its result packet. In order for the specified hand motion to bring the hand's image to the retina center, the retina mapping must happen to be centered at the hand's destination. In C, the correlation process has added this condition to the context packet.

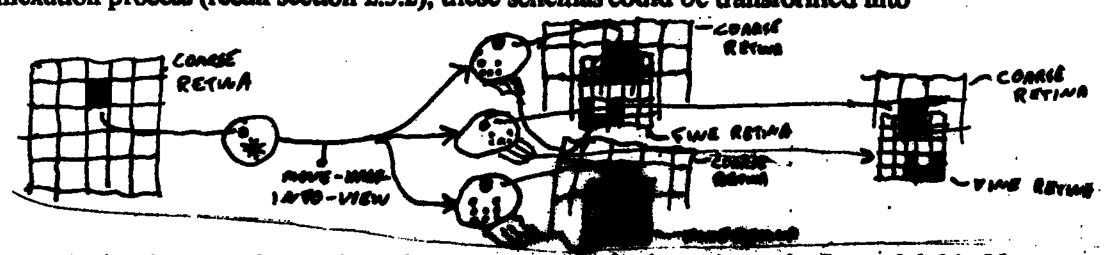
Schemas like this could develop for all glance-mappings, joining to form a compound action capable of moving the hand into view regardless of where the gaze was directed. Let's imagine that positional hand actions are somewhat inaccurate, so that say these few schemas are needed to cover the likely possible results:

Figure 3.3-26



By the extended annexation process (recall section 2.3.2), these schemas could be transformed into

Figure 3.3-27




in order to be linked to the local-grasp schema via such components as the last schema in figure 3.3-24. Now, an object can be grasped whenever it appears just above center in the coarse retina. If we have a map-object-to-just-above-center schema (just like map-to-center but with a different target), this can be linked to the above addition to local-grasp. The resulting action says: first, shift gaze to map the object to its canonical retina position; then, proceed to touch and grasp it. Thus, we finally have the capability to grasp whatever is seen nearby.

Two subtleties are worthy of mention here:

* I neglected to point out a complication in the development of the move-hand-beside-object action. The move-hand-into-view action would inevitably acquire "look-at-hand" components-- that is, components which result in the hand's appearance on the fine retina not by moving the hand, but by shifting the gaze. The compound action, after all, is organized around the realization of its terminal packet, and cares little *how* that result is arrived at. But this is a problem for move-hand-beside-object, since gazing at the hand will not bring the hand closer to the object. Happily, there is a reasonable fix: precisely because they move the eye rather than the hand, look-at-hand components would tend to violate move-hand-beside-object's context condition: the coarse retina item would turn off. Such components would thus interrupt the schema and, eventually, be excluded from a customized move-hand-beside-object schema (recall section 2.3.4).

This illustrates an important point about the coordination of sight and prehension. It is not enough simply to tabulate some cross-consequences of visual and tactile events. In order to abstract above particular sensory modalities, certain differentiations must be realized as well. In particular, the customization of move-hand-into-view reflects the discovery that hand motions alter but *part* of visual space, while eye motions shift it all. Complementary discoveries remain to be made about the effects on *tactile* space of eye and hand motions.

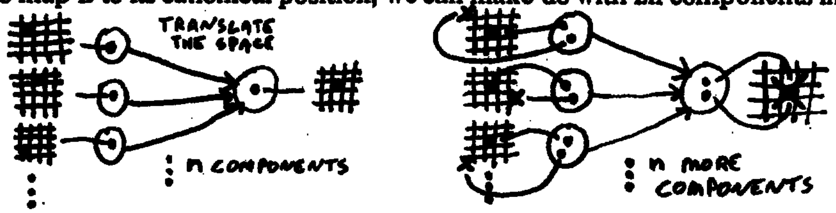
* Note that the grasp-what's-seen schema can move the hand from any of n body-relative positions to an object in any of n such positions, without requiring $n^2$ components to handle the individual cases-- this despite the Schema mechanism's characteristic insistence on separate, virtually-equivalent versions of schemas (such as the components of the visual map-to-center action). It seems that once a collection of items is organized into a coherent field by a network of actions (as for example the retina items are organized by such schemas as map-to-center), the organization can be exploited in subsequent development. In particular, grasp-what's-seen illustrates a kind of exploitation which I call the **canonical mapping** method. Its general form is this: to transform thing A from any of n positions in a space, to the vicinity of thing B in the same space, we can

* translate the space itself to be centered at B; this requires ::n components to take different actions depending on B's initial position in the space; and

* move A to the center of the new space; this again requires ::n components to branch on the possible initial positions of A in the new space.

(Of course any other position could do just as well as the "center".) Thus, at the expense of an extra action to map B to its canonical position, we can make do with 2n components instead of $n^2$.

Figure 3.3-28




Thus it is especially important to influence the schema selection process-- through goal biasing, accessibility indices, etc.-- to promote the exercise of a variety of ways to arrive at a familiar situation, so that that situation can be canonical in the above sense.

(If a compound action can run certain components concurrently, an extension of canonical-mapping is possible whereby the preferred perspective is maintained dynamically. For example, in an attempt to grasp a moving object, one branch of a compound action might continuously adjust the gaze to center on the object, while another branch keeps moving the hand towards the changing center of view.)

Just as grasp-what's-seen can develop via move-hand-into-view, a look-at-hand schema can lead to the formation of look-at-what's-touched; details are left as an exercise. Schemas to grasp what's touched and look at what's seen fulfill the second stage's coordination of sight and prehension.

### 3.3.3 Third Stage and Beyond

I haven't worked out substantive anticipations of the Schema mechanism's behavior beyond this point. This is just as well, for an extended detailed scenario would be ponderous, and progressively unconvincing: lengthy informal pronouncements on the powers of complicated unimplemented systems are greeted with suspicion, as well they should be. It is now appropriate to supplant prose with programming. However, let me wrap up the Scenario with some general aims for the Schema mechanism's conquest of later Sensorimotor stages.

The business of the third stage is to use the coordination of sight and prehension to develop a network of schemas which resemble the first stage's visual schemas-- except, these new schemas deal with synthetic items designating objects, which get organized into a (body-relative) physical space. The first-stage visual potential object items can serve as a point of departure for the third stage items. It can be discovered how hand movements alter potential-object items: if the hand is in sight and moves an object, the effect on potential-object items (via visual clues) can be learned through correlation. The look-at-what's-touched schema can lead to the proposal of tactile clues for the (formerly visual-) potential-object synthetic items, enabling phenomena like Piaget's deferred circular reactions. At this point, a variety of sensory perspectives will yield equivalent interpretations in potential-object items, paving the way for secondary-circular-reaction-type knowledge about objects to be attained through correlation.

Fourth and fifth stage developments might be expected to synthesize yet another level of potential-object items. These could express the recoverability of an object with respect to more complicated actions involving the displacement of an obstacle. The requisite compound action might develop by repair when a grasp schema was interrupted by an obstacle's intervention. Similarly, the cumulative gropings of 5th-stage inventions might come about in the repair of more primitive object-manipulating actions. The need for schemas to represent one object's position relative to another's raises some questions that I haven't delved into yet.

These vague post-second-stage remarks are intended only to be enticing, not convincing. You may wish to re-read the Sensorimotor synopsis of section 1, with a view towards imagining how observed behaviors through the six stages might fit in with a Schema-like mechanism.

"Ineffable, holy, enshrined:
so the spirit has come to be known.
But the magical imagery born of the mind
is just the same carved out of stone."
--Thomas Zimmerman

## 4. Conclusion

I claim that trying to figure out how humans work is perhaps the most promising way to invent a human-like intelligence; further, that it is well to do this by studying human intelligence from its earliest activity. Intelligence is nothing if not a Great Emulator: a new technique, mastered by a person, soon functions as smoothly and effortlessly as if it were built in. Looking at several years' accumulation of such things, one can scarcely hope to distinguish the "given" from the "gotten". Conceivably, a great deal of time could be spent investigating the nooks and crannies of a vast acquired structure, without discerning the underlying thread of acquisition. Attempting to model the *genesis* of intelligence affords, if not by itself a superior chance, at least a worthwhile complement to efforts to identify that thread by replicating later activity.

### Future Directions

I've now presented a sketch of a proposed mechanism to explore, and a rough Scenario arguing that such a mechanism might reproduce some fundamental Sensorimotor phenomena. Two paths of effort suggest themselves at this point:

* fill in the details of the mechanism, and implement it in software along with a microworld and a monitor system to examine schemas interactively; then test and revise the mechanism.
* develop a more comprehensive interpretation of key Sensorimotor developments in terms of the Schema mechanism's method of representation.

These efforts are complementary. The first starts at the bottom and works upward; the second starts with the Piagetian specification and works down towards the mechanism. The paths interact more strongly than by a prospective meeting in the middle: the partial accomplishment of each is important to the progress of the other. Experimental refinement of the mechanism would hold little chance of success without key Sensorimotor vignettes to give direction-- to guide interpretations of what the mechanism is doing, and suggest expectations of what it ought to do. Conversely, a mechanistic interpretation of Sensorimotor phenomena requires some conception of the nature of the mechanism and its data; experimentation with the mechanism can perhaps discover unanticipated ways of achieving some of these phenomena, and will certainly identify unforseen problems.

Evaluation

As discussed in the introduction, it is important that a mechanism for learning from scratch be based on a powerful theory of how humans might do that. This is true not only because such a theory is a valuable source of ideas for the mechanism, but also because it would otherwise be hard to tell what performance to desire of the mechanism. It would not be obvious that the intermediate abilities of human infants were on the right track, if not for the fact that these levels are known to culminate in indisputable intelligence; we would be similarly in the dark as to how an "infantile" artificial intelligence should progress, unless guided by the milestones of human development. But this guidance is not enough: when trying to simulate the beginnings of human intellectual development, it is important to be wary of the *scale* of the simulation.

A common failing of learn-from-scratch programs is that they propose mechanisms which may "work" in grossly constrained test situations (in which, say, only one "event" can occur at a time, making possible the association of consecutive events), but which would be helpless in a realistically complicated universe. But "learning" in a sufficiently trivial universe can be done in so many trivial ways that there is no inherent reason to expect a mechanism for such learning to bear any meaningful resemblence to what is needed to learn in an interesting world. At least three kinds of distortions result from a drastic reduction of the scale of complexity of the universe:

* combinatorics: a system which handles a situation with a trivial number of elements may suffer an explosion of time or space requirements if more elements are introduced.

* noise sensitivity: in an idealized universe, anomolous events may never occur; a mechanism unable to correct for erroneous or anomolous observations may thrive there, but be buried under accumulated garbage in a realistic world.

* abstraction: when a universe is conveniently abstracted into just the kinds of elements that form the basis for meaningful predictions, a major problem in drawing empirical conclusions is bypassed. A mechanism for learning in such a universe need not be able to disregard information on irrelevant levels of abstraction, or create new levels.

To be sure, the microworld of my Scenario is idealized and trivial; this is an unavoidable first step. Thus the Scenario, even if plausible, does not in itself offer any refutation of the possiblity of fatal scale problems in the Schema mechanism. However, the discussion of each feature of the mechanism has included consideration of how to implement the feature without excessive time/space requirements (allowing only for the the extravagance of a very large crossbar). Correlation (and other, incidental features which likewise induce assertions from emperical data) are argued to be quite correctable when they err; another kind of "noise immunity" is conferred by the ability of compound actions to adapt to and compensate for various irregularities in the world. And as for abstraction, the correlation process allows items on inappropriate levels of abstraction to be discarded along with all other irrelevant items; further, an attempt is made (via compound actions and synthetic items) to allow the mechanism to organize levels of abstraction above that given by the microworld and its sensory and motor primitives.

None of this, of course, is to argue that the Schema mechanism, as it stands, is devoid problems of scale. Rather, the point is simply that enough attention is being devoted to such matters that my approach cannot be immediately disqualified for reasons of scale.

One may be troubled by the basic thrust of Genetic A.I. for the following reason: There is a strong impression that it would be easier to endow a mechanism from the outset with primitive knowledge about objects and space than to design the mechanism to *learn* such things; why, then, should we think that evolution built humans the hard way? Perhaps observed Sensorimotor development largely reflects the physiological maturing of pre-encoded capabilities, rather than a cumulative acquisition from experience.

Piaget argues convincingly, but not conclusively, that Sensorimotor development is not preprogrammed. He does this by an extensive collection of examples showing how a wide variety of new intellectual features are plausibly interpreted as resulting from specific experiences of the infant. While there is surely room to doubt this interpretation, Genetic A.I. uses Piagetian theory only as a hypothesis to be explored, so it would be unwarranted to insist now on greater certainty. As to the intuitive argument that this would have been the wrong way to "design" human intelligence, consider the following speculation. Suppose that our species' ancestors indeed had most of their knowledge hardwired, but capabilities gradually evolved that extended the initial endowment with learned information. Probably these early capabilities were oriented towards highly specialized functions. But at some point, these capabilities reached such a degree of flexibility and generality that they could be applied to domains far removed from what they were designed for. It would then be desirable for Sensorimotor knowledge in general to be extended by the new learning mechanism, but there is no reason to expect that the early hardwired structures would have been designed to be interfaceable to the later learning mechanism. If functionally equivalent constructs could be derived in software, the original structures might atrophy as the initial convenience of preprogrammedness yielded to the power of extensibility.

This speculation is especially favorable to Genetic A.I. in that the speculation suggests that the mechanism for Sensorimotor learning plays a role in more general intelligence as well-- since, by this hypothesis, it was for compatability with more general processes that this mechanism was brought to bear on Sensorimotor knowledge. The speculation is thus consonant with Piaget's empirical claim that the processes of Sensorimotor development can indeed be identified throughout later periods.

Other reasons for reservation about Genetic A. I. are specific to the Schema mechanism's interpretation of Piagetian theory. The Schema mechanism's representation of the world violates many intuitive expectations. One tends to take for granted such things as the fact that moving to *there* implies no longer being *here*, that a rule which holds in one part of a space generalizes to other positions, etc. Customary schemes of representation are structured around notions of objects and space in such a way as to automatically embody such basic assumptions. A representation which must be struggled with to approach these assumptions seems alien and inappropriate.

But the key claim here is that the Schema representation (plausibly) *can* be struggled with successfully, by a mechanism resembling that presented in this paper. Though it may be uncustomary, it is not unreasonable to try to design a representation whose paramount feature is ease and power of automatic extensibility, even at the sacrifice of direct representational convenience. (Of course, the virtue of this is bounded by the plausibility of the mechanism proposed to do the extending; many simple representations built around associationist mechanisms are rightfully discounted.) And many of the Schema mechanism's limitations are no stranger than what is being modelled: the bugs which Piaget demonstrates in children's concepts are often bizarre and counterintuitive, and make little sense if we assume that children's internal representations resemble those which seem most obvious to us.

In compensation for the Schema mechanism's limitations, there is beginning to emerge a collection of techniques, not directly built into the mechanism, but apparently at the mechanism's disposal in its efforts to acquire new knowledge and skills. The coordination of coarse-then-fine tuning, and the canonical-mapping strategem, number among these. Hopefully a richer and more coherent set of techniques awaits discovery.

## Summary

There emerges from Piaget's theory a view of the child bootstrapping herself to intellectual competence by a long series of reconstructions of the world in progressively more abstract and less egocentric terms. Developments up to the third stage provide the first bootstrap step, which this paper's Scenario analyzes with respect to two dimensions of development: the organization of higher-level actions, and the synthesis of higher-level items. Other important dimensions are not even touched on: for example, the extension of knowledge to embrace events unrelated to personal action, or the extension of learning whereby problems can be solved by internal contemplation rather than by physical experimentation. It is my hope that further work on the theory of the two easier dimensions of development will create a framework from which the more difficult dimensions can fruitfully be explored.

I'm far less committed to the Schema mechanism itself than to the methodology of Genetic A.I. Although I believe that the Schema mechanism embodies several good ideas, some of which will survive to later versions of the theory, the current proposal can be at best an early point of departure towards a mechanism that might eventually work. Hopefully, though, the Schema mechanism lends weight to Genetic A.I. by showing that plausible and fertile ideas can arise from this approach.

## Appendix I - Related Work

I'm aware of three earlier works somewhat along the lines of this paper.

M. Cunningham [Cunningham] describes a mechanism intended to implement aspects of several psychological theories, including Piaget's. His presentation includes a hypothetical Sensorimotor scenario more extensive than the one here. Cunningham's mechanism seems fundamentally associationist-- it's based on the construction of a link between all the "active elements" of one moment, and the active elements of the next moment. Therefore I consider Cunningham's mechanism and scenario implausible. Nonetheless, it was his effort that first suggested to me the pursuit of Genetic A.I., and provided a point of departure for the Schema mechanism.

T. Jones [Jones] also presents a model of some Piagetian Sensorimotor phenomena. Jones' system makes empirical associations, but only between items which are related to each other via one of a small number of pre-supplied "patterns". Valid associations within the bounds of these relations can indeed be discovered and verified. But Jones encounters the other side of the coin of associationist combinatorics problems: in a realistic universe, only a tiny subset of related consecutive events will match the relations, and the rest will escape detection. Jones speculates briefly on how new relations might be generated-- by transitivity among old ones, or by relating variables which are "most important" by some criteria-- but I doubt that less than a combinatoric explosion of such relations would encompass the associations that need to be discovered. Jones' program INSIM1 has exhibited partial success in learning to coordinate two consecutive primitive actions in order to suck its thumb; some intermediate levels of its evolving thumb-sucking ability correspond to those described by Piaget.

J. Becker [Becker] doesn't explicitly address Piagetian theory, but his "intermediate level cognition" explores similar issues of bootstrapping (via interaction with a simple environment) from initially sparse knowledge, on a level of abstraction similar to what I call here the Piagetian level. Becker's schemas, like mine, are organized into a context, action, and result. The contents of these have more built-in structure than my "items", and Becker presents ways to generalize and differentiate schemas by comparing these structures. Unfortunately, Becker too seems to rely on an intrinsically associationist fuel for the formation and revision of schemas: in his microworld, events occur one at a time; full sequences of events are stored away and form schemas which predict that sequences which match each other so far will continue to match.

# References


[Becker]
Becker, Joseph. "A Model for the Encoding of Experiential Information" in Computer Models of Thought and Language eds. Roger Schank, Kenneth Colby. pp 396-434. Freeman (San Francisco, 1973).

[Cunningham]
Cunningham, Michael. Intelligence: Its Origins and Development. Academic Press (New York, 1972).

[Fahlman]
Fahlman, Scott. A System for Representing and Using Real-World Knowledge. Ph.D. thesis. MIT (Cambridge, 1977).

[Jones]
Jones, Thomas. A Computer Model of Simple Forms of Learning. Ph.D. thesis. MIT (Cambridge, 1970).

[Perceptrons]
Minsky, Marvin and Seymour Papert. Perceptrons. MIT Press (Cambridge, 1969).

[C.R.]
Piaget, Jean. The Construction of Reality in the Child. Ballantine (New York, 1954).

[O.I.]
Piaget, Jean. The Origins of Intelligence in Children. Norton (New York, 1952).

[P.D.I.]
Piaget, Jean. Play, Dreams, and Imitation in Childhood. Norton (New York, 1951).

[P.C.]
Piaget, Jean and Barbel Inhelder. The Psychology of the Child. Basic Books (New York, 1969).