Title

Learning and Memory in Machines and Animals: An AI Model That Accounts for Some Neurobiological Data

Permalink

Journal

Authors

Granger, Richard H.
McNulty, Dale M.

Publication Date

Peer reviewed

Learning and memory in machines and animals:
An AI model that accounts for some neurobiological data[1]

Richard H. Granger
and
Dale M. McNulty

Artificial Intelligence Project
Computer Science Dept.
Cognitive Sciences Program
and
Center for the Neurobiology of Learning and Memory
University of California
Irvine, California 92717

ABSTRACT

The CEL model of learning and memory (Components of Episodic Learning) *[Granger 1982, 1983a, 1983b]* provides a process model of certain aspects of learning and memory in animals and humans. The model consists of a set of asynchronous and semi-independent functional operators that collectively create and modify memory traces as a result of experience. The model conforms to relevant results in the learning literature of psychology and neurobiology. There are two goals to this work: one is to create a set of working learning systems that will improve their performance on the basis of experience, and the other is to compare these systems' performance with that of living systems, as a step towards the eventual comparative characterization of different learning systems.

Parts of the model have been implemented in the CEL-0 program, which operates in a 'Maze-World' simulated maze environment. The program exhibits simple exploratory behavior that leads to the acquisition of predictive and discriminatory schemata. A number of interesting theoretical predictions have arisen in part from observation of the operation of the program, some of which are currently being tested in neurobiological experiments. In particular, some neurobiological evidence for the existence of *multiple, seperable memory systems* in humans and animals is interpreted in terms of the model, and some new experiments are suggested arising from the model's predictions.

1. Introduction to the problem

1.1 Characterization of learning processes

The amnesic patient identified by his initials 'H.M.' is apparently incapable of learning any new information; since the operation that removed a part of the limbic system of his brain, he has been unable to learn to recognize new people or situations. For instance, he re-introduces himself to his doctor Brenda Milner every time she visits him, even though she has visited him many times a week for many years! In contrast, his pre-operation memories appear not to be impaired, nor is his ability to carry on a relatively normal conversation or other everyday functions.

However, H.M. *can* acquire certain categories of new abilities. For instance, he has been tested on the 'mirror-writing' task of writing while seeing only a mirror image of what he writes. Every time the experimenter came in the room, once a day for several weeks, H.M. had to be re-introduced to both the experimenter and the experiment, and insisted that he had of course never seen either before, and that he didn't know how to do this (mirror-writing) task. *Yet his performance on the task improved steadily over the several-week period; in fact, he learned the task at about the same rate that control subjects did.* When confronted with examples of his poor early trials compared with his much-improved recent trials, he is unable to explain how the differences arose, and doesn't remember ever performing those experiments.

These and other results in humans and animals have led inescapably to the hypothesis that there are *multiple memory systems*, i.e., separable biological systems that semi-independently establish

long-term memories from experience. As suggested by H.M.'s behavior, these two systems have distinct characteristics; i.e., each is only capable of learning certain types of information.

However, an accurate *characterization* of precisely *which* tasks are learnable by which mechanism has proven elusive; there currently exist a number of competing neuropsychological hypotheses characterizing the different memory systems (see e.g., *[Squire 1980; Squire, Cohen and Nadel 1982; Mishkin 1982]*).

A long-term goal of the research described here is to attempt to characterize these different learning systems in terms of the types of learning behavior they produce. Our current subgoal is to create a system in which the behavior of learning and memory systems can be characterized. We hope to be able to build two different systems out of similar functional components, each of which has a particular set of learning abilities. We would then be able to show what differences in the models gave rise to the differences in learning abilities.

Recent research on learning and memory in AI has focused primarily on advanced human abilities (see e.g., *[Schank 1983; Schank and Burstein 1981; Lebowitz 1982; Kolodner 1983; Carbonell 1982; Langley 1981, 1982; Mitchell 1981]*). We have adapted some ideas on MOPs *[Schank 1983]* and the indexing of E-MOPs *[Kolodner 1983a, 1983b]* to the tasks we are modeling. Our focus is on much lower-level domains of learning and memory, especially 'subcognitive' tasks that lower mammals (e.g., rats) can learn. This has enabled us to concentrate our models on the *processes* underlying learning and memory, rather than on complex memory *structures*; our approach has been to attempt to identify a candidate set of mechanisms sufficient to allow the acquisition, storage and retrieval of simple episodic information, and to compare our results against experimental data on learning and memory processes in animals.

A key point here is that these 'subcognitive' learning and memory tasks, as far below 'higher' human abilities as they are, are nonetheless still difficult and elusive, and therefore eminently worthy of being the focus of an AI learning mechanism. These and related tasks have been extensively studied by cognitive psychologists and neurobiologists in their experimental approaches to learning and memory; yet their theories of human and animal learning and memory have been insufficiently precise to allow for the construction of computer models for testing the theories. Still missing is a bridge between AI models of learning, and psychological and neurobiological experiments on learning.

1.2 Introduction to **CEL**

The CEL model of learning and memory *[Granger 1982, 1983a, 1983b]* provides a process model of the acquisition and operation of certain aspects of learning and memory in animals and humans. The model conforms to constraints provided by relevant results in psychology, neuropsychology and neurobiology; a number of behavioral data are explained in terms of the model, and certain specific theoretical lesions and modulations of the model predict behavioral effects that correspond to observed behaviors in similarly manipulated animals.

Parts of the model have been implemented in a computer program called CEL-0. CEL-0 takes as input a sequence of experiential sensory events coded in terms of sensory modality and feature sets. The program operates on the inputs, building a memory database of information derived from the input streams. Sample domains that have been worked on include a simple 'feeding' microworld in which the model learns to predict (via classical conditioning) which events reliably and predictably lead up to its being fed; and a 'maze' microworld in which the model explores and learns (operantly) to identify where 'interesting' and rewarding areas of the maze are, and to create a simple 'cognitive map' *[Tolman 1932, O'Keefe and Nadel 1979]* of the maze environment. In the maze microworld, CEL-0 interacts with MazeWorld, a simulated 'maze environment' program, that receives CEL-0 input moves, and returns a value indicating CEL-0's new location in the maze; hence, each move of CEL-0 causes 'feedback' from the simulated maze, which in turn triggers CEL-0's next move.

Some of CEL-0's unexpected behavior in the MazeWorld has triggered some new theoretical ideas which are presented here. For instance, we have identified seven different categories of learning, i.e., seven different ways that new memory traces can be created in CEL-0, each corresponding to a different 'calling sequence' of operators, each of which in turn seems to correspond to a logical class of training situations that might arise in the real world. These seven classes of learning will be briefly discussed in a later section of this paper.

Other examples of theoretical ideas that have arisen from working with the program in the MazeWorld include: a mechanism for active 'exploratory behavior' during learning, a mechanism for creating subgoals from goals during learning, acquisition of 'landmarks' during learning that serve as useful index points, and a comparison of 'efficient' learned behavior vs. 'superstitious' learned behavior; some of these are described in some detail in *[Granger and McNulty 1984]*.

Attempts to find detailed correspondences between the model and experimental data in neurobiology have so far been fruitful. A number of specific predictions arising from work with the model are in the disparate areas of selective attention, modulation of memory, and rapid forgetting and learning deficits associated with certain limbic lesions ([Granger 1983b] presents detailed analyses of these three substantive areas of CEL's modeling efforts).

This paper will present first a description of CEL-0's behavior in the MazeWorld simulation, and then a specific neurobiological prediction dealing with multiple memory systems; the prediction is currently being tested in a neurobiological lab at the Center for the Neurobiology of Learning and Memory at Irvine.

2. Introduction to the model: The twelve CEL operators and their functions

The **CEL** model proposes a characterization of the constituent functional operators that comprise learning processes, in the hope that these primitive operators may each have specific instantiations that can be identified in the neural substrate. The model identifies a set of twelve 'primitive' memory operators which operate in parallel to collectively perform five classes of memory manipulation: reception, recording, retrieval, reconstruction and refinement. The model consists of the operation of these twelve operators on memory representations we term episodic schemata. Detailed descriptions of the functions of these operators and their (often nonintuitive) interactions are provided in [Granger 1982, 1983a, 1983b].

In brief, the twelve operators have the following functions:

- Reception operators:

  DETECT - *set of sensory input channels and any associated hard-wired preprocessing performed by those input channels, such as visual and auditory processing;*
  SELECT - *'tunable' input filter to selectively attend to some inputs over others on the basis of prior experience;*

- Recording operators:

  NOTICE - *matches inputs against known desirable and undesirable states; triggers COLLECT when a match occurs;*
  COLLECT - *packages recent stream of inputs into a kernel episodic schema;*
  INDEX - *creates new indices, and hooks into existing indices, for each new episodic schema;*

- Retrieval operators:

  REMIND - *matches inputs against indices for existing schemas; triggers ACTIVATE when match occurs;*
  ACTIVATE - *incorporates REMINDed schemas into current predicive schema; triggers the Reconstruction operators;*

- Reconstruction operators:

  ENACT - *performs any efferent actions in current predictive schema; 'tunes' SELECT's filter to attend to predicted afferent events;*
  SYNTHESIZE - *matches inputs against predicted events in current predictive schema; triggers Refinement operators to modify schema in response to matches and mismatches;*

- Refinement operators:

  REINFORCE - *incrementally strengthens current schema(ta) according to SYNTHESIZE's judgment of its successful predictiveness (i.e., matches);*
  BRANCH - *creates a branch in current schema(ta) according to SYNTHESIZE's judgment of unsuccessful predictiveness (i.e., mismatches);*
  DETOUR - *creates a non-pursuable branch in current schema according to NOTICEings of undesirable events, predicted or not.*

These operators act in parallel, asynchronously and semi-independently in the **CEL** model, and complex interactions among them at run time enable these relatively straightforward operators to give rise to a rich set of learning and memory behaviors.

3. A brief example of the operation of the CEL-0 program

3.1 Introduction to the CEL-0 environment

The setting described here for CEL-0's operation is a relatively simple maze, that CEL-0 moves through by interacting with MazeWorld, a simulated 'maze environment' program. MazeWorld receives CEL-0 input moves, and returns a value indicating CEL-0's new location in the maze; hence, each move of CEL-0 causes 'feedback' from the simulated maze environment, which in turn triggers CEL-0's next move. Following is a schematic view of the relatively simple MazeWorld maze that we will use for the examples in this section; 'M0' is the entry point into the maze, and 'M4' contains water, which will be used for a 'reward' under circumstances to be described later.
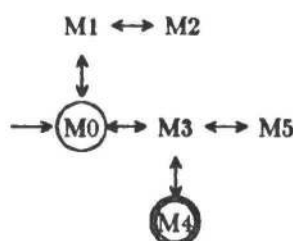
The following sections describe a connected set of examples of CEL-0's operation in this maze. The description will be in three phases:

Phase 1 ('Exploration' phase): CEL-0 uses 'innate' (built-in) episodic schemas to move through the maze, establishing episodic traces corresponding to its 'routes' through the maze.

Phase 2 ('Effectiveness' phase): CEL-0 has an added desired state (satisfy-thirst) that drives its behavior; it searches for and finds a (not necessarily most efficient) route through the maze to any location of water.

Phase 3 ('Efficiency' phase): CEL-0 refines its already-effective routes through the maze to reward locations. (This 'phase' is actually going on in parallel with the other two).

In each phase, CEL-0's behavior can be described in terms of three lists: a sequence of CEL operators, the corresponding sequence of overt moves in the simulated environment (if any), and the corresponding additions or changes to long-term memory (if any).

3.2 CEL-0's exploratory behavior in MazeWorld

For the purposes of this example, CEL-0 will start at location M0 in the maze, facing towards M3. The internal representation is described in *[Granger and McNulty 1984]*; it simply consists of information about what views are in front of, to the right of, behind and to the left of the current position of CEL-0 in the maze. Hence, the starting position has a view of M3 in front, walls to the right and behind, and M1 to the left.

(This is an admittedly huge oversimplification of a 'realistic' maze situation, but it seems justified for two important reasons: (1) selective attention to relevant features is the key thing that gets slighted by this oversimplification, and we have already done some analyses of selective attention in complex environments (see *[Granger 1983b]*); and (2) there are a number of interesting and complex processing problems that arise even *with* this simplification, and these problems would be difficult to present without first simplifying away the selective attention problems for pedagogical reasons.)

Because of the extremely simplified inputs for this example, DETECT and SELECT essentially just attend to everything here; see *[Granger 1983b]* and the selective attention section in this paper for an explanation of how these operators become much more complex in the face of more complex inputs.

Once SELECT has entered a representation into temporary memory, NOTICE attempts to match it against desirable and undesirable states, and REMIND attempts to match it against any existing schemas that might be relevant to the situation. There are three built-in 'exploratory schemas' in CEL-0, two of which get REMINDed by this input. Each of the three schemas (ES1, ES2 and ES3) is simply two events long, each corresponding to the 'impetus' to move in a particular type of situation, essentially corresponding to the following sequences:

ES1: see front opening $\Rightarrow$ go straight

ES2: see obstruction $\Rightarrow$ look around (360°)

ES3: see side opening $\Rightarrow$ turn towards opening

So at location M0, REMIND will find both ES1 and ES3a, then ACTIVATE will have to choose at most one of them to pursue; for this example let it choose ES1. (ACTIVATE in fact contains a set of (currently six) 'preference metrics' that it uses to decide among proposed (REMINDed) alternative schemas – *[Granger and McNulty 1984]* describe these in detail). ENACT and SYN-THESIZE then begin to reconstruct ES1. ENACT does so by performing any events in the schema, and SYNTHESIZE by comparing new inputs that result from successive ENACTed events against the 'predicted' inputs in the schema itself. SYNTHESIZE notes that the match between the event and the more generalized representation in ES1 is only a partial match, and because it's not an exact match, calls BRANCH to create a new branch of the schema, and begins recording this new branch.

CEL-0 continues in this fashion, making the following moves through the maze: M0 - M3 - M5 - M3 - M0 - M1 - M2 - M1 - M0 - M3 - M5 - M3 - M4 - M3 - M0 - M1 - M2 - etc. An extensive description and explanation of the operator sequences driving these moves can be found in *[Granger and McNulty 1984]*.

Note that when M4 is arrived at, the fact that there is water there will cause a REMIND of another innate (built-in) schema that essentially says when water is seen, drink it. However, this schema may not be reconstructively ENACTed unless ACTIVATE lets it be (or unless there are no alternative schemas that get REMINDed); one of ACTIVATE's preference metrics says not to prefer schemas that do not match any currently desirable state, as specified on the 'Desirable State List' (DSL).

3.3 Effective goal pursuit in **CEL-0**

The result of the 'exploration' phase is the creation of a number of schemas describing various 'routes' through the maze, indexed by their starting and ending positions (more detail on INDEX is provided in *[Granger 1983b]* and *[Granger and McNulty 1984]*).

In phase 2, we simply add a desired event to CEL-0's Desirable State List (DSL) – this is the list that NOTICE matches incoming events against, and that ACTIVATE checks to see whether or not to bother to ENACT a REMINDed schema. Hence, if we add 'drink water' to the DSL, **CEL-0** will now 'act thirsty', in the following three senses: (1) it will drink water if it sees any (via REMIND, ACTIVATE and ENACT of the built-in schema that says when water is seen, drink it); (2) it will tend to prefer sequences that lead to seeing water (via ACTIVATEs preference metric for currently-desirable states); and (3) it will store any sequences of events that lead to water (via NOTICE and COLLECT). Hence, via all three of these mechanisms, CEL-0's memory will now contain schemas that are 'effective' with respect to the achievement of its goal of finding and drinking water.

3.4 Active exploration by **CEL-0**: Sensitivity analysis

A schema that leads to M4 at the end of the exploration phase is: M0 - M3 - M5 - M3 - M4. Note that while *effective*, this schema is not maximally *efficient* – it could simply go M0 - M3 - M4. The fact that it doesn't is simply an accident of the exploration phase (see Section 4.2 below). **CEL-0** has a process that causes schemas to be tested for their sensitivity to changes in the sequence; the process makes multiple variations of schemas by deleting various features or events from the event sequence, and then tests the resulting variations for their effectiveness.

The process effectively establishes a set of multiple internal 'hypotheses' as to which of the features of the episode are the most critical and predictive. Hence, this process amounts to a test of the sensitivity of the new episode to changes in those features. This process of testing episodes for their sensitivity to changes is termed 'sensitivity analysis'. The following subsections briefly outline the process.

3.4.1 Introduction to sensitivity analysis

When the model senses an instance of an episode, say, a pursuit-type episode such as M0 - M3 - M5 - M3 - M4, which results in some NOTICEd desirable state, that episode is COLLECTed into a long-term memory trace. The INDEX operator then begins to choose features of the events in the episode to use as indices, which will be used as recognition cues at retrieval time, i.e., whenever similar events happen subsequently. Depending on which features are chosen as indices, of course, subsequent retrieval either will or will not take place based on the presence or absence of any particular feature in the new input trace. Hence, the effective recognition of any new instance of a learned episode is sensitive to the feature-indices that are created at INDEX time during recording of the episode.

During the establishment of these feature indices on a new trace, the INDEX operator performs a multi-step process which has the effect of creating multiple traces of the episode, each with a different feature or set of features deleted from the trace. The multiple versions of the episode that result from this process serve the purpose of enabling CEL-0 to test instances of the episode for their sensitivity to changes in the constituents of the episode: the long-term trace undergoes ongoing modification and refinement depending on which versions of the episode turn out to accurately match subsequent instances of the input stimuli.

Intuitively, what is happening is that the INDEX operator is hypothesizing a series of variations of the instance of the episode, implicitly predicting that these versions might serve as useful predictors of subsequent instances of the episode. Those predictive hypotheses are tested each time the set of variations of episodes are retrieved and reconstructively compared to a new input instance (via ENACT and SYNTHESIZE). In this way, the sensitivity analysis process allows the model to learn more than was contained in the single instance of the episode: it learns the ways in which that episodic instance might be sensitive to changes and variations. Furthermore, the process has the effect of robustly reducing any dependance on the order of presentation of events, making the model eventually learn the same things about the maze regardless of what order it happens to acquire them in.

3.4.2 The five steps of sensitivity analysis

1. When a new long-term trace is written, INDEX's first step is to search for any existing feature indices that match any features in the new trace. If so, then those indices are 'attached' to the trace, i.e., each index now points to the new trace in addition to any other traces it may already be pointing to.

2. For each sensory feature in the input, create a new index for that feature, that points to the episode.

3. For each feature-index pointing to the episode, either found by step 1 or created by step 2, begin creating variations of the episode by leaving out one or more of the features contained in the initial copy. Each variation is written into memory as a 'near-miss' copy of the episode.

4. For each of the new episode-variations, search for an existing index that has the new subset-features of the variation; if found, attach it to the episode.

5. For each feature-set index created, attempt to find others with subsets of the same features. For each such index set found, create a new higher-level index (see [Granger 1983b]), corresponding to the shared features that points to each of the members of the index set.

The combined effect of these steps will be to create a growing set of indices pointing to the episode, each of which will be triggered by a different set of feature cues at retrieval time. At the same time, multiple copies of the episode itself are being created, each a slight variation of the others; i.e., no two are exactly alike. The indices will slowly become a hierarchical set, because step 5 creates higher-level or 'second-order' indices, each of which points only to other indices (see [Granger 1983b]). For instance, 'template indices' are examples of higher indexes that contain only event-sequence information, with specific sensory information deleted.

4. Some insights resulting from experience with CEL-0

There are a number of difficulties that have arisen during the programming of CEL-0 that have the form of interesting theoretical problems that were not obvious until the implementation difficulties arose. Some of these are discussed here, with the focus on the emergence of seven categories of learning, based on seven different 'calling sequences' of CEL operators all of which are capable of establishing or modifying a memory trace, i.e., learning.

4.1 Seven ways to establish a memory trace in CEL

The twelve CEL operators do not call each other serially; hence, although COLLECT is the primary way for episodic traces to be established in permanent memory, there are four distinct calling sequences that may result in the creation of a new trace, each of which constitutes a category of learning in CEL; in turn, these four categories have between them a number of different subcategories, for a total of seven. These are listed here, followed by a set of brief descriptions and examples of each subcategory.

Goal-based establishment:

  1. *pursuit* of desirable result *(Pursuit-based learning)*
  2. *avoidance* of undesirable result *(Avoidance-based learning)*

Expectation-based establishment:

  1. match between expectation and environment *(Success-driven learning)*
  2. mismatch between expectation and environment *(Failure-driven learning)*

Exploration-based establishment:

  1. analysis of relevance of schema features *(Sensitivity analysis)*

Coincidence-based establishment:

  1. schema activated simultaneously with newly-created schema *Append-driven learning)*
  2. two schemas concurrently activated *Splice-driven learning)*

4.1.1 Goal-based trace establishment

When the NOTICE operator finds that an incoming event matches something on either the Desirable or Undesirable state list (*DSL* or *USL*) (see *[Granger 1983b]*), NOTICE triggers the COLLECT and INDEX operator to make a record of the sequence of events that led up to the desirable or undesirable event.

Case one: Pursuit-based learning

In the *desirable* case, the INDEX operator simply indexes the sequence of events by SELECTed features (see *[Granger 1983b]*).

Case two: Avoidance-based learning

In the *undesirable* case, INDEX calls the DETOUR operator to attempt to create a link pointing to potential *alternatives* to the undesirable result, so that that path won't be pursued in the future.

4.1.2 Expectation-based trace establishment

While a schema is being reconstructively ENACTed after having been triggered (REMINDed and ACTIVATEd) by some cue, the SYNTHESIZE operator is constantly matching incoming real-world events against events in the schema (i.e., it is checking the schema's implicit expectations). Both matches and mismatches can cause new things to be written into memory.

Case three: success-driven learning

If SYNTHESIZE finds a match, then it calls REINFORCE to add 'strength' to the links pointing to the successfully predictive schema.

Case four: failure-driven learning

If a *mismatch* is found, BRANCH is called to create a new link between the index and the new sequence of events (whatever just actually happened), thereby effectively reducing the relative strength of the link from the index to the previously-expected result.

4.1.3 Exploration-based trace establishment

Apparent exploratory behavior by CEL arises from the operation of the 'sensitivity analysis' procedure described above (and described in more depth in *[Granger 1983b]*), combined with the existence of the set of simple 'exploratory schemata'. Recall that sensitivity analysis causes a number of variations of each schema to be created, each of which will be tested and either strengthened or weakened according to its success or failure. These will operate on the schemas collected during CEL-0's 'wandering' through the maze, to refine the model's representation of pathways through the maze, eliminate some redundancies, and identify some 'landmarks' that make useful indices to the set of pathways (see *[Granger and McNulty 1984]*).

As it collects sequences of paths through pieces of the maze, sensitivity analysis refines them by testing the relevance of their consituent events.

Case five: Sensitivity analysis

For instance, if an initial route through the maze is the sequence M0 - M3 - M5 - M3 - M4, a diminution of the route yields M3 - M5 - M3 - M4, which will work when the starting point is M3. Further diminution cuases the eventual creation of the route M3 - M4, which is actually an improvement over the original in terms of efficiency, since it can get to the presumably desirable state M4 without bothering to go through M5 and doubling back through M3. Note that in light of this new schema, the initial five-step route can be viewed as 'superstitious' behavior; i.e., the model is acting as though it 'thinks' that just because it went through M5 to get to M4 the first time, it must do so on subsequent trials. It is crucial to note that efficiency is not always best; in fact, mammals can be trained to repeat long sequences of otherwise 'superstitious' behavior, as long as that behavior is rewarded, while any variations go unrewarded (see e.g., *[Hilgard and Bower 1970]*).

4.1.4 Coincidence-based trace establishment

There are two cases of 'coincidence' that can arise in the model: either an existing schema gets REMINDed during the COLLECTion of a new schema, or a schema gets REMINDed during the ENACTing of another schema that has been previously REMINDed and ACTIVATEd.

Case six: Append-driven learning

If the model is COLLECTing a new schema that leads to, say, an undesirable result, such as an unpleasant taste, that NOTICEd taste may simultaneously cause a REMIND of, say an innate 'gag reflex' schema (i.e., it says to spit out after sensing a bad taste). In such a case, the INDEX (and DETOUR) operators create index links to both the sequence of events leading up to the bad taste, so that it might be avoided in the future, and to the sequence of events REMINDed by the event, so that this sequence might be substituted for the undesirable sequence the next time it happens; this is an instance of an 'active avoidance' situation.

Case seven: Splice-driven learning

If the model is currently ENACTing an active schema, e.g., running a maze toward a food reward, and during this, another schema gets REMINDed (e.g., a light flash that is known to lead to some different reward), then both schemas are indexed together by the same initiating feature, giving that feature added predictive power.

4.2 Note: Design decisions affecting CEL-0's performance

It should be noted here that a number of design decisions in CEL-0 (including the specifics of the ACTIVATE preference metrics, the details of the built-in exploratory schemas, and the details of the functions of the operators, notably SYNTHESIZE, REMIND and ACTIVATE) will affect the path it will take through the maze, and in many cases will affect whether or not the correct learning will take place at all. We have been experimenting with versions of CEL-0 to see which changes cause which behaviors, but we intend to continue to compare the resulting behaviors against the learning literature wherever possible (see esp *[Rescorla and Wagner 1972]*), and to suggest new experiments (and their predicted outcomes) when the literature doesn't provide the necessary data on some specific point about how a rat, for instance, should run the maze. Section 5 of this paper makes some brief remarks about our use of some results in animal learning as a 'requirements specification' for CEL-0's performance; *[Granger and McNulty 1984]* contains more discussion of this.

5. The neurobiology of multiple memory systems

5.1 The constellation of deficits in the amnesic syndrome

The patient H.M., like most other amnesics, exhibits a whole constellation of related deficits. The key deficit is the inability to consciously store new information, as described earlier in this paper. Two of the other major components of the overall amnesic syndrome are:

- *Retrograde amnesia:* H.M not only is incapable of consciously storing new information *since* his operation; he also has lost some of the memories that happened to him immediately *preceding* the operation, up to about two years before the operation, while memories older than that remain unimpaired. This striking finding *[Squire 1980]* is used as evidence that memory *consolidation* takes time (perhaps up to two years) before it becomes a permanent part of memory; hence, perhaps memories that were still being consolidated at the time of the operation were disrupted, and never got firmly established as permanent memories.

- *Rapid Forgetting:* H.M. is able to carry on conversations, and perform other tasks of long duration, *as long as the task isn't interrupted;* when interrupted for more than a few minutes, he completely forgets where he was, and starts over again 'from scratch', e.g., he might then have the exact same conversation all over again without realizing he's just done it.

5.2 Of rats and men

There are recently-discovered situations in which rats in a maze exhibit forms of learning previously only attributed to primates and humans. *'Learning-set learning'* refers to very rapid (usually just a single trial) learning of new situations that are similar to previously-learned ones, i.e., the animal seems to form a *'template'* that it can use to expidite the learning of subsequent situations. The rats' learning-set learning (*LSL*) system apparently is entirely separable from its more standard, slower 'associative learning' (*AL*) system – there are specific drugs and lesions that have been used to entirely eliminate abilities associated with the *LSL* system without affecting the performance of the *AL* system, and vice versa. This constitutes evidence that rats have multiple memory systems.

Furthermore, recent experimentation *[Staubli and Lynch 1989]* has shown that rats can be given amnesic symptoms strikingly similar to those in humans, by making corresponding lesions to the hippocampus and another limbic structure, the thalamus. In particular, rats are trained to select one of two odors for a water reward. This initially requires 50-100 trials before a minimal criterion of learning is met (i.e., associative learning (AL)). Over successive pairs of odors, the rats' behavior changes such that they come to learn the correct odor in subsequent odor-pairs in only 3-4 trials (learning-set learning (LSL)). Two forms of learning are thought to be involved: (1) abstract 'template-driven' (LSL-type) information about the task (e.g., the fact that it contains a 'correct' and an 'incorrect' olfactory cue), and (2) specific memory (AL-type) as to which particular odor was correct for a given pair.

One specific type of lesion (lesions of the connection between the dorsomedial nucleus (DMN) of the thalamus and the frontal cortical system) eliminates the animals' ability to go from the many-trial (AL) mode to the subsequent rapid-learning (LSL) mode over successive pairs of odors. This suggests that the rats are learning the *specific* memories for correct odors, but are failing to learn the *template* information about the existence of correct and incorrect odors in each pair.

Disconnection or lesions of the hippocampus, on the other hand, produces an apparent inverse of this result, with a time-dependency as well: the rats acquire the rapid learning mode (i.e., they appear to learn the abstract correct-incorrect information), but for any given pair of odors they cannot recall the right specific odor (i.e., cannot perform the task) if delays of more than about 5 minutes are interposed between trials (i.e., a deficit similar to rapid forgetting). Hence it seems that these rats are acquiring the abstract memory, but are failing to create a long-term trace of the specific memory.

5.3 Interpretation of the data

What ability, i.e., what specific knowledge or process, is available to the rat the LSL situation, but not in the AL situation, to enable template-driven learning? The problem for CEL (or for any other model of learning and memory) in attempting to provide a consistent account of these results, is that apparently the templates are learned but the specific memories leading to those templates are lost. We do not have a complete solution, but we have come up with a set of opposing hypotheses, either of which could potentially explain the data. These opposing hypotheses have been used to

design an experiment that is currently being run to help further clarify the the question, and to narrow down the set of possible consistent models of these two learning systems.

In the language of the CEL model, there are two classes of possible explanations: (1) the hippocampal (AL) losses are due to a 'storage-side' failure to either COLLECT or INDEX the specific information, or else (2) these losses are due to a 'retrieval-side' failure to correctly *use* the specific odor memory to find the water reward; i.e., perhaps REMIND finds both the template memory and the specific-odor memory, but ACTIVATE is not correctly using the specific-odor memory to instantiate the template memory in order to find the reward.

The articulation of these two opposing possibilities has suggested an experiment to try to test whether the specific odor was in fact present in memory at all. The memory seems not to show up in the odor-choice situation, but if explanation (2) above is correct, then the memory may be there but just not being used correctly in that situation. It turns out that there is a relatively simple experimental methodology for testing for 'raw memories' like this. Details are provided in *[Granger 1989b]*, but briefly, the experiment allows us to see whether a rat has any memory of a particular event (such as a specific odor) or no memory of that event. That is, the rat's behavior in the presence of some previously-seen event can be reliably distinguished from its behavior in the presence of an unrecognized event; hence, we should be able to tell whether or not the specific odor is in memory or not. This experiment, described in *[Granger 1989b]*, is currently being run at the Center for the Neurobiology of Learning and Memory at Irvine.

If it turns out that the memory shows up in this experiment, then we may hypothesize that the deficit is on the retrieval side, that is, the memory is present, but it cannot be correctly used to perform the choice behavior. In CEL terms, it is possible that ACTIVATE cannot instantiate the memory into the template that can use it to find the water reward. If, on the other hand, the rats exhibit no recognition of the specific odors, we will hypothesize that the deficit may indeed be a storage-side deficit, and we will have to attempt to alter the model to account for the loss of a spcific memory after the creation of a template from it.

Either way, the CEL model will have aided in suggesting a key experiment that can deicde the question of whether the rapid-forgetting phenomenon is a storage-side or a retrieval-side deficit. This brings us a step closer to an understanding of the nature of the multiple (LSL and AL) learning and memory systems.

6. Conclusions: Artificial and natural learning mechanisms

There exist many theoretical questions in learning and memory that rely on the consistent interpretation of an almost bewildering array of interrelated experimental results. The field of multiple memory systems is one particularly exciting current example of this; a battle over the characteristics of these systems is currently raging among memory researchers in the neurosciences *[Squire 1980; Squire, Cohen and Nadel 1982; Mishkin 1982; Tulving 1984]*.

The search for consistent interpretations of these data can be aided by artificial models of learning and memory, and, reciprocally, the development of consistent models can be furthered by the experimental testing of the models' predictions against natural learning systems. While it is not necessary for an artificial learning system to precisely account for all available psychological data on learning, it has happened time and again in AI that sincere attempts to provide consistent interpretations of problematic psychological results have resulted in both better psychological theory and richer and more productive computer systems.

There are certain specific processing problems that any learning system, natural or artificial, must have a way of solving. We are trying to characterize some of those processing problems in specific learning situations, in hopes of identifying the similarities among, and differences between, different instances of learning systems. The CEL model has so far been helpful in identifying and clarifying some of the possible theoretical interpretations of results in the area of multiple memory systems. We hope that by continuing to iterate the loop from theoretical suggestion to experimental result and back, we can further refine and narrow down the range of possible interpretations of multiple learning and memory systems, so that the study of artificial and natural learning mechanisms can productively use each others' results.

References

Carbonell, J., "Derivational Analogy and its Role in Problem Solving," in *Proceedings of the 1983 National Conference on Artificial Intelligence*, 64-69 , 1983.

Granger, R.H., "Identification of Mental Operators Underlying the Acquisition of Simple Predictive Behavior.," *Department of Computer Science Technical Report 191*, University of California, Irvine, November, 1982.

Granger, R.H., "Identification of Components of Episodic Learning: The CEL Process Model of Early Learning and Memory.," *Journal of Cognition and Brain Theory*, 6, 1 (February 1983a), 5-38.

Granger, R.H., "An Artificial-Intelligence Model of Learning and Memory that Provides a Theoretical Framework for the Interpretation of Experimental Data in Psychology and Neurobiology," *Department of Computer Science Technical Report 210*, University of California, Irvine, September, 1983b.

Granger, R.H., and McNulty, D.M., "The CEL-0 system: Experience with a computer model that learns to run a maze," *Department of Computer Science Technical Report 220*, University of California, Irvine, March, 1984.

Hilgard, E.R. and Bower, G.H., *Theories of Learning*, New York: Appleton, 1966.

Kolodner, J.K., "Maintaining Organization in a Dynamic Long-Term Memory," *Journal of Cognitive Science*, 7, 4 (1983a), 243-280.

Kolodner, J.K., "Reconstructive Memory: A Computer Model," *Journal of Cognitive Science*, 7, 4 (1983b), 281-328.

Langley, P., Bradshaw, G.,and Simon, H.A., "BACON.5: The discovery of conservation laws.," in Proceedings of the Third International Joint Conference on Artificial Intelligence, 121-126 , 1981.

Lebowitz, M., "Generalization and memory in an integrated understanding system," *Department of Computer Science Technical Report 186*, Yale University, New Haven, Conn., 1980.

Mishkin, M., "Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus," *Nature*, 273, (1978), 297-298.

O'Keefe, J., and Nadel, L., *The Hippocampus as a Cognitive Map*, Oxford: The Clarendon Press, 1978.

Schank, R.C., "Failure-Driven Memory," *Journal of Cognition and Brain Theory*, 4, (February 1981), 41-60.

Simon, H.A., Langley, P. and Bradshaw, G., "Scientific discovery as problem solving," *Synthese*, 47, (1981), 1-27.

Slotnick, B., "Olfactory learning in rats," *Science*, 185, (1974), 796-798.

Squire, L., Cohen, N., and Nadel, L., "The medial temporal region and memory consolidation: A new hypothesis," in *Memory Consolidation*, Hillsdale, NJ: Erlbaum Assoc., 1982.

Squire, L.R., "Two forms of human amnesia: an analysis of forgetting," *Journal of Neuroscience*, 1, (1981), 635-640.

Tolman, E.C., *Purposive Behavior in Animals and Men*, New York: Century, 1932.

Tulving, E., "Episodic and semantic memory.," in *Organization of Memory*, E. Tulving & W. Donaldson (Eds.), New York: Academic Press, 1972.