LEVELS OF GOAL DIRECTION AND THE CAUSES OF LEARNING

Dale M. McNulty

Irvine Computational Intelligence Project
and
Center for the Neurobiology of Learning and Memory
University of California
Irvine, CA 92717

ABSTRACT

A general purpose model and learning program are described which account for the phenomena of latent learning and irrelevant-incentive learning. The model is composed of three separate goal pursuit levels. At the lowest (latent) level are the constant, implicit goals associated with the system's memory management mechanisms. At a higher (overt) level are the dynamic, explicit behavioral goals which the implicit goals enable by manipulating memory representations to conform to the external surroundings. At the highest (meta-overt) level the latent level employs knowledge as meta-schemata. Functional specifics which enable the program to demonstrate the learning behavior in a maze environment are discussed.

1.0 Introduction and Relevant Historical Perspective

Prior to the 1930's psychologists associated learning with performance. For Pavlov, Thorndike, Hull, Spence, etc. learning was a reinforcement of stimulus-response associations. In 1924 Simmons observed a contrary phenomenon. Unrewarded rats allowed to explore a maze were just as capable of finding their way to food during subsequent rewarded trials as animals which were rewarded continuously. Blodgett (1929) termed this learning "latent" because the animals acquired maze knowledge in the absence of any demonstration of learning. Blodgett's conclusion contradicted the firmly entrenched Law of Effect, and Hull's theory of drive reduction, which require behavioral reinforcement to strengthen the S-R connections of learning. Early cognitivists (Tolman and Honzik, 1930) explained Blodgett's conclusion in terms of cognitive knowledge of goals and goal acquisition. The vast majority of modern theorists continue to consider learning to be goal directed.

The split from earlier S-R theory is a result of two important observations from the latent learning experiments (Thistlethwaite, 1951). First, reinforcement, such as food or water, is not necessary for learning to occur. Second, learning should be distinguished from performance. Knowledge need not be manifest for learning to occur. Information is often 'known' and not demonstrated until the proper motivation is presented. Bearing these distinctions in mind, it is possible to recognize two different paradigms which need to be explained by learning models. One type of experiment utilizes unrewarded trials, or exploratory behavior, to test the subject's ability to learn how to arrive at a location that has been associated with a reward. Experiments of this type are testing latent learning. A second type of experiment tests the subjects' ability to learn how to arrive at the location of an object that was irrelevant at the time of exposure. An example of an irrelevant object is food which has been rendered irrelevant by feeding subjects to the point of satiation. Experiments of this sort test irrelevant-incentive learning.

A program capable of both types of behavior will be described here. For the sake of simplicity, these two experimental paradigms will only be distinguished in this discussion where specific program operations distinguish the two.

2.0 Levels of Goals and the Causes of Learning

Since learning does occur in the absence of reinforcement and without associated performance, previous learning theories relying on reinforcement are not satisfactory. In formulating new models the obvious question to ask is: what are the causes of learning? Four possible explanations come to mind:

[1] Learning is not goal directed. This position is difficult to conclusively prove or disprove, however the nonrandom nature of learning suggests that learning is either goal directed, or, that animals attend to and store every experiential event. The later does not seem supportable, so it must be concluded that learning is, in some way, associated with goals.

This research was supported in part by the National Science Foundation under grant IST-81-20685 and the Naval Ocean Systems Center under contracts N00123-81-C-1078 and N66001-83-C-0255.

[2] Subjects in the latent learning and irrelevant-incentive learning paradigms receive reinforcement in ways other than feeding or drinking and, thus, the Hullian Model of stimulus-response adequately explains the phenomena. A number of authors have argued that the rats learn the maze because they are driven by such things as a need for handling or a desire to escape the maze. The experimental results are very complex and the possibility for many types of reinforcement can not be totally ruled out, but as Thistlethwaite shows, reinforcement does not adequately explain the entire latent learning effect. Many others, including Mackintosh (1974) and Bower and Hilgard (1981), have also effectively dismissed the reinforcement arguments, thus, the consensus is that latent learning is a real phenomenon which is not explained by the Law of Effect or reinforcement theory.

[3] A modified version of the Hullian Model of stimulus-response explains the latent learning phenomena. This position is not supported by the evidence. Again Thistlethwaite, Mackintosh, and others effectively argue against this. The S-R model of learning is generally not accepted today, so the argument will not be carried on here.

[4] Learning is goal based, but there are multiple levels of goal pursuit, some of which have not been previously recognized. (Goal is used here to mean objective, including the pursuit of an endeavor, such as Hull's drive reduction.) At one level the organism pursues goals such as eating and drinking, and at another level the pursuit of a different type of goal explains phenomena such as latent learning and irrelevant-incentive learning. At least one publication has speculated on the possibility of another, less obvious, level of goal pursuit. O'Keefe and Nadel (1978) argued, without detailing the mechanisms whereby it occurs, that rats are driven to create cognitive maps.

The most plausible explanation for learning seems to involve multiple levels of goal pursuit, but if multiple levels of goal direction do exist, then the earlier question can be rephrased to ask: what are the various goal levels associated with learning?

The CEL model (Granger, 1983a, 1983b) proposes three levels of goal pursuit. At the lowest level, innate, latent mechanisms semi-autonomously pursue constant goals which create and manipulate memory structures in such a way that knowledge and, thus, performance, eventually conform to the external environment. The goals at this level are functionally invariant, but the structures that the mechanisms operate on vary dynamically with the goals of the higher, overt level which change in response to the varying needs of the organism. This distinction between functional invariance with dynamic targets at the latent level and dynamic goals at the overt level can be likened to a packing machine in a soda factory. One day the machine may be employed to place orange soda in the cartons. The next day, depending on the (overt) needs of the plant, cherry soda may be packed into cartons with different labels. The machine continues from day to day performing the same function (packing 24 bottles at a time into a carton), but using the bottles and boxes targeted by the higher, overt level of the plant. The overt goals of the plant vary in the same way as the overt goals of the organism change to reflect a dynamic environment and the way the organism must respond to it. If the animal has been starved, the latent mechanisms assist the animal in finding food. If it has been deprived of water, the latent mechanisms still search and match data, but now, instead of matching memory structures that pertain to food, the mechanisms will be concerned with knowledge related to water.

The third level of goal pursuit accounts for still higher level goals such as reasoning, planning, and possibly learning set learning. At this level the organism manipulates low level knowledge with other low level knowledge. This level is called the meta-overt level because the low level knowledge serves a meta role in the acquisition of overt goals.

3.0 Examples of Latent Mechanisms

3.1 Introduction to the Program

CEL-0 is a program embodiment of the CEL (Components of Experiential Learning) model of learning and memory. The program, like the model, consists of twelve parallel and semi-independent primitives which build and manipulate hierarchical memories. These operators collectively perform five classes of data manipulation on memory representations termed episodic schemata. In brief, the twelve operators have the following functions:

Reception Operators
DETECT - sensory events from the various modalities (sight, touch, etc.).
SELECT - ively attend to sensory inputs on the basis of past experience.
NOTICE - input which matches the desirable and undesirable states and trigger COLLECT.
COLLECT - sensory events into packets (episodic schemata) for memory storage.
INDEX - schemata into the memory hierarchy for future reference.

Retrieval Operators
REMIND - the system of past experiences which are related to the present situation.
ACTIVATE - REMINDed schemata based on predictive values of schemata and behavior desired. ACTIVATE triggers reconstruction.

Reconstruction Operators
ENACT - the appropriate efferent actions in the active schema; tune SELECT to attend to predicted afferent events.
SYNTHESIZE - new schemata by matching inputs against predicted events; trigger refinement operators to modify schema based on matches or mismatches.

Refinement Operators
REINFORCE - strengthen successful schemata to reflect their predictive success.
BRANCH - within a schemata to indicate mismatches
DETOUR - within a schemata to indicate branches to be avoided.

In the maze environment these mechanisms combine to enable CEL-0 to demonstrate a latent learning ability enable CEL-0 to demonstrate to demonstrate a latent learning ability similar to rats. A run-time listing of CEL-0's maze performance is available in [McNulty and Granger, 1985].

To understand CEL-0's operation it is convenient to divide its activity into two functionally distinct classes: the acquisition of knowledge and the use of that learned knowledge (i.e. performance). This distinction, however, is completely artificial. CEL-0 is constantly learning, and thus, it is difficult to distinguish between CEL-0's learning and performance, but the distinction is useful for purposes of analysis. A detailed explanation of how CEL-0 performs in the maze is provided in [Granger and McNulty, 1984a, 1984b] and [McNulty and Granger, 1985]. This discussion will focus on the latent roles that some of the operators play and how they interact with the overt level.

3.2 The Role of Latent Mechanisms in the Acquisition of Knowledge

During the exploratory phase of the latent learning experiment CEL-0 is satiated and, thus, possesses no overt goal. When first placed in the maze CEL-0 'knows' nothing of the environment itself. It begins with four innate schemata (move forward, turn left, turn right, and turn around) and knowledge of when those schemata apply. CEL-0's knowledge of the maze is constructed from the sensory input it encounters as it negotiates the maze.

3.2.1 The NOTICE Operator

NOTICE has the innate goal of comparing afferent sensory events with events on the Desirable State List(DSL) and the Undesirable State List(USL). The DSL and USL are dynamic structures which change as CEL-0's overt goals change in response to the ambient environment and innate needs (e.g., hunger, thirst). Any match indicates that CEL-0 has experienced something of note and should store away for possible later use the experiences which lead up to that event.

3.2.2 The COLLECT and INDEX Operators

COLLECT and INDEX are the operators responsible for creating episodic schemata and placing them in memory for later use. COLLECT groups the most recent sensory events as a schema, places the schema in the memory hierarchy, and calls INDEX to create pointers to the new piece of knowledge so that it is recallable. Because the pointers are the only links between present and past, it is critical that the correct attributes be indexed. INDEX uses the first and last sensory events of the episode as pointers. The former allows CEL-0 to find episodes related to current sensory events. The latter, called the goal index, enables CEL-0 to find schemata associated with overt goals. In the latent learning paradigm two circumstances cause the creation of new memories (see [Granger and McNulty, 1984] for a more complete discussion):

a)SYNTHESIZE detects a successful completion of an ACTIVATEd schema

b)a NOTICE of an event on the DSL or USL

INDEX is vested with an additional goal which plays a very important part in CEL-0's ability to improve its performance during rewarded trials. A COLLECTed schema is composed of seven sensory events representing the most recent CEL-0 experiences. Frequently, only a subset of that group is pertinent to the overt goal which caused the schema to be laid down in memory. The events not pertinent to the goal are extra baggage which CEL-0 may, but probably not, find useful during performance. INDEX helps the refinement of performance by creating many schemata from the one original. Each newly created schema is a sequential subset of the original events CEL-0 experienced. This function is called 'sensitivity analysis' since the net effect is to test the individual sensory events for relevance to the outcome of the schema.

3.3 The Role of Latent Mechanisms in Performance

During initial trials in the maze CEL-0 acquires knowledge of the maze as it 'explores' with no overt goal. Exploration is driven by CEL-0's latent goal level where the innate mechanisms autonomously go about satisfying their goals. In subsequent trials CEL-0 is made hungry and food is placed at one end of the T-maze. During the rewarded trials, CEL-0 must:

a)associate the need (satisfy hunger) with the reward

b)associate the location of the reward with stored schemata

c)ACTIVATE the appropriate schemata

d)Improve performance by refining learned schemata into more appropriate new schemata

3.3.1 The NOTICE Operator

On the first rewarded trial CEL-0 proceeds as on previous trials because it has no knowledge of the reward until it is encountered. Once detected, NOTICE initiates a COLLECT and INDEX which results in episodic schemata indexed by and terminating in 'the sight of food'. If allowed to eat, NOTICE will also COLLECT and INDEX schemata which result in 'consumption of food'. The distinction between these two schemata is important to CEL-0's performance in the irrelevant-incentive learning experiments because schemata resulting in the 'sighting of food' when satiated must be associated with schemata resulting in the 'consumption of food' when hungry.

3.3.2 The REMIND Operator

REMIND's goal is to match afferent sensory events and goals against previously stored indices. A match means that a current sensory event is related to a previous experience.

3.3.3 ACTIVATion of Schemata

REMINDed schemata might be predictions of future experiences. ACTIVATE's job is to choose the best possible predictor (schema) given CEL-0's current goals. ACTIVATE proceeds autonomously applying metrics of comparison and matching schemata results to goal relationships. The goals it matches against are the dynamic goals communicated from the overt goal level.

3.3.4 ENACTment of Efferent Actions

Episodic schemata are composed of two types of events. One type of event results from environmental sensory data which CEL-0 has encountered. The other type of event represents the sensory data CEL-0 has perceived of its own actions. CEL-0 responds to its environment, and as a result perceives new sensory data which represent a new situation in the environment. CEL-0 must now respond to that new situation. Thus, for a schema to be an accurate predictor of an outcome, CEL-0 must usually effect actions comparable to those contained in the stored schema. ENACT's goal is to search the active schema for these efferent acts and initiate them.

3.3.5 The SYNTHESIZE Operator

SYNTHESIZE has responsibility for constructing new schemata from previously stored episodes. It does this in concert with REMIND and ACTIVATE. Recall that new sensory inputs are constantly causing REMINDs of previous experiences and ACTIVATE is selecting schemata which might serve as tools for achieving overt goals. The ENACTed efferent events cause new sensory events and the REMIND/ACTIVATE process continues. As CEL-0's location in the maze changes, ACTIVATE may select a more appropriate schema over the one that is currently active. The new schema will have new efferent events which will cause new sensory stimulation. All of these sensory events, new and old, will be incorporated by SYNTHESIZE into new schemata which will be COLLECTed and INDEXed. In this way, new, refined schemata are constructed from previous experiences.. These new schemata can, in turn, be further refined by the same process during later trials in the maze. [McNulty and Granger, 1985] discusses this in more elaborate detail and shows how these mechanisms work with sensitivity analysis to produce a stepwise refinement of behavior.

3.3.6 REINFORCEment

ACTIVATEd schemata are predictions. Some predictions prove to be successful and some do not. A successful predictor is more useful than an unsuccessful predictor and, as such, deserves special status among CEL-0's stored experiences. When SYNTHESIZE detects the successful completion of an ACTIVATEd schema, REINFORCE is triggered to mark that schema as successful. ACTIVATE will use that strengthening as one of its metrics when selecting the best available predictor.

4.0 Conclusions

Experiments from the psychology literature suggest that some types of learning can not be explained by the unidimensional goal direction many previous learning theorists have employed. Two such types of learning are latent learning and irrelevant-incentive learning. A multidimensional system involving three levels of goal direction seems a more reasonable explanation of learning. The lowest level is the latent level where semi-autonomous mechanisms operate to map the external environment onto the internal data structures such that the organism's behavior conforms to the external surroundings. The goals at this level are implicit, innate and, thus, unchanging. At a higher level, and working in concert with the latent mechanisms, is the overt level where goals are dynamic, reflecting the changing desires and changing environment the organism finds itself in. At the highest level, existing knowledge structures (i.e. episodic schemata) are used as meta-schemata to pursue overt goals. Typical functions at the meta-overt level are: reasoning about prior experience (such as deduction and induction), planning, understanding, selection of schemata for activation, and 3-D rotation of spatial knowledge. These three levels working in concert enable simple structures to exhibit complex behavior.

REFERENCES

[1] Blodgett, H.C., "The Effect of the Introduction of Reward Upon the Maze Performance of Rats", Univ. Calif. Psychol., 4:113-34.

[2] Bower, G.H. and Hilgard E.R., Theories of Learning, Englewood Cliffs, Prentice-Hall, 1981.

[3] 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, pp.5-38, February, 1983a.

[4] 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.

[5] 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, 1984a.

[6] Granger, R.H. and McNulty, D.M., "Learning and Memory in Machines and Animals: An AI Model That Accounts for Some Neurobiological Data", in Proceedings of the Sixth Annual Conference of the Cognitive Science Society, pp. 161, 1984b. (also in: Memory , Experience, and Reasoning, Kolodner, J.L. and Riesbeck, C.K. (eds.), Erlbaum and Associates, Hillsdale, NJ, to appear 1985).

[7] McNulty, D.M. and Granger, R.H., "Elements of Latent Learning in a Maze Environment", Information and Computer Science Technical Report #85-03, University of California, Irvine, 1985.

[8] Mackintosh, N.J., The Psychology of Animal Learning, Academic Press, London, New York, and San Francisco, 1974.

[9] O'Keefe, J. and Nadel, L., The Hippocampus as a Cognitive Map, Clarendon Press, Oxford, 1978.

[10] Simmons, R., "The Relative Effectiveness of Certain Incentives in Animal Learning", Comp. Psychol. Monogr., 1924, 2, 1-79.

[11] Tolman, E.C. and Honzik, C.H., "Introduction and Removal of Reward, and Maze Performance in Rats", Univ. Calif. Publ. Psychol., 4:257-75.