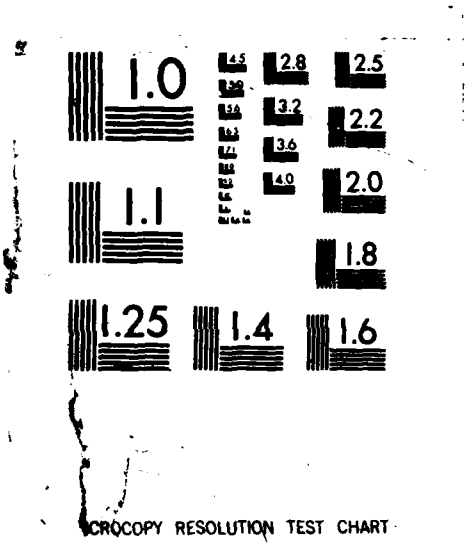
AD-A174 603 GENETIC AI: TRANSLATING PIAGET INTO LISP(U) 1/1
MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL
INTELLIGENCE LAB G L DRESCHER FEB 86 AI-M-890
UNCLASSIFIED N00014-80-C-0505 F/G 5/10 NL


1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8
CROCOPY RESOLUTION TEST CHART

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER: 890 | 2. GOVT ACCESSION NO. |
| 3. RECIPIENT'S CATALOG NUMBER | |
| 4. TITLE (and Subtitle): Genetic AI: Translating Piaget into Lisp | 5. TYPE OF REPORT & PERIOD COVERED: AI-Memo |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s): Gary L. Drescher | 8. CONTRACT OR GRANT NUMBER(s): N00014-80-C-0505 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209 | 12. REPORT DATE: Feb 1986<br>13. NUMBER OF PAGES: 16 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, VA 22217 | 15. SECURITY CLASS. (of this report): UNCLASSIFIED<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)



18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

machine learning
constructivism
Piaget

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(see reverse side)

ABSTRACT. This paper presents a constructivist model of human cognitive development during infancy. According to constructivism, the elements of mental representation— even such basic elements as the concept of physical object— are constructed afresh by each individual, rather than being innately supplied. Here I propose a (partially specified, not yet implemented) mechanism, the Schema Mechanism; this mechanism is intended to achieve a series of cognitive constructions characteristic of infants' sensorimotor-stage development, primarily as described by Piaget. In reference to Piaget's "genetic epistemology", I call this approach genetic AI— "genetic" not in the sense of genes, but in the sense of genesis: development from the point of origin.

The Schema Mechanism focuses on Piaget's concept of the activity and evolution of cognitive *schemas*. The schema is construed here as a prediction that a specified action results in a certain state of the world, provided that some other specified state holds when the action is taken. A schema is used both as an assertion about the world, and as an element of plans to achieve goals. A mechanism of *attribution* causes a schema's assertion to be extended or revised according to the observed effects of the schema's action, adding new elements to the initial (context) or final (result) state specifications. Atypically of associationist mechanisms, the attribution facility is able to sort through the combinatorial explosion of hypotheses resulting from the need to identify relevent *conjunctions* of state elements for a schema's context or result.

Crucially, the Schema Mechanism does not only construct new schemas in terms of extant actions and state elements; actions and state elements are themselves constructed by the Schema Mechanism, and then used to express further schemas. The mechanism constructs a new action to represent the process of achieving a certain goal state, abstracting above the details of how the state is achieved. Most importantly, the mechanism builds a new state element to represent a condition not yet described by other state elements, when such a condition is needed to account for certain anomolies in a schema's prediction; in effect, the mechanism constructs new ontological elements, based on existing schemas, providing new representational vocabulary for future schemas.

Included here is a sketch of the proposed Schema Mechanism, and highlights of a hypothetical scenario of the mechanism's operation. The Schema Mechanism starts with a set of sensory and motor primitives as its sole units of representation. As with the Piagetian neonate, this leads to a "solipsist" conception: the world consists of sensory impressions transformed by motor actions. My scenario suggests how the mechanism might progress from there to conceiving of objects in space— representing an object independently of how it is currently perceived, or even *whether* it is currently perceived. The details of this progression parallel the Piagetian development of object conception from the first through fifth sensorimotor stage.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 890 February, 1986

# Genetic AI
## —Translating Piaget into Lisp—

Gary L. Drescher

ABSTRACT. This paper presents a constructivist model of human cognitive development during infancy. According to constructivism, the elements of mental representation— even such basic elements as the concept of physical object— are constructed afresh by each individual, rather than being innately supplied. Here I propose a (partially specified, not yet implemented) mechanism, the Schema Mechanism; this mechanism is intended to achieve a series of cognitive constructions characteristic of infants' sensorimotor-stage development, primarily as described by Piaget. In reference to Piaget's "genetic epistemology" this approach is called genetic AI—"genetic" not in the sense of genes, but in the sense of genesis: development from the point of origin.

The Schema Mechanism focuses on Piaget's concept of the activity and evolution of cognitive *schemas*. The schema is construed here as a prediction that a specified action results in a certain state of the world, provided that some other specified state holds when the action is taken. A schema is used both as an assertion about the world, and as an element of plans to achieve goals. A mechanism of *attribution* causes a schema's assertion to be extended or revised according to the observed effects of the schema's action, adding new elements to the initial (context) or final (result) state specifications. Atypically of associationist mechanisms, the attribution facility is able to sort through the combinatorial explosion of hypotheses resulting from the need to identify relevent *conjunctions* of state elements for a schema's context or result.

Crucially, the Schema Mechanism does not only construct new schemas in terms of extant actions and state elements; actions and state elements are themselves constructed by the Schema Mechanism, and then used to express further schemas. The mechanism constructs a new action to represent the process of achieving a certain goal state, abstracting above the details of how the state is achieved. Most importantly, the mechanism builds a new state element to represent a condition not yet described by other state elements, when such a condition is needed to account for certain anomolies in a schema's prediction; in effect, the mechanism constructs new ontological elements, based on existing schemas, providing new representational vocabulary for future schemas.

Included here is a sketch of the proposed Schema Mechanism, and highlights of a hypothetical scenario of the mechanism's operation. The Schema Mechanism starts with a set of sensory and motor primitives as its sole units of representation. As with the Piagetian neonate, this leads to a "solipsist" conception: the world consists of sensory impressions transformed by motor actions. My scenario suggests how the mechanism might progress from there to conceiving of objects in space— representing an object independently of how it is currently perceived, or even *whether* it is currently perceived. The details of this progression parallel the Piagetian development of object conception from the first through fifth sensorimotor stage.

# 1 Introduction: Why Genetic AI

According to Piaget's theory of cognitive development, virtually all of our concepts are acquired. Even something as basic as the idea of physical object is constructed afresh by each individual, rather than being built in at birth. Piaget stands in opposition not only to nativism, but also to traditional empiricism: Piaget stresses the active structuring, interpretation, and organization of experience by the individual, in contrast with the more passive, "mechanical" tabulation characteristic of empiricist theories. Thus Piaget's position is called *constructivist*.

Piaget's constructivism tries to trace the development of intelligence and concepts in a typical individual, based in part on detailed, long-term observation of the ordinary behavior of infants and children. Consider an infant who sees an object, then reaches out and grasps it. This could be due to the infant's understanding that there are objects, that an object has a spatial location, that it has visual and tactile manifestations, that a certain visual pattern means object A is at position X, and that moving the hand to position X will therefore result in touching the object, which the infant desires. Alternatively, the infant might have no suspicion of the existence of objects, but might have noticed that a certain sensation, followed by a certain action, results in another particular sensation (which the infant desires). A third possibility is that the infant is just exhibiting a reflex consisting of a motor response to a visual stimulus, without specifically desiring the result of that response, without even anticipating what the result will be, indeed without even knowing that there *is* any result.

Taken in isolation, the act of grasping an object would be utterly ambiguous as to these (and other) interpretations. But Piaget's observations of the extent and limitations of infants' abilities, and the gradual elaboration of these abilities over months and years, permit reasonable inferences as to the development of underlying cognitive structures. In the present example, the Piagetian view is that all three interpretations are correct, each at a different stage of development. Mindless reflex activity yields to learned predictions that can be harnessed to pursue goals. These predictions are at first in drastically subjective form, expressed exclusively in terms of primitive perceptual inputs and motor actions. The predictions are then reformulated in gradually more objective terms of representation, terms that become progressively independent of personal action and perception.

Crucially, the earlier formulations are not merely replaced by the more sophisticated ones; often, the later structures *incorporate* earlier ones. A new concept— such as the physical object concept— forms as a synthesis of various *fragments* of the concept— such as the possibility of going from visual to tactile sensation by the appropriate hand motion, or the possibility of recovering the sight or touch of something that is not now perceived. The original fragments are bound to the details of particular perspectives and actions; the synthesis abstracts beyond these details, becoming independent of them, and also generates new detailed fragments as needed. This synthesis and abstraction, I believe, is a central theme of construcing novel concepts, concepts that transcend their precursors, that are of a fundamentally different nature than their precursors. A major goal of my research is to elucidate a mechanism for the creation of novel elements of representation.

In place of computational concepts, Piaget relied on biological metaphors (assimilation, accomodation, evolution) to think about the mind. This approach was fruitful because of the deep similarities in the design of all complicated, adaptable systems. But metaphorical descriptions do not allow precise characterization of cognitive structures, of how these structures behave, and especially, of how they change. To understand Piagetian development in detail, and to be able to verify whether underlying structures indeed develop in humans as Piaget infers, we need a less vague description of those structures and their developmental mechanism.

My effort is to submit Piaget's "genetic epistemology" to the method of artificial intelligence, which is to investigate cognitive processes by trying to engineer replicas of them; hence the name, genetic AI. My hope is to specify, and eventually implement, a mechanism (which I call the Schema Mechanism) that can recapitulate key constructions of Piagetian development in infancy, with emphasis on the development of the concept of physical objects.[1]

At present, the Schema Mechanism is unimplemented, and only partially specified. Nonetheless, I have elaborated a detailed scenario of the mechanism's anticipated behavior.[2] Pending actual implementation,

[1] As described in [Piaget52], [Piaget54], and [Piaget62].

[2] For expository convenience, the scenario is narrated herein as though its hypothetical events were real.

this anticipation remains subjective. But even at this preliminary point, the scenario is, to my knowledge, a far more concrete plausible construal of early Piagetian development than has been proposed before.[3] The Schema Mechanism, and associated scenario, are presented at length in [Drescher85], and summarized in this paper.

- Section 2 introduces Piaget's notion of the schema, and the corresponding formalism in the Schema Mechanism: basically, an assertion that a given action results in a certain state of the world, provided that certain conditions hold when the action is taken.
- Section 3 describes the microworld that the Schema Mechanism runs in, and catalogs the structures with which the Schema Mechanism is initially endowed. This section launches the hypothetical scenario, which continues through the following sections.
- Section 4 describes the *attribution* facility, which builds new schemas in terms of existing representations of states and actions.
- Section 5 introduces the facility for building a *composite action*, a subroutine for achieving a certain goal state, independently of how the state is achieved.
- Section 6 explains how the Schema Mechanism constructs new state elements, called *synthetic items*. A synthetic item is defined to represent the condition under which a certain schema is valid, when that condition seems inexpressible in terms of existing state elements. This is the Schema Mechanism's primary facility for extending its ontology, adding new elements to its representational vocabulary.
- By this point, the hypothetical scenario has recounted how the Schema Mechanism, at first supplied only with state elements and actions corresponding to sensorimotor primitives, learns of the correspondence between sensory modalities, and discovers that objects persist even when not perceived. Section 7 shows how the mechanism's first object-conception is oblivious to the possibile existence of *hidden* objects, and how this omission is corrected, all in accordance with the classic Piagetian progression.

# 2 The Schema

The *schema* is the central construct in Piagetian development. A schema is a structure that organizes some pattern of activity, such as reaching to touch something that one sees, or opening a box to retrieve its contents, or putting two sets of objects in pairwise correspondence. The evolution of schemas has several themes:

- differentiation and generalization, whereby a schema is adapted to special cases and extended to similar cases;
- coordination, whereby schemas are used to facilitate one another; and especially
- abstraction, whereby detail-bound schemas are synthesised together, transcending the details of the components.

To paraphrase Piaget in computational terms, a schema has both declarative and procedural content. Declaratively, a schema usually embodies a prediction or expectation of what will happen next, contingent on some action. Procedurally, a schema can be used as a component of a plan, to help achieve a goal. (A schema can also be exercised spontaneously, as though for practice or exploration.) These two aspects feed on each other, the procedural use appealing to a schema's predicted result, the prediction being revised according to observed effects of the schema's actual use.

A schema is formulated in terms of other representational elements that designate conditions and actions. At first, the infant is limited to subjective— in fact, virtually "solipsist"— representations corresponding to

[3] [Cunningham72] sought to explain early development in terms of an essentially associationist mechanism that seems unconvincing to me. But it was Cunningham's impressive vision of a Piagetian sensorimotor scenario that inspired my own effort.

sensory impressions and basic motor actions. Consequently, initial schemas are closely bound to personal activity. But the infant gradually constructs its own terms of representation, terms that are more and more abstract and objective; this leads to the formation of schemas which, while they retain a procedural flavor, no longer depend literally on physical action.

In the Schema Mechanism, a schema is a structure that has three main parts: a context, action, and result. A schema asserts that if its context is satisfied, taking its action is expected to bring about its result. (A schema asserts nothing about what will happen when its context is *not* satisfied.) Context and result are expressed as conjuctions of *items*. An item is a binary (On-Off) state element that represents some condition in the world; an item's state asserts whether or not that condition is now satisfied— On if so, Off if not. As an abbreviation, I often speak of the item itself being satisfied, referring really to the condition it represents; similarly, a conjunction of items is said to be satisfied when all the conjuncts are.

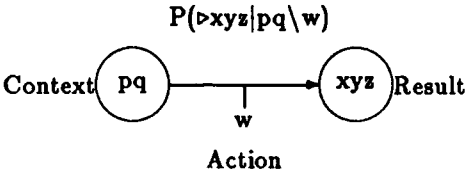



Figure 1: A schema.

Figure 1 shows a schema with items p and q in its context, action w, and items x, y, z in its result. A schema also maintains some auxiliary data, such as the schema's reliability— that is, the reliability with which the predicted result will actually follow the schema's action (provided that the context is satisfied). Reliability is measured by recording:

- $P(\triangleright R|CA)$, the conditional probability of a transition to the result state ($R$) given context conditions ($C$) and action ($A$).
- $P(\triangleright R|C\neg A)$, the conditional probability of a transition to the result under the same conditions except without the action.

Actually, of course, frequencies are recorded, not probabilities; when a sample size is large enough, the frequency is presumed to approximate the probability. $P(\triangleright R|C\backslash A)$ is notation for the pair $(P(\triangleright R|CA), P(\triangleright R|C\neg A))$. The difference between the two members of the pair is regarded as a measure of the efficacy of the schema's action in bringing about the result given the context— it says how much more likely the result is if the action is taken than if not. (The "\" symbol is read here as "with or without"; so $P(\triangleright R|C\backslash A)$ is the probability of a transition to $R$ given $C$ with or without $A$.)

The Schema Mechanism uses schemas to pursue goals. If a schema's context is satisfied and its result conjunction includes a current goal, the Schema Mechanism can achieve the goal by *activating* the schema, that is, initiating the schema's action. Or it may be that several schemas chain together, the result of each including the context items of the next, culminating at a current goal (Fig. 2). If, at some moment, the

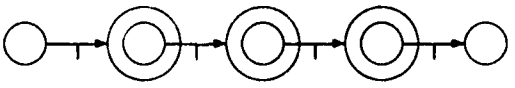


Figure 2: Schemas chain from a satisfied context to a goal.

context of any schema along the chain is satisfied, and if the schemas in the chain are all reliable, a series of activations, starting with the context-satisfied schema, will satisfy each next schema's context in turn, until the goal is reached. A schema's activation is said to succeed if the predicted goal in fact obtains, to fail if not.

The bulk of a schema is actually in two auxiliary structures, the extended context and extended result, shown in Fig. 3. The extended context and extended result each have an entry for every extant item, and for every *established conjunction* of items; an established conjunction is one that appears as the context or result

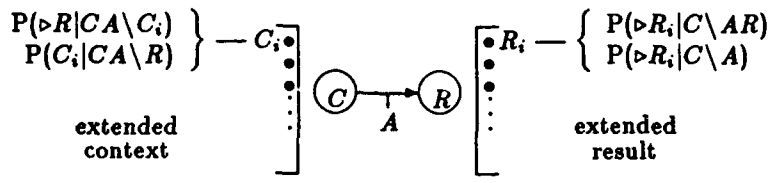


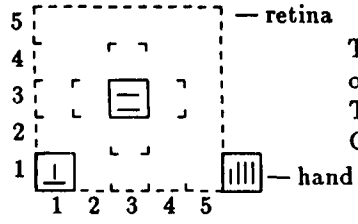

Figure 3: A schema has an extended context and extended result.

of some schema. Each extended context entry is scrutinized as to its possible relevance as an additional context condition. For each entry $C_i$, the schema continuously monitors the state of the corresponding item (or conjunction), recording the extent to which the schema's reliability is increased by the entry's satisfaction— this is measured by $P(\triangleright R|CA\backslash C_i)$— and the extent to which the schema's successful activation means the entry was more likely to have been satisfied— $P(C_i|CA\backslash R)$. This information is used, first of all, to override the schema's usual prediction in exceptional circumstances: a normally reliable schema might fail to achieve its result whenever a particular unusual condition holds; if that condition corresponds to an extended context entry, the exception will be noted at that entry. Secondly, the extended context information is used to identify conditions under which an ordinarily untrustworthy schema is more than usually reliable. A facility of *attribution* then creates a new, "spinoff" schema, like the prior schema but with the newly-identified conditions added to its context, making the new schema more reliable than the old one.

Similarly, each extended result entry $R_i$ records data as to the possibility that that entry is an additional (or alternative) result of the schema's activation, as measured by $P(\triangleright R_i|C\backslash AR)$ and $P(\triangleright R_i|C\backslash A)$. This, too, can give rise to a spinoff schema, like the schema that spawned it, but with the new item(s) included in the result.

# 3 Initial Structures

The Schema Mechanism lives in a simulated, two-dimensional microworld[4] that is intended to be crudely analogous to an infant's environment, at least with respect to the presence of various objects to interact with and learn about. The microworld includes a (simulated) robot body (including a mobile hand) that the Schema Mechanism controls. The sensory interface between the robot body and the Schema Mechanism includes tactile sensors that report the hand's contact with other objects, and a visual system with a *retina* that includes a *fovea*. The retina maps onto a region near the robot body, as shown in Fig. 4 for a typical microworld scene. Each cell in the retina grid simply reports whether or not an object appears there[5]. The




The hand is to the retina-projection's lower right, just out of its range; the rest of the body is "off-camera" below. Two other objects are present, both appearing on the retina. One is on the fovea, the other at the periphery.

Figure 4: The retina projects onto a typical microworld scene.

fovea, comprising five cells at the center of the retina (highlighted in Fig. 4), conveys more detailed visual information about any objects appearing on those cells.

[4]More correctly, a *token* of the Schema Mechanism lives there; "Schema Mechanism" is a generic term, like "internal combustion engine". But having now made this distinction, I will proceed to ignore it.

[5]These retina cells are *not* intended to correspond to individual receptors in the human retina. Rather, they are supposed to be analogous to lowlevel visual-processing output identifying objects or regions at places in the visual field. In order to gloss over the difficult details of real-world vision, the SM's microworld is designed to be so simple that trivial sensory inputs can provide the same sort of information as sophisticated real-world processing.

For each retina cell $(i, j)$, the Schema Mechanism has a corresponding primitive item[6] Ret $i, j$ which is On whenever that retina cell detects an object. For each fovea cell $(i, j)$, there is also a small collection of items Ret$_k$ $i, j$ which report the presence of each of $k$ visual features in any object appearing at that cell. These features are in lieu of real-world shape, texture, color, and so forth; in this microworld, there are just $k$ unspecified visual features whose presence or absence serve to identify an object (or kind of object). Note that the names used here to refer to items— Ret1, 3 etc— are for our convenience only. The Schema Mechanism's representation of each these items is atomic— there is no internal structure to suggest a family of related items with different parameters; nor is the Schema Mechanism supplied with any indication that these items have anything to do with vision. As far as the mechanism is concerned, these items might as well be named Ytrewq, Uiopjk, etc. The items' semantic content is to be *constructed* by the mechanism.

The retina (together with the fovea) can be mapped onto a range of different orientations relative to the body. Altering this mapping corresponds shifting one's gaze by eye movement. The Schema Mechanism is supplied with four primitive actions to change the retina-mapping: GlanceFd, GlanceBk, GlanceRt, and GlanceLt. Taking these actions shifts the visual orientation one unit forward, back, right, or left, respectively, unless this would move the mapping beyond its body-relative range; the unit of motion is the size of a retina cell. (Naturally, actions' names, like items', are meaningless to the mechanism.) For each of the, say, $5 \times 5$ possible body-relative glance orientations, there is a *proprioceptive* sensory primitive item Glance@$i, j$ that is On when the retina is in that orientation.[7]

Similarly, there are four primitive actions— HandFd, HandBk, HandRt, and HandLt— for moving the hand incrementally, within its body-relative range. There is a collection of manual proprioceptive items Hand@$i, j$ that report the hand's orientation (in the same body-relative coordinate system as the visual proprioceptive items: when, say, Hand@3, 1 and Glance@3, 1 are both On, the retina is centered on the hand). A single Touch item reports an object's contact with the front of the hand, and a small collection of items Touch$_k$ report specific (but unspecified) tactile properties, analogous to the fovea's visual properties.

Finally, for each primitive action, a "bare" schema is supplied, a schema that uses the action and that has empty context and result (Fig. 5). These schemas assert no prediction, and there is little for the Schema

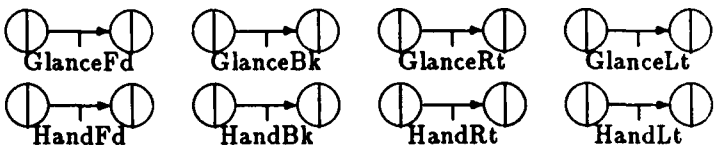



Figure 5: Initial schemas are bare.

Mechanism to do but activate them at random. Still, they are the Schema Mechanism's point of departure for exploring and understanding. In the hypothetical scenario that follows, the Schema Mechanism builds on its meager initial endowment to construct an elaborate representation of its world. New schemas are spun off from old ones, according to the attribution facility mentioned above. The mechanism also augments its representational power by creating new actions and items, from which further schemas are built.

# 4 Attribution: Building New Schemas

The Schema Mechanism's attribution facility constructs new schemas that are variants of existing ones. Information from a schema's extended result is used to identify additional results of a schema's activity (beyond any already noted in the result proper). A new schema is formed that includes the extra item(s) in its result conjunction. Typically this schema is unreliable— the additional result may be only slightly (though statistically significantly) more likely when the action is taken than when it is not. Then, the

[6] The Schema Mechanism's initially supplied items and actions are "primitive" both in the sense of being a built-in foundation, and in the sense of being atomic, unstructured, as far as the Schema Mechanism is concerned.

[7] Be careful not to confuse the proprioceptive Glance@$i, j$ items with the Ret$i, j$ items. The retina items refer to images on the retina; their coordinates specify positions on the retina. The proprioceptive items refer to where the retina is oriented, relative to the body; their coordinates specify the body-relative position where the retina is centered. These coordinates range from one to five, so (3,3) is the "central" orientation.

new schema's extended context is used to identify any items whose satisfaction contributes to the schema's reliability; often, the tentative result becomes definite provided that certain additional conditions are met.

An early application of attribution is the Schema Mechanism's organization of its visual field. Consider the bare schema whose action is GlanceLt. Occasionally this action results in some retina item turning On (because the glance action shifts some object's image to that retina cell from the adjoining cell to the left). The GlanceLt schema's extended result records this occasional response for each retina item— for instance, P(▷Ret3, 2\GlanceLt) shows that GlanceLt boosts the likelihood that Ret3, 2 will turn On. For each retina item, a new schema is spun off with that item in its result, as Fig. 6 illustrates.

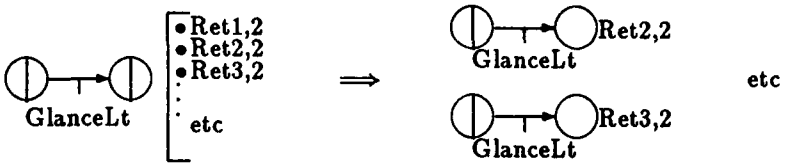



Figure 6: Schemas express visual effects of glance actions.

These schemas have low reliability. In order for, say, Ret3, 2 to turn On as a result of glancing left, there must have been an object that was not only within view, but that happened to be at Ret2, 2 just before the glance action; this happens only occasionally. Not only is the result unusual— even when it happens, it is likely to be buried among many purely coincidental, unrelated events. Still, the result is significantly more likely to occur if the glance action is taken than if it is not,[8] and this is what is measured by P(▷Ret3, 2\GlanceLt).

Due to their unreliability, these new schemas are not useful in themselves. But they are vital stepping-stones. Each new schema's extended context identifies the extra condition that makes the schema reliable— the condition that the appropriately adjoining retina item is On[9]. For example, Ret2, 2 is the condition needed for GlanceLt to reliably result in Ret3, 2; this is demonstrated by

$$P(\triangleright \text{Ret3, 2} | \text{GlanceLt} \setminus \text{Ret2, 2}) = (\text{hi, lo})$$

where $1 - hi$ is the probability that the image does not shift to Ret3, 2 despite Ret2, 2 being On (say, because the object itself is moving left, along with the glance), and *lo* is the likelihood that the image will shift there even without Ret2, 2 being On (perhaps, again, because of the object's own motion). A new schema is spun off, incorporating Ret2, 2 into its context (Fig. 7). This schema, finally, is reliable and useful.[10]

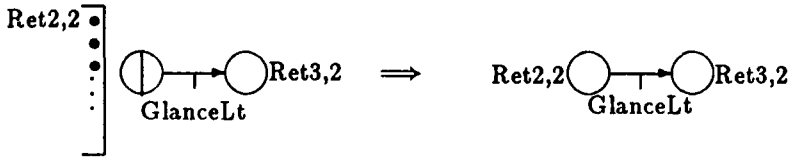


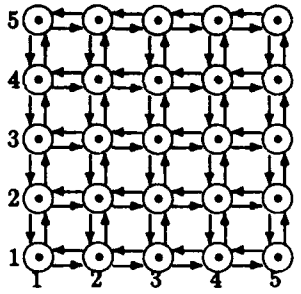

Figure 7: The schema finds a missing context condition.

The Schema Mechanism builds such schemas for all retina positions and for all four glance actions. The retina items are thereby organized into an array that matches their spatial structure (Fig. 8). The schemas chained together in this array are useful for shifting an object's image from one retina position to another.

[8] This is true even considering that an object's motion can cause the same result in the absence of any glance action. Moving objects contribute equally to the likelihood of the result whether a glance action is taken or not, while stationary objects only figure when the action is taken. Nearby objects are usually stationary, which makes the likelihood of the result several times greater if a glance action is taken.

[9] This doesn't happen though for those schemas where an object is shifted onto the *edge* of the retina from out of view, where the object had no prior sensory manifestation.

[10] If some schema's context needs the addition of several extended context entries' items in order to confer reliability, the Schema Mechanism indentifies the relevant entries individually. If, say, items p and q are required, then $P(\triangleright R|CA \setminus p)$ is (P(q|p),0), and $P(\triangleright R|CA \setminus q)$ is (P(p|q),0). Either p or q (whichever contributes more to the schema's reliability) is first added to the context of a new (still unreliable) schema; the relevance of the other condition is then noted by that schema's extended context, spinning off a reliable schema with both p and q in its context.




Each of these 80 schemas has item Ret $i, j$ as its context or result, as suggested by the spatial arrangement of the diagram. The actions are GlanceFd, GlanceBk, GlanceRt, and GlanceLt, corresponding inversely to the directions of the arrows (inversely, because shifting the retina moves an image the other way).

Figure 8: Schemas link retina items in a visual array.

An important special case is *foveation*, bringing an image from the retina's periphery to the fovea, gaining access to its visual details.

Similarly, the proprioceptive items Glance@$i, j$ get linked together in a network of schemas with the four glance-shifting actions; and the items Hand@$i, j$ are joined by schemas with the four hand-moving actions.

The schemas shown so far have at most one item in their contexts and results. Other schemas have more. For instance, some of the visual-array schemas involve an image passing from one fovea cell to another— eg, from Ret2, 3 to Ret3, 3 (as a result of glancing left). For each kind of visual detail $k$ reported by the fovea, whenever that detail is shown at (2, 3), glancing left has the additional effect of turning On the fovea item Ret$_k$3, 3. As Fig. 9 illustrates (for details a and b), these extra effects are noted in the extended results of the

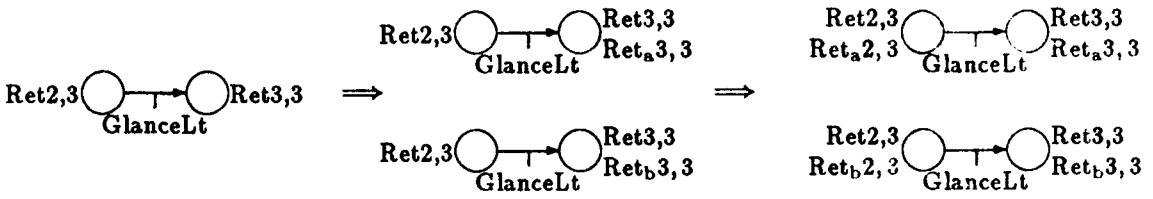



Figure 9: Schemas take note of shifting visual details.

schema in question. New schemas are formed that include the extra effects in their results. The new schemas are unreliable because each lacks a necessary additional context condition (Ret$_a$2, 3 for one schema shown, Ret$_b$2, 3 for the other). But each of these schemas' extended context notes the relevance of the missing item, which is then incorporated into the context of a new, reliable schema. Similarly of course for other visual details, fovea positions, and directions of glance-shifting.

The foveal migration of typical *conjunctions* of visual details— each conjunction comprising the visual features of a commonly-seen object, or kind of object— is captured by other, similar schemas (Fig. 10). To build each such conjunction, result items are added one at a time to a succession of spinoff schemas; for each such schema, the appropriate item is identified, then added to the context of another new schema (only the final products are shown in Fig. 10). A noteworthy related development is the discovery of visual effects of hand movement, culminating in schemas such as the one in Fig. 11.

When the hand moves forward while being watched, Touch sometimes results. A condition under which this result is reliable is that an object is seen in front of the hand prior to the hand's motion. This is expressed by the schema in Fig. 12, which is spun off from the one in Fig. 11 (via an intermediate schema, not shown, that lacks the context condition Ret3, 4). Schemas such as this confer the ability to touch an

---

If many conjuncts (or rarely-satisfied ones) need to be added, the relevance of each individually may be too small to notice. In that case, the other measure maintained by the extended context entry, P(p|$CA\backslash R$), may be helpful— if p is one of many necessary conditions, it will always be On when the schema's activation succeeds in achieving the schema's result, but only sometimes be On when the schema fails.

If there is a disjunction of many possible ways of conferring reliability, each consisting of a large conjunction of conditions, then both extended context measures may fail to detect the individual components' relevance. Even then, certain embellishments (not presented here) of the basic attribution machinery are often effective. But in the examples in this paper's scenario, only one extended context entry needs to be identified at a time to progress from an intermediate schema to a reliable one.


Ret$_{objx}i, j$ is notation for the conjunction of fovea items corresponding to the visual details of some Object X at Ret$i, j$. Similarly, Ret$_{hand}i, j$ is the conjunction of fovea items for the visual details of the hand appearing at Ret$i, j$.

Figure 10: Visual details characterize specific objects.

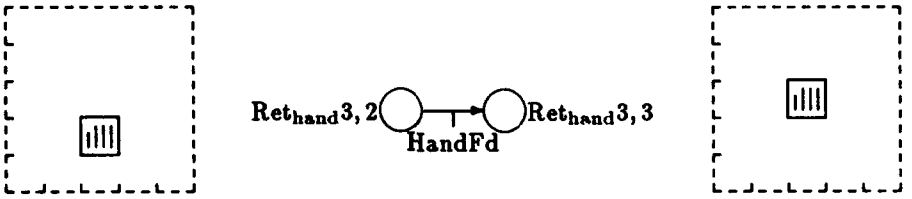



Figure 11: A schema describes the visual effects of hand motion.

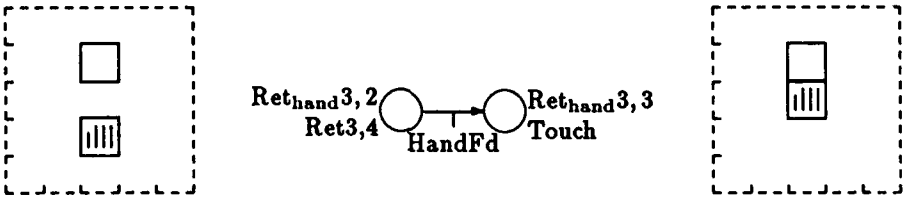



Figure 12: This schema knows how to touch an object seen close to the hand.

object that is seen close to the hand. In actual Piagetian development, this is a significant precursor of the broader ability to touch a nearby visible object regardless of whether the hand is also in view. In general, the schemas discussed so far typify the abilities, and limitations, of infants in early (Piaget's first and second) stages of development.

The attribution facility builds schemas to propose and verify hypotheses about the effects of actions. The goal is chiefly to construct reliable schemas, not probabilistic ones. But attention to small (but significant) differences in probabilities is a way to avert a combinatorial explosion in the search for reliable schemas. If n items exist, then for a given action, there are $2^{2n}$ expressible schemas using that action ($2^n$ possible context conjunctions times $2^n$ possible result conjunctions); even if we only consider, say, conjunctions of 5 items or less, there are still about $n^{10}$ possibilities. If n is 1,000,000— or even 1,000— it is impossible to build, and simultaneously monitor, all of these schemas to verify which ones are reliable. With the attribution facility, this enormous space of hypotheses is searched by verifying the relevance of each context or result element individually, building the required conjunctions incrementally; but the relevance of individual conjuncts is only subtly, probabilistically manifested (eg, recall Fig. 6). Thanks to the ability to detect such manifestations, each schema needs to monitor only n possibilities at once, not $2^{2n}$ or $n^{10}$; so even if the number of extant schemas is on the same order as the number of items, only $n^2$ worth of brute force is required. This is still substantial, but within the realm of possibility.

# 5 Composite Actions

Schemas' actions play two roles in the Schema Mechanism. With respect to planning, an action serves as something that the Schema Mechanism can cause to happen in order to move towards a goal. With respect to learning, an action is something whose effects are looked for, via the attribution facility. These roles support each other: being able to take an action at will aids the investigation of its results, enabling a kind

of deliberate experiment; and knowing the results of an action is particularly *useful* when the action is (at least partly) under control, and can be taken, or avoided, according to the desirability of the results.

In the Schema Mechanism, every newly-achievable result (ie, every novel result conjunction of a newly created schema) defines a new *composite action*— the action of achieving that result. For example, for every proprioceptive item Hand@$i, j$, there are several schemas (from the manual proprioceptive array discussed above) with that item as their result— schemas that move the hand to that position by shifting the hand from an adjoining position. The first time, say, Hand@2, 1 appears as the result of some new schema, a new action Hand@2, 1 is defined. As with the original, primitive actions, a new, bare schema that uses the action is also created (Fig. 13). And of course, similar actions and schemas are created for each of the other hand positions, designated by the other items Hand@$i, j$.

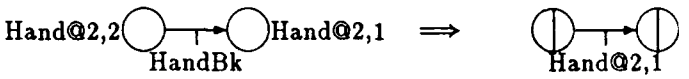



Figure 13: A new result gives rise to a new composite action.

Initiating a *primitive* action triggers some motor effector under the Schema Mechanism's control (eg one that shifts the hand or the glance). When a composite action is initiated, a chain of schemas is identified leading from some currently-satisfied context to the composite action's *goal state*— in the present example, Hand@2, 1. (If no such chain is found, the action cannot be taken on that occasion.) The schemas of the chain are activated in succession to achieve the action's goal state. For example, the action of satisfying Hand@2, 1 is achievable by the Schema Mechanism via the schema at left in Fig. 13; and by other schemas for moving the hand there from other adjoining positions; and via the many other proprioceptive-array schemas that chain to the ones that achieve Hand@2, 1 directly.

One way to identify a chain of schemas to an action's goal is to *broadcast* an inquiry from the goal conjunction, to all schemas whose results include the goal, to all schemas whose results include any of *those* schemas' contexts, and so on until a schema is found whose context is satisfied. Although the Schema Mechanism's architecture allows this broadcast to proceed in parallel through a tree of chained schemas, it is advantageous to "compile" the results of prior broadcasts for quicker future reference. The information is stored in each action's *arbiter*. The arbiter remembers which schemas lie along chains to the action's goal, and how "close" each link is to the goal; and the arbiter always monitors (simultaneously) which schemas now have satisfied contexts. The arbiter functions as a special-purpose machine for achieving the action's goal; when a composite action is initiated, its arbiter activates the closest schema to the goal (among those reliable schemas whose contexts are now satisfied), repeating this until the goal is reached. Occasional new broadcasts update the arbiter's information.

Designating an event as an action in its own right is a way of abstracting above the details of how to achieve that event. One consequence is that a goal can be *invoked* at a higher level of representation-- just initiate the action designating that goal, and the details follow automatically. A deeper but subtler consequence is that the *effects* of the goal event can also be expressed with respect to the more abstract representation, as the following examples illustrate.

Consider the bare schema with action Hand@2, 1, shown in Fig. 13. Sometimes this schema's action results in the hand being brought into view, say at Ret3, 2. This result is independent of the hand's orientation prior to the action, but it does depend on the glance orientation happening to be just right— $\text{Ret}_{\text{hand}}3, 2$ results from Hand@2, 1 only if Glance@2, 2 is satisfied.[11] The tentative result, and then the necessary condition, are identified in succession by attribution, finally yielding the schema of Fig. 14. Similar schemas prescribe how to bring the hand into view for other glance orientations, or at other places on the retina.

Bringing the hand into view, say at Ret3, 2, is itself an achievable result, which gives rise to a composite action; and a schema for bringing the hand into view chains to a schema for touching an object that's beside the hand to produce the ability to touch whatever is seen nearby, whether close to hand the hand or not (Fig. 15).[12] Finally, a similar seqence of developments constructs schemas that chain between sensory modalities in the other direction, enabling the Schema Mechanism to turn its gaze to what it touches,

[11] Recall that the same body-relative coordinate system is used in the names of the visual and manual proprioceptive items;

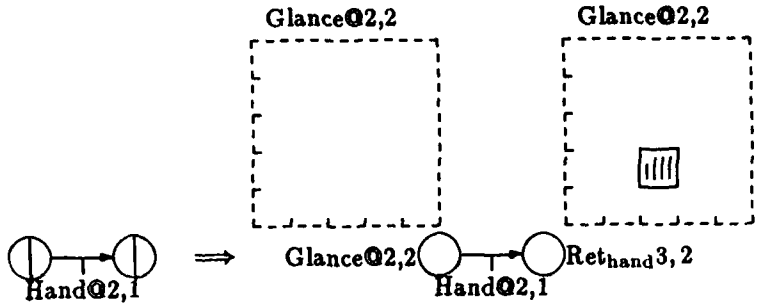



Figure 14: The unseen hand moves into view.

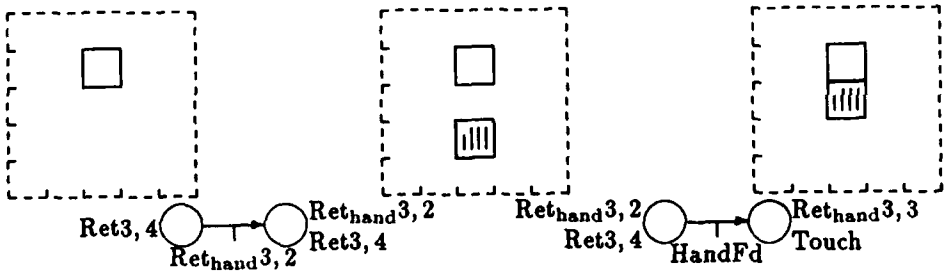



Figure 15: The hand reaches for what's seen.

expecting to see something. Note that there is not (yet) any representation of the fact that sight and touch are *two manifestations of the same thing*; the mechanism just knows now how to go from one to the other. But this is a start.

Notice that the (context-dependent) result of Hand@2, 1— seeing the hand— could instead be represented as a (further) result of an action, such as HandLt, that the mechanism uses to obtain Hand@2, 1. But this representation would be inadequate:

- If seeing the hand when it's at (2,1) is predicted as a result of, say, HandLt, that prediction doesn't generalize to situations where Hand@2, 1 is brought about by, say HandFd. The effect of each way of causing Hand@2, 1 would have to be learned separately. The more ways there are to cause an event, the more important it is to look for results of the event itself, not just of particular ways of causing it.

- Hand@2, 1, and other achievable results, are often caused externally, rather than by an action of the Schema Mechanism. For purposes of gathering extended context and result data, the attribution facility regards a composite action as having been taken whenever the action's goal is satisfied, regardless of the cause. Thus, by using composite actions, the Schema Mechanism can learn about the results of external events as well as its own actions. (Examples of this appear in sections 6 and 7.)

What is truly relevant to seeing the hand is where the hand is put, not how it gets there. Building a composite action lets an event be represented at the relevant level of description, which makes it possible to express, discover, and choose to pursue its effects at that level.

# 6 Synthetic Items and Conservation

Much interesting, practical knowledge can be expressed in terms of objects' sensory manifestations, but much more cannot. Very little of the world is within range of our senses at a given moment; it is important that we can know about things we do not now perceive. Even for something we do see, it is often important

and when the glance is at orientation $(i, j)$, an object at $(i, j)$ appears at the center of the retina, Ret3, 3.

[12]More straightforwardly, the hand might be placed beside the object directly. But if there is imprecision in that placement, the hand will often wind up merely close by; the next step in the chain then achieves actual contact.

to represent it at a level of description that abstracts beyond its present sensory manifestation, so we can formulate assertions that apply to it regardless of our current sensory perspective.

This section describes two mechanisms for representing something independently of its sensory manifestation. The first is an important, but limited, technique for identifying what is preserved across a pair of inverse actions. The second is a more far-reaching facility for constructing new elements of representation, *synthetic items*, to surpass the sensory-bound primitive items.

The Schema Mechanism recognizes a pair of actions as mutual inverses when there is some item (or established conjunction of items) which, if it starts out being On, is turned Off, then On again, when the two actions are taken in succession. (I won't describe here just how this recognition is implemented; it involves comparing probability measures in the spirit of the attribution facility.) For example, HandBk and HandFd are mutual inverses, because if, say, Hand@3, 2 is On, it is turned Off by HandBk, then back On by HandFd.

More significantly, HandBk and HandFd are inverses with respect to Touch— if the hand is touching something and moves back from it, then forward again, the object is again in contact. A new schema is formed to express this recoverability (Fig. 16). This is significant because Touch's recovery could not have

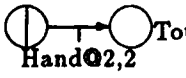

Figure 16: The hand moves back from an object, then regains contact.

been (reliably) predicted on the basis of HandFd alone— HandFd does not usually result in Touch, but does when it directly follows HandBk, provided that Touch was On prior to moving the hand back. In contrast, the recovery of Hand@3, 2 is predictable on the basis of HandFd, together with the Hand@3, 1; the previous HandBk need not be taken into account. The recovery of Touch might be similarly predicted, ignoring the previous HandBk, *if*, say, the object is in view in front of the hand. But this discussion addresses the general case where the object and the hand are not necessarily being watched; hence, after the hand is moved back, there may remain no sensory manifestation of the object by which to predict the recovery of the original perception.

This anticipation of perceiving something again after it disappears is a rudimentary form of *conservation*, knowing that something is "still there" even when it is not perceived. Visual conservation, as well as haptic, is promoted by inverse-action recognition— schemas analogous to the one in Fig. 16 assert that an image reappears when the glance is shifted away, then back. The ability to recover objects in this manner is observed very early (first Piagetian stage) in infants' development.

But if conservation across a pair of inverse actions goes beyond immediate sensory data, it does not go very far beyond. Unless the recovering action immediately follows its inverse (so that the hand or eye orientation is likely still to be adjacent to the original position), there is no basis for predicting recovery. Eventually (Piaget's third stage), infants exhibit a more general kind of conservation: an infant may explore an object visually or haptically, turn away for a few minutes for some other activity, then deliberately return to the object at its original position. This requires knowing *where* to return to the object, rather than just reversing an incremental departure. In the Schema Mechanism, synthetic items make this possible.

Consider the bare schema with composite action Hand@2, 2. Unreliably, but noticeably to the attribution facility, this action results in Touch; this is expressed by the schema in Fig. 17. This schema is reliable just




characteristic item: HapticObj@2,3

Figure 17: This item asserts roughly that an object can be touched at (2,3).

in case there is an object at (2,3); that condition, of course, is unknown to the Schema Mechanism, which so far cannot even *represent* the assertion that an object is there, let alone knowing what that entails.

But this schema furnishes an important clue. Despite its unreliabity, the schema is *locally consistent*: if it successfully achieves its result, it is likely to succeed again if it is activated in the next little while, and if it fails, it is likely to fail again. That is because nearby objects are usually stationary. The Schema Mechanism, of course, is not equipped to appreciate this reason; but each schema does determine, by simple empirical tabulation, whether its own success is locally consistent.

If an unreliable schema is found to be locally consistent— and if no context conditions can be identified that reliably distinguish the successful situations from the unsuccessful ones— it is plausible that some condition *not represented by extant items* determines the schema's success or failure. In the present example, there is (often) no sensory manifestation of the presence (or absence) of an object at (2,3) prior to moving the hand there; hence, there are indeed no primitive, sensory items whose state can be relied on to predict whether Touch will result.

Under these circumstances, the Schema Mechanism constructs a new synthetic item. (It is called the *characteristic item* of the schema for which it was built, which is called the item's *host schema*.) The synthetic item represents *whatever mystery condition* governs the success or failure of its locally consistent host schema; the item might be said to *reify* the schema's conditions of validity. In this example, let's call the new item HapticObj@2, 3. (As always, the item's name has significance only to us, not to the Schema Mechanism.) Of course, such haptic object items develop for other positions as well. And analogously, an item VisualObj@$i, j$ is spawned from each schema whose action is Glance@$i, j$, and whose occasional result is Ret3, 3 (seeing something at retina center[13]).

A synthetic item, like a primitive one, is supposed to be On if, and only if, the condition it represents is satisfied. A primitive item's state is maintained according to input from some sensory device directly "wired" to the item. Maintaining the state of a synthetic item is more elaborate; the Schema Mechanism turns a synthetic item On or Off according to several kinds of evidence as to whether the item's condition is currently met.

The first piece of evidence is whether or not the most recent activation of the item's host schema was successful. When a successful activation occurs, a schema turns its characteristic item On; when there is a failure, the item is turned Off. If the host schema usually fails, its characteristic item automatically reverts to the Off state if the schema has not been activated for a period of time; the time required is based on empirical determination of the expected duration of the schema's temporary reliability. A schema is presumed reliable while its characteristic item is On, unreliable when it is Off.

Recall that a composite action is regarded as having been implicitly taken whenever its goal state is satisfied. If a schema's context is satisfied when its action is implicitly taken, the schema is regarded as having been implicitly activated. Based on the implicit activation's success or failure in obtaining the schema's result, the schema's characteristic item is turned On or Off, just as with explicit activation. In the example above, HapticObj@2, 3's host schema's context is empty, hence always trivially satisfied. That schema's action is regarded as having been taken whenever Hand@2, 2 is On, and so the schema has in that case been implicitly activated. The activation is successful or not according to whether the schema's result, Touch, obtains. So whenever the hand is at (2,2), HapticObj@2, 3 is turned On if something is touched, otherwise Off.

An interesting property of the visual- and haptic-object synthetic items is that they are not, at first, connected to one another (just as visual and tactile primitives had been unconnected until intermodal coordination was constructed). Touching something at (3,2) turns On HapticObj@3, 2, but not VisualObj@3, 2; looking directly at an object at (3,2) turns on the visual item, but not the haptic one. This is encouragingly consistent with infants' initial behavior towards objects' permanence: an infant who, a while ago, had held a toy without looking at it, will reach back for it but not try to *see* it (until after touching it[14]); an infant who had looked at a toy, but not held it, will not *reach* for it before seeing it again. At the next stage of development, though, visual and haptic persistance become interchangible.

This interchangibility develops in the Schema Mechanism by means of a second kind of evidence as to a synthetic item's state. A synthetic item's host schema's extended context identifies items (or established conjunctions) that represent *manifestations* of the condition represented by the synthetic item. A condition's manifestation is a state that implies that the condition now holds. For example, HapticObj@3, 2 is a

[13] Items can develop with respect to off-center views, too, but the ones I'm emphasising are the most interesting.

[14] Once the toy has been touched, the earlier intermodal schemas for looking at what's touched are applicable.

manifestation of VisualObj@3, 2, and vice versa. Identifying manifestations is almost identical to identifying new context candidates by attribution: an extended context entry is deemed to represent a manifestation of the schema's characteristic item if the ordinarily unreliable schema is likely to succeed whenever that entry has been satisfied recently— "recency" depending, again, on the empirically determined expected duration of the schema's temporary reliability. When any condition that manifests a synthetic item is found to be satisfied, the synthetic item is turned On. Likewise, manifestations of a schema's invalidity— conditions whose recent satisfaction implies that the schema will *fail*— are identified, and used to turn a synthetic item Off.

Once a haptic-object item and the corresponding visual-object item (and their negations) have been identified by their respective host schemas as mutual manifestations, they become effectively synonomous, always in the same state as one another. I will refer to Obj@$i, j$ to designate the conjunction of HapticObj@$i, j$ and VisualObj@$i, j$ once the Schema Mechanism has established their synonymity.

A synthetic item, like a primitive one, can be incorporated by attribution into the contexts and results of other schemas. For example, the schema in Fig. 18 shows how to shift Obj@3, 2 to an adjoining position by grasping it and moving the hand.[15] Of course, this is only one of many such schemas; others pertain to


Figure 18: The hand moves an object: a sensory-independent representation.

different positions, or different directions of hand motion. These schemas elaborate the spatial structure of the Obj@$i, j$ items, just as the retina and proprioceptive arrays were elaborated in earlier development. And the Schema Mechanism exploits such schemas as a third way of maintaining the state of synthetic items: if a reliable schema is activated, and that schema's result contains a synthetic item (or its negation), the Schema Mechanism then turns the item On(or Off), in accordance with the schema's prediction.

# 7 Recovering Hidden Objects

The Schema Mechanism has now come a long way on its hypothetical journey. Originally, the only extant elements of representation were designations of immediate sensory inputs; and even those elements were not yet *understood* to mean anything whatsoever. By now, the spatial structure and intermodal coordination of the sensory elements is represented by the Schema Mechanism, in terms of practical schemas joining those elements via actions. And the mechanism has built new elements to represent the presence of an object at a certain position, independently of how, or whether, the object is currently perceived.

But this progress should not be exaggerated. The mechanism's concept of a permanent object is far from complete. This is illustrated by what happens when an object is *hidden*.

Certain objects in the Schema Mechanism's microworld are *visual obstacles*. When an object sits directly in front of a visual obstacle, that object cannot be seen— the retina does not detect it. For simplicity, let's say that all visual obstacles are identifiable by their characteristic visual details; and Ret$_{obstacle}i, j$ designates the conjunction of fovea items that identifies a visual obstacle object at Ret $i, j$. Figure 19 shows a situation where an object is hidden by an obstacle.

Note that in Fig. 19, the retina is centered on the (unseen) object; say the object is at (2,2), so Glance@2, 2 is On. But if Glance@2, 2 is On and Ret3, 3 is Off, then Obj@2,2 is turned Off: the Schema Mechanism— like an infant early in the acquisition of object conservation— reacts as though the hidden object no longer exists! When an object is unseen despite looking directly at it, the Schema Mechanism, at this point, cannot even *express* the object's continued existence there— there are no extant items whose state reflects that existence.[16] This, of course, is reminiscent of the earlier inability to express *any* unperceived object's

[15] I've neglected to mention before now that the Schema Mechanism includes a Grasp action that engages grasping for a while (or until the Ungrasp action is taken). The proprioceptive item Grasping reports when grasping is engaged. A grasped object moves when the hand does.

[16] I am overlooking the case where the hidden object is in contact with the hand.

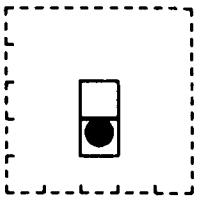


Figure 19: The object at retina center is hidden from view by the visual obstacle behind it.

continued existence. The remedy, too, is similar: the Schema Mechanism constructs synthetic items to represent hidden objects. The details in this case are a bit cumbersome, but here is the general idea:

- A precursor step is to note the object's recovery across the pair of inverse actions consisting of placing the obstacle to hide the object, then displacing the obstacle. As with the earlier examples of inverse-action conservation, this is limited to disappearance and recovery caused by a specific pair of actions, and to situations where the recovery immediately follows the disappearance.

- Consider a schema whose action is to move the hand in the context of grasping a visual obstacle that's at, say, position (2,2) (Fig. 20[17]); the action thus displaces the obstacle from that position. The schema's result is the appearance of Obj@2,3. This result obtains just in case an object happened to be

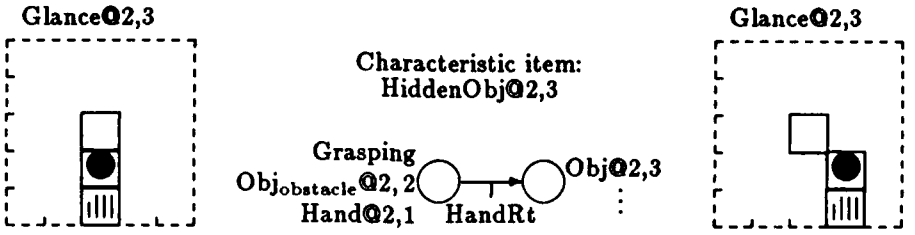



Figure 20: A visual obstacle is moved, revealing an object it had hidden.

sitting, hidden, at (2,3). The result is unreliable, but locally consistent: objects are usually stationary, so the object at (2,3) is likely to stay put for a while; so if this unhiding schema is tried again soon after succeeding once,[18] the schema is likely to succeed again. Thus, the schema gives rise to a synthetic item, which we can call HiddenObj@2,3.[19] When the unhiding schema successfully obtains its result, Obj@2,3, HiddenObj@2,3 is turned On. If the obstacle is soon replaced, hiding the object again, the Schema Mechanism can activate the unhiding schema again, with the anticipation that the schema's result, Obj@2,3, will obtain— because the schema's characteristic item is now On, which asserts that the schema is believed reliable at the moment.

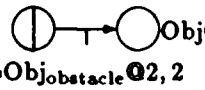
- Here is a fascinating subtlety. Suppose the above unhiding schema is activated successfully, recovering an object at (2,3), turning On the item HiddenObj@2,3. Suppose the object is then moved *away* from (2,3), and then the obstacle is replaced at (2,2). If the object was moved away in direct view of the retina, then the Schema Mechanism properly turns Obj@2,3 Off. But notice: nothing turns HiddenObj@2,3 Off! Thus, the unhiding schema continues, spuriously, to be deemed reliable. If, say, the object is now hidden again, at a different location— and if the object's presence is a current goal of the Schema Mechanism— the obstacle-displacement schema may be (wrongly) trusted to pursue this goal: the obstacle at (2,2) will be removed again in order to recover the object that is no longer there! Bizarre as it may seem, infants beginning to learn about hidden objects routinely exhibit just this behavior.

[17] The synthetic item $Obj_{obstacle}@2,2$ (in Fig. 20) is to Obj@2,2 as $Ret_{obstacle}i,j$ is to Reti,j: $Obj_{obstacle}@2,2$ represents the assertion that if the glance is directed to (2,2), $Ret_{obstacle}3,3$ will result— thus, that there is a visual obstacle at (2,2).

[18] Of course, this schema can be tried again only if the obstacle is replaced in the meantime.

[19] This is a slight misnomer, since the item says On even when the object is uncovered; the item really denotes a *possibly*-hidden object.

There is an alternative formulation of hidden-object recovery that does not succumb to this mistake. In Fig. 21, the obstacle's displacement from position (2,2) is expressed less subjectively, not in terms of a hand movement, but in terms of a composite action whose goal state is the obstacle's absence. (In Fig. 21, $\neg \text{Obj}_{\text{obstacle}}@2,2$ is, of course, the negation of $\text{Obj}_{\text{obstacle}}@2,2$, asserting that there is no such object there.) The schema in Fig. 21, like the previous hidden object schema, is unreliable, but locally consistent; and




characteristic item: BetterHiddenObj@2,3

Figure 21: Here, unhiding is attributed to the action of displacing the obstacle.

when it is, temporarily, reliable, it too prescribes how to unhide an object at (2,3). Let's call this schema's characteristic item BetterHiddenObj@2,3.

This schema's alternative, less subjective formulation of the obstacle-displacement action has a surprising advantage. Whenever there is no obstacle at (2,2), this new schema has (at least implicitly) just been activated (since its action's goal state is then satisfied, and the schema's empty context is always satisfied). Thus, whenever there is no obstacle at (2,2), BetterHiddenObj@2,3 is turned On if Obj@2,3 is On, Off if Obj@2,3 is Off. Thus, in the situation just discussed, when the unhidden object is moved away from (2,3), BetterHiddenObj@2,3 (unlike HiddenObj@2,3) turns Off along with Obj@2,3. This still doesn't quite solve the problem— it remains to correct the flawed original unhiding schema, which otherwise continues to assert its faulty expectation of recovering the object at its original position.[20] But once the item BetterHiddenObj@2,3 exists, the extended context of the original unhiding schema discovers that that item is an additional condition that must be met if the original schema is to succeed. This progress is consistent with the behavior of infants at Piaget's fifth stage of sensorimotor development.

# 8 Conclusion

For brevity, I have omitted much here. I've discussed features of the Schema Mechanism in somewhat simplified form, ignoring several technical problems and fixes. And I have neglected to discuss several important facilities; for example, the Schema Mechanism's way of assesing the *value* of achievable states, which helps determine what goals it pursues and what structures it builds; an embellishment of the attribution facility, whereby result and context conditions are sometimes discovered in just one trial; and the facility of *subactivation*, which uses extant schemas to run a "simulation" of a contemplated sequence of actions— subactivation is a "thought experiment" from which lessons can be extracted, by attribution and by other facilities, just as from actual experience.

In addition, the hypothetical scenario is much compressed from the version in [Drescher85]. One symptom of this is the introduction here of attribution in connection with Piagetian first-stage development, composite actions at stage two, and synthetic items at the third stage. It is important to note that this sequence is just an expository device; in the extended version of the scenario, all facets of the mechanism are productive almost immediately. Thus, I do not propose that the structure of Piagetian stages just mirrors the architecture of the developmental mechanism, each stage corresponding to another piece of the machinery. On the contrary, the features of the mechanism all pertain quite generally to the knowledge and abilities of many different stages. What the progression of stages reflects is the dependency of some constructs on others; the structures that come later are those that need to incorporate earlier ones, or that depend on knowledge gained through the activity of earlier ones.

The Schema Mechanism owes an obvious, enormous debt to Piaget's theory, and offers a bit of repayment. For one thing, as mentioned in the introduction, any plausible concrete interpretation of Piaget's

[20] In fact, it is only at this point— when *both* forms of hidden-object representation are functioning— that the Schema Mechanism fully mirrors the behavior of an infant making a place error. The more objective representation is needed to inform the mechanism that the hidden object is recoverable in the first place, before any attempt has been made to remove the obstacle; only after the recovery has taken place does the subjective representation assert continued recoverability. Note, too, that separate hidden-object representations are needed for distinct positions, and these need not develop simultaneously; recoverablility may be learned precociously at familiar positions, tardily at others.

constructivist mechanism aids the discussion and development of constructivism. But there is another basic contribution. In recent years, constructivism has been challenged by new evidence about neonates' knowledge— knowledge that is said to be inexplicably precocious in Piagetian terms.[21] Usually, though, this knowledge is exhibited in tightly circumscribed ways, not apparently amenable to extension and adaptation; this suggests that we may be seeing an amalgam of special-purpose reflexes, internal details of peripheral modules, and even evolutionary vestiges of pre-constructivist intelligence, whose existence is compatible with the (re)construction of basic concepts by a central, general, Piagetian mechanism. But just pointing out that the evidence against constructivism is inconclusive is hardly a compelling defense. What would be compelling is to show that there is a *reason* for human intelligence to be designed to construct its basic concepts in accordance with Piaget's progression: to exhibit a reasonably engineered general learning mechanism which, starting with no concept of objects, does develop such a concept, and, crucially, needs to pass through intermediate stages that exhibit just the ablities, limitations, and mistakes that characterize the steps of the Piagetian sequence. Of course, however, such a mechanism has yet to be built; and until it is, I have offered no objective evidence, just a plausibility argument as to what evidence we might expect.

Inventing a new concept, one that is very different from its precursors, is arguably the highest manifestation of intelligence. We should therefore be very impressed at the conceptual inventions of infants, if, as Piaget argues, they come equipped with only sensory representations, and build the rest themselves. The sort of thing that an object is is really *nothing like* the sort of thing that a sensory impression is— despite the fact that direct perception is among the kinds of evidence by which we know an object. Even the rudimentary approximation to objects that the Schema Mechanism develops in its hypothetical scenario is far removed from a mere sense datum. It is encouraging to have even a sketch of a mechanism that might achieve such invention, and doubly encouraging if it is shown to do so in accordance with observed human development.

## Acknowledgements

This is a revision of a paper presented at a conference of the Society for the Study of Artificial Intelligence and the Simulation of Behavior (AISB) in Exeter, England in August, 1985. I am grateful to Bob Lawler for instigating my writing this paper. Phil Agre and Ken Haase contributed helpful suggestions. All errors, large or small, are mine.

## References

[Bower74] Bower, T.G.R. (1974) *Development in Infancy*. San Francisco: W.H. Freeman.

[Bower77] Bower, T.G.R. (1977) *A Primer of Infant Development*. San Francisco: W.H. Freeman.

[Cunningham72] Cunningham, M. (1972) *Intelligence: Its Origins and Development*. New York: Academic Press.

[Drescher85] Drescher, G. (1985) *The Schema Mechanism: A Conception of Constructivist Intelligence*. MS thesis, Cambridge: MIT.

[Flavell85] Flavell, J. (1985) *Cognitive Development*. (2nd edition) Englewood Cliffs, N.J.: Prentice-Hall.

[Piaget52] Piaget, J. (1952) *The Origins of Intelligence in Children*. New York: Norton.

[Piaget54] Piaget, J. (1954) *The Construction of Reality in the Child*. New York: Ballentine.

[Piaget62] Piaget, J. (1962) *Play, Dreams, and Imitiation in Childhood*. New York: Norton.

[21] See, for example, [Bower74], [Bower77], and [Flavell85].

END

1-87

DTIC