

A Mechanism for Early Piagetian Learning

Gary L. Drescher
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Ma 02139

Abstract

I propose a mechanism to model aspects of Piagetian development in infants. The mechanism combines a powerful empirical learning technique with an unusual facility for constructing novel elements of representation—elements designating states that are not mere logical combinations of other represented states. I sketch how this mechanism might recapitulate the infant's gradual recognition that there exist physical objects that persist even when the infant does not perceive them. I also report results of a preliminary, partial implementation.¹

I. Statement of the Problem

According to Piaget's *constructivist* theory of mind, 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 [Piaget 1952,1954]. At first, the infant's conception of the world is virtually solipsist: the infant represents the world only in terms that correspond to basic sensory impressions and motor actions. As the infant interacts with the world, it learns that some actions affect some sensations. But the infant does not understand that there are objects "out there", objects that its actions affect, that can be perceived by sight or touch, and that persist even when not perceived.

Crucially, the infant later transcends this limitation, inventing for itself the idea of physical object, constructing new terms of representation to augment the innate sensorimotor ones. The infant constructs the concept gradually, in stages; along the way, intermediate representations become less subjective, less tied to the infant's own perspective and activity. Progression from subjective to objective or abstract representations is a central theme of Piagetian development; the physical-object concept provides an early, paradigmatic example.

¹This research was done at the MIT AI Laboratory. This work is nonmilitary, but my use of laboratory computers obliges me to state that the laboratory's AI research has been supported in part by the Advanced Research Projects Agency of the Department of Defense under ONR contract N00014-85-K-0124. This does not imply my approval of the United States policy of terrorizing civilians to impose repressive regimes for US strategic or economic advantage.

Piaget supplies elaborate observations of characteristic behaviors at each developmental stage as reflections of the infant's underlying representations of the world. But Piaget stops short of explaining what mechanism underpins the development he describes; that is the goal of my present effort. I take Piagetian development as a working hypothesis; trying to implement it is a way to test and refine the hypothesis.

II. Overview of the Theory

The Schema Mechanism controls a body in a microworld. Based on interaction with the world, the mechanism accumulates and organizes knowledge, and uses this knowledge to guide its activity in pursuit of goals, foremost of which is the acquisition of further knowledge.

The Schema Mechanism's data structures comprise *items*, which are binary state elements; *actions*, which can produce state transitions; and *schemas*, which designate effects of actions on states. Primitively supplied items and actions designate simple sensory inputs and motor outputs; nonprimitive items and actions, autonomously constructed by the mechanism, designate states and transitions at higher levels of abstraction. Schemas are defined in terms of actions and items; nonprimitive actions and items are defined in terms of schemas, and also serve as elements of subsequently made schemas.

Starting from a small primitive base, the Schema Mechanism constructs *spinoff schemas*, as it discovers relations among actions and items; *composite actions*, each of which is defined by a *goal state*, and implemented by schemas that coordinate to achieve the goal state; and, most importantly, *synthetic items*, each of which designates a state that is novel in that it may be inexpressible as any function of the current state of other extant items. The facility for constructing new schemas is noteworthy in that it has due regard for the combinatoric problem of discovering regularities when many things typically happen at once, mostly coincidentally. Synthetic items permit the invention of radically new concepts; they facilitate Piagetian *conservation*, wherein an individual realizes that there is something that stays constant even though its manifestation (in terms of any concepts previously at the individual's disposal) changes or ceases.

A schema has a context, action, and result. The con-

text is a (possibly empty) set of items, as is the result. A schema asserts that if its context is satisfied— if the designated states obtain— then taking the action makes the result *more likely* to obtain than if the action weren't taken. A *reliable* schema asserts, further, that the action makes the result likely (not just *more likely*); schemas keep track empirically of their reliability. Only reliable schemas can serve as elements of a "plan" (a set of schemas coordinated to achieve a goal); unreliable schemas are stepping stones to finding reliable ones, as described below. Note that a schema's context is not a precondition for taking the action; the same action might be taken in a number of contexts, with different expected results. Note also that a schema, even if reliable, is not a rule that says to take the action when the context is satisfied; rather, the schema just asserts what would happen if the action were taken then.

The Schema Mechanism lives in a two-dimensional microworld, populated by objects that can be seen, felt, grasped, and moved. The mechanism controls a body that has a hand and an eye. Each primitive item corresponds to a sensory input; for example, for each of 25 regions in the visual field, there is an item that is in the On state whenever an object appears in that region. (This is meant to be analogous to an output of lowlevel vision in humans, rather than, say, to the state of a retinal cell.) Other visual primitive items provide detail about the appearance of objects at the central, *foveal* region of the visual field. Tactile primitive items report contact with the and, and other parts of the body. Finally, *proprioceptive* primitive items report the body-relative position of the hand, and the glance orientation. For each of 25 glance orientations, there is a visual-proprioceptive item that is On whenever that orientation is current; similarly, there are 25 haptic-proprioceptive items that report hand position. There are ten primitive actions: four actions for moving the hand incrementally forward, back, right, or left; four for incrementally changing where the visual field maps to; and opening and closing the hand.

A schema whose context conditions are currently satisfied competes for *activation*— having its action taken— based in part on its leading to the satisfaction of some goal. (Also, a schema can suppress its action if the schema predicts an undesirable result in the current situation.) Schemas can form an implicit chain from a current state to a goal state, the result component of each schema in the chain including the elements of the next schema's context; the mechanism's parallel architecture lets such chains be found quickly. The mechanism's built-in goals include mundane ones (eg eating), as well as curiosity-based goals, which appeal to heuristic assessments of the usefulness and interestingness of the mechanism's constructs. In addition to built-in goals, some states become valued as goals because of their strategic facilitation of other things of value. I omit further discussion here of criteria for activation and valuation, to emphasize instead the machinery for building new structures.

III. Building New Schemas

The Schema Mechanism looks for results that follow from actions; and, if a result follows unreliably, the mechanism seeks conditions under which the reliability improves. The mechanism builds schemas that reflect these discoveries. Typically, the derivation of a reliable schema involves building a series of intermediate ones, which alternate between discovering intermittent results of a schema's activation, and finding additional conditions that must hold for the results to follow reliably. In the beginning, for each primitive action, there is also a built-in schema with that action, and with empty context and result. These initial schemas, which assert nothing, are points of departure for building contentful schemas.

The Schema Mechanism builds new schemas from existing ones by extending the context or result of an existing schema. The old schema doesn't change, but a copy, or *spinoff schema*, appears, with a new item added to its context or result.

Every schema has an *extended context* and an *extended result*, in addition to the context and result proper. Each extended context or result has an slot for every item, primitive and nonprimitive, in the mechanism's database. For each schema, each extended result slot keeps track of whether the associated item turns On more often if the schema has just been activated than if not. If so, the mechanism attributes that state transition to the action, and builds a spinoff schema, with that item included in the result. (If a schema's activation makes some item more likely to turn Off, the item's *negation* joins the result of a spinoff schema.)

A result attributed to a schema's activation may be arbitrarily unlikely to follow the schema's activation; the result must only be significantly *more likely* than if the schema isn't activated. A spinoff schema can thus be arbitrarily unreliable. But a schema's extended context tries to identify conditions under which the result more reliably follows the action. Each extended context slot keeps track of whether the schema is significantly more reliable when the associated item is On (or Off). When the mechanism thus discovers an item whose state is relevant to the schema's reliability, it adds that item (or its negation) to the context of a spinoff schema. (Extensions of this scheme, described in [Drescher 1985,1986], increase its sensitivity to certain kinds of context conditions, reduce the proliferation of effectively redundant spinoffs, and suppress otherwise-reliable schemas when exceptional, overriding conditions hold.)

For purposes of execution only, three-part schemas could instead be two-part production rules, context and action collapsing into the left-hand part of a rule. But a bipartite structure is inadequate for building new schemas by marginal attribution, which needs to treat context, action, and result differently.

The Schema Mechanism uses only reliable schemas to pursue goals. But the mechanism needs to be sensitive

to intermittent results, because a reliable effect can seem arbitrarily unreliable until the relevant context conditions have been identified. Consider, for example, the action of shifting the glance incrementally to the left. This reliably turns On the item designating an object at, say, the center of the visual field— provided that an object was seen just left of center beforehand. Until that prior condition is recognized as such, the result will be seen to follow from the action only infrequently.

Moreover, the same action, in other contexts, has different results (eg, making other visual-field items change state); furthermore, the given result often occurs without the action in question, caused instead by another glance action, by a hand action, or by an external event; and, whether or not the result obtains, the action typically accompanies many other, coincidental transitions. Despite all this, the result is more likely to occur at a given moment if the glance-left action is taken than if not (presuming, realistically, that objects' images spend somewhat more time being approximately stationary in the visual field than they spend moving). Thus, the initial glance-left schema can identify the visual-field-center item as a tentative result; this prompts the construction of a spinoff schema, whose extended context then finds the condition (namely, the visual-field left-of-center item being On) that confers reliability. This discovery spawns another schema, this time a reliable one.

An alternative mechanism might look for reliable results already paired with appropriate contexts, rather than trying to identify infrequent results independently first. But there are too many such pairs to consider them exhaustively. Usually some *conjunction* of conditions must hold for a result to follow an action reliably; hence, contexts of reliable schemas typically include more than one item. Results, too, must be able to include multiple items, in order to chain to multiple-item contexts. With m actions and n items (and their negations), there are $m3^{2n}$ expressible schemas; even if contexts and results were limited to, say, five items each, there would be about $m(2n)^{10}$ expressible schemas. If there are to be thousands, or perhaps millions, of actions and items, even $m(2n)^{10}$ possibilities are far too many for exhaustive search.

One might try to relieve the combinatoric problem by partitioning actions and items into categories, designing the mechanism to seek connections within categories, not between them. Indeed, it seems plausible that most actions are irrelevant to most items. But I am skeptical that many categories of mutual relevance can be usefully characterized in advance. Among the primitives, for example, hand actions have haptic proprioceptive results, tactile results, and visual results; thus, we can exclude neither inter- nor intra-modal connections. As for nonprimitive items and actions, it seems even less plausible to be able to impose *a priori* constraints on the mutual relevance of constructs that themselves are not known *a priori*.

Thus, I propose instead the present *marginal attribution* scheme, whereby the mechanism can identify an

action's contribution to a result before hypothesizing the corresponding context conditions, even if, out of context, the result follows the action only infrequently, and amid many other, irrelevant events. This approach is not inexpensive; exhaustive cross-connectivity between schemas and items may seem an exorbitant, brute-force solution. But it is a bargain compared to the size of the space being searched, the space of expressible schemas.

IV. Building New Actions

The Schema Mechanism builds new actions, called *composite actions*. Each composite action has a *goal state* which, like a context or result, is a set of items. A composite action identifies schemas that can help to achieve the goal state— schemas that chain to the goal state from various other states. When a composite action is initiated, it coordinates the successive activation of schemas to reach the goal state (if possible from the initial state); these schemas need not independently compete for activation.

Any newly-achievable result— any conjunction of items that appears for the first time in some reliable schema's result— is a candidate goal state for a new composite action. As with each primitive action, the mechanism builds for each new composite action a schema with empty context and result that uses that action. For example, if the mechanism has built schemas that say how to turn on a lightswitch, then the mechanism could also define a composite action whose goal state is lightswitch-on. The schema with that action can then discover what results from the lightswitch being on.

It is important to be able to represent the action at the right level of abstraction— as lightswitch-on, rather than just as whatever primitive motor action is responsible for pushing the switch on. A schema that looks for results of lightswitch-on per se discovers and represents effects that are independent of the particular motor sequences responsible; hence, in the absence of contrary evidence, the discovery automatically generalizes to other motor implementations of the same higher-level action. Furthermore, the mechanism regards a composite action as having been taken whenever its goal state is satisfied, even if external events are responsible; hence, composite actions let the mechanism look for the effects of external events, not just of its own actions.

V. Building New Items

It is important for a learning mechanism to discover relations among existing representational elements, and to organize such knowledge at appropriate levels of abstraction. But a constructivist system's greatest challenge is to synthesize new elements of representation, to designate what had been inexpressible. *Synthetic items* enable the Schema Mechanism to do this.

Each synthetic item is based on some schema that says, in effect, how to recover a manifestation of something

that is no longer shown; we can say that the synthetic item *reifies* this recoverability, construing the potential-to-recover as a thing in itself.

For example, suppose the Schema Mechanism moves its hand away from some stationary object directly in front of its body (and suppose the mechanism's eye is directed away from the object). Presumably the object is still present; but, at first, the mechanism has no way even to represent this fact, since the object now has no manifestation in the state of any primitive items. The Schema Mechanism, like a four-month-old infant in Piaget's theory, is simply oblivious to the possibility of reaching again for the unperceived object, or of turning to look at it.

But suppose there is a schema with empty context, whose (nonprimitive) action is moving the hand directly in front of the body (as indicated by a haptic proprioceptive item), and whose result is touching-something. (In other words, this schema says: if I reach directly in front of me, I'll touch something there.) Of course, this schema is unreliable; it only works when there is in fact an object sitting there. But, significantly, this schema is *locally consistent*: if it activates and achieves its result, it is likely to be reliable if activated again in the next little while (because, typically, objects tend to stay put for a while).

The Schema Mechanism keeps track of the observed local consistency of each schema. If a schema is unreliable, but locally consistent, the mechanism constructs a synthetic item for it. This item designates *whatever unknown condition in the world* governs the schema's success or failure; the schema's local consistency implies that this condition is slow to change state. In the present example, the mechanism creates a synthetic item that designates a palpable object directly in front of the body. While a synthetic item is On, the mechanism regards the item's *host schema* (the schema for which the item was created) as reliable.

An item is useful only to the extent that some machinery turns the item On or Off when the condition it designates does or doesn't obtain. Each primitive item, of course, is simply wired to some peripheral apparatus that maintains the item's state. To maintain the state of a given synthetic item, the mechanism exploits three kinds of clues:

- When the item's host schema activates successfully, the item turns On; unsuccessfully, Off. The item reverts to the Off state a while after being turned On; the length of the while is the empirically established expected duration of the host schema's local consistency.
- The host schema's extended context looks for items whose states correlate empirically with the synthetic item's: an item whose being On implies that the synthetic item is On (or Off); or whose being Off implies that the synthetic item is On (or Off). If the mechanism finds a strong correlation, it thereafter turns the synthetic item On or Off according to the correlated

item's state.

- The synthetic item, like any other item, may appear in the context or result of subsequently constructed schemas. If the item (or its negation) is in the result of a reliable schema, the mechanism turns the item On (or Off, respectively) when that schema has completed activation.

A synthetic item's state-maintaining criteria bootstrap from one another: as these criteria accumulate, the item becomes more likely to turn On or Off when appropriate, increasing the mechanism's opportunity to discover further correlates of the item's state. The crucial step is the first one: synthesizing an item to reify an unknown condition lets the mechanism start to learn about that condition.

VI. The Hypothetical Scenario

In [Drescher 1985,1986], I present a detailed hypothetical scenario in which the Schema Mechanism builds its way toward a late-sensorimotor-stage conception of physical objects. First, the mechanism assembles a substrate of spatial knowledge: it builds a network of schemas that denote the adjacency of pairs of visual field items by noting the transformation from an item to an adjacent one by the appropriate incremental-glance action (as in the example above). A similar network shows the relations among visual proprioceptive items, again with respect to glance actions; and another network relates the haptic proprioceptive items via incremental hand-motion actions.

Early intermodal coordinations appear. Some schemas anticipate contact between hand and body when the hand moves incrementally from certain proprioceptively-designated starting places. Other schemas predict visual effects of moving the hand when it is in view; still others anticipate tactile contact when, for example, the hand is seen just left of some object, and moves left. This anticipation corresponds to the earliest form of Piagetian visual-tactile coordination in infants.

The visual-field schemas chain together to enable the mechanism to *foveate*: to look directly at some object that appears at the visual periphery. The visual proprioceptive schemas chain together to enable the mechanism to shift from any glance orientation to any other; similarly, chains of haptic proprioceptive schemas lead from any hand position to any other. Each proprioceptive item is by now an achievable result; hence, each such item is the goal state of a "positional" (as opposed to incremental) action, the action of moving the hand or eye to a certain orientation. Positional hand actions facilitate knowing how to move the hand into view: each visual proprioceptive item serves as the context of a schema whose action is moving the hand to a certain position, and whose result is seeing the hand. When an object is in view, schemas for moving the hand into view near the object chain to schemas that say, based on the visual appearance of the hand and object, how to move the hand to touch the object. This coordination extends the mechanism's earlier, cruder visual-tactile coord-

dination. Other schemas chain in the opposite direction, enabling the mechanism to shift its gaze to look at what the hand touches.

The positional actions also facilitate the construction of synthetic items that designate objects at the various positions. For instance, as in the example above, there is a schema with empty context, whose action is moving the hand directly in front of the body, and whose result is touching; the corresponding synthetic item designates a palpable object directly in front of the body. Other synthetic items designate palpable objects at other body-relative positions. Analogous schemas, with positional eye (rather than hand) actions, give rise to synthetic items that designate visible objects at various positions. At first, nothing prevents, say, a palpable-object synthetic item from being On while the visual-object item for the same position is Off; in that case, the mechanism knows that it can reach back to the object, but is oblivious to the possibility of glancing at it. Later, though, each visual-object item's state-maintaining apparatus recognizes the corresponding palpable-object item's state as an indicator of the item's own state (and vice versa). Later still, the mechanism synthesizes items that designate objects hidden by obstacles; each such item's host schema shows how to recover the object by displacing the obstacle (in contrast with the simpler action of reaching or glancing back to an unhidden object). Each of these elaborations of the concept of physical objects corresponds to a milestone in Piagetian development.

The Schema Mechanism's earliest acquired abilities are probably innate in humans. My proposal of a *tabula rasa* learning mechanism is not meant to deny the extensive innate domain-specific competence in human *peripheral* modules. But by the present hypothesis, this competence is not available to the central system, which sees only the outputs of the peripheral modules, presented as "gensyms". The central system must reconstruct in its own scheme of representation much of the innate peripheral competence, as a first step to surpassing the innate competence.

VII. The Implementation

An existing program, called MARCSYST (Marginal Attribution and Representation Construction System), partially implements the Schema Mechanism. So far, MARCSYST builds spinoff schemas, but does not yet implement composite actions or synthetic items. In a typical run of 20,000 simulated seconds, the mechanism builds schemas that accord with the beginning of the hypothetical scenario: there are schemas that comprise much of the eventual visual field network, and the visual and haptic proprioceptive networks; and many schemas that designate hand-eye, and hand-body coordination. These results, while quite preliminary, are on the right track.

After accumulating several hundred schemas, the present Lisp Machine version of MARCSYST slows to sev-

eral real-time seconds per simulated second. I plan to move MARCSYST to a massively parallel machine before completing the implementation.

VIII. Related Work

[Jones 1970], [Cunningham 1972], [Becker 1973], and [Klahr, Wallace 1976] propose mechanisms for aspects of Piagetian or sensorimotor development; Cunningham's work, including a detailed sensorimotor scenario, inspired my own effort. Becker's schemas, like mine, have a context, action, and result. None of these systems addresses the combinatoric problem in finding empirical associations; and none except Klahr and Wallace's constructs nontrivially novel elements of representation. Klahr and Wallace's system builds tokens that designate the applicability of production subsystems; these tokens are similar in spirit to synthetic items.

Acknowledgements

I am grateful to Seymour Papert, and Hal Abelson, Phil Agre, David Chapman, Marvin Minsky, Ron Rivest, Gerry Sussman, and other friends and colleagues for illuminating discussions and support. Ron Rivest showed me how to halve the implementation's memory needs. David Chapman helped with comments on this paper.

References

- [Becker 1973] Becker, J. "A Model for the Encoding of Experiential Information", *Computer Models of Thought and Language*, eds. Schanck, R. and Colby, K. pp 396-434. San Francisco: Freeman, 1973.
- [Cunningham 1972] Cunningham, M. *Intelligence: Its Origins and Development* New York: Academic Press, 1972.
- [Drescher 1985] Drescher, G. *The Schema Mechanism: A Conception of Constructivist Intelligence*. MS thesis, MIT, 1985.
- [Drescher 1986] Drescher, G. "Genetic AI: Translating Piaget into LISP", MIT AI Laboratory Memo 890, 1986.
- [Jones 1970] Jones, T. *A Computer Model of Simple Forms of Learning*. Phd thesis, MIT, 1970.
- [Klahr, Wallace 1976] Klahr, D. and Wallace, J. *Cognitive Development: An Information Processing View*. New York: Lawrence Erlbaum, 1976.
- [Piaget 1952] Piaget, J. *The Origins of Intelligence in Children*. New York: Norton, 1952.
- [Piaget 1954] Piaget, J. *The Construction of Reality in the Child*. New York: Ballantine, 1954.