Conditioning and Concept Formation in Embodied Agents

Terry Stewart
Institute of Interdisciplinary Studies:
Cognitive Science
Carleton University
Ottawa, K1S 5B6, Canada
tcstewar@chat.carleton.ca

Sharon Wood
School of Cognitive and Computing Sciences
University of Sussex
Brighton, BN1 9QH, UK
sharonw@cogs.susx.ac.uk

Abstract

Learning algorithms typically model the acquisition of conceptual knowledge from some start state to some fixed learned end state. Natural associative learning demonstrates a more comprehensive range of processes which complement this static view of learning. An experimental regimen is presented for evaluating learning algorithms against this wider remit. This approach provides a general basis for analysing performance and measuring concept formation. We use it here to examine the Distributed Adaptive Control (DAC2) model.

Introduction

Various techniques lend themselves to modelling particular aspects of learning, from sensorimotor concepts (Thelen and Smith, 1994) to language (Anderson, 1983). Invariably, that part of learning modelled is acquisition: the process of acquiring 'new' concepts or knowledge, to the extent that concept acquisition and learning are often viewed as synonymous.

Here, we make the case that psychology provides us with a much richer view of learning, in which acquisition plays a key part, yet which is balanced by complementary processes which actively modify previous learning. These other processes, it might be argued, are essential for maintaining the long term viability of a learning entity in a complex and dynamic environment.

Although the strong case needs to be made, it would seem to be productive to consider how we decide whether, and in what way, a given learning algorithm comprehensively satisfies all those aspects of learning revealed from a psychological perspective.

Background

The work described here is based on a review of neural models of associative learning, with particular focus on two features: the ability to form 'concepts' by extracting patterns from the sensory data, and the capability of dealing with embodied agents (Stewart, 2000a). These features are not readily apparent in any one model, but the architecture of these models is such that they may be able to complement one another in a combined system. The key models investigated were Theory of Neuronal Group Selection (Edelman, 1987), Rectified Gaussian Belief Networks (Hinton *et al*, 1997), and Distributed Adaptive Control (Verschure, 1998).

The methodology proposed is to subject a given embodied implementation of a learning agent to an evaluation of its observable behaviour, in order to determine whether its performance evidences the typical characteristics of human and animal learning. Identifying where performance conforms and/or deviates from paradigmatic patterns of behaviour provides a basis for identifying those aspects of natural learning not yet encompassed by the learning algorithm.

This might then go on to suggest either where appropriate modification of the algorithm could lead to improved conformity; or why the learning algorithm is, in principle, unable to support the full spectrum of typical behaviour.

Furthermore, this approach provides a context for evaluating concept formation. The ability to perform some consistent action whenever confronted with a situation that the observer would identify as a particular concept, is something that can be tested experimentally. If the response of an embodied agent demonstrates that it can differentiate between situations which reflect the absence or presence of this underlying concept, then one can argue that the agent has formed a representation of that concept.

Learning Processes

The learning paradigm selected for initial application of this methodology is that of Classical Conditioning. The arguments for this (more fully documented in Stewart, 2000b) are that it is an example of natural associative learning found in even the simplest of organisms.

Classical conditioning is, in essence, the forming of associations. Initially, a certain stimulus (*unconditioned*-US) is associated with a certain response. The experimenter then presents a new stimulus (*conditioned*-CS) at the same time as the initial stimulus (US). After a few repetitions, the subject will give the response when presented with the new stimulus (CS) by itself.

The simplest rule for forming these associations in an artificial learning system is Hebb's rule. With this, the connection between any two components in the system should be strengthened if both of those components are active at the same time. This rule would therefore seem to be an appropriate candidate for a learning algorithm emulating associative learning in the classical conditioning sense.

Let us consider the stimulus of a ringing bell which activates a node in the network at the same time as a presence-of-food node is active. Initially, the food node is strongly connected to the salivation node (perhaps by genetic design, or perhaps by previous learning), causing that node to become active. With the bell node active at the same time as the presence-of-food node, the connection between them would increase. If this happens a few times, then the new connection will become strong enough to activate the salivation node by itself.

Note that this will only happen if the association is a regular occurrence. Thus, this sort of association formation is actually an identification of regularities in the environment. This observation leads us to believe that a real model of classical conditioning will include concept formation/detection as an integral part.

In this overly simplified example, the concepts that the agent is creating are directly specified in its sensory data. A more complex learning algorithm would be needed to allow the agent to handle more subtle concepts, where those features of its environment that are relevant are not neatly presented to individual sensors. In real life, relevant data is distributed across many sensors and multiple sensory modalities.

Also it is observed that many learning algorithms (including the one discussed above) generally start from an uninitialized (or random) start state, then proceed to some final ‘learned’ state via acquisition, stopping there. It is not clear that, as a learning process, this is sufficient. In particular, what may be deemed appropriate associations are likely to change over the lifetime of an individual, rather than remaining fixed or static. In contrast, classical conditioning includes the ability to un-learn associations that are no longer valid, and to learn to discriminate between initially similar patterns. It is certainly arguable that these features are necessary for a creature to deal with a complex and changing environment.

Experiments with an Embodied Agent

The methodological approach is demonstrated through its application to the reimplementation of Distributed Adaptive Control (DAC2). The roots of DAC2 start in a paper by Verschure and Coolen (1991), where they develop a connectionist model based on classical conditioning.

What makes this model relevant is that it is not limited to the sort of local representation found in the purely Hebbian example above. The stimulus patterns between which associations are formed can be completely arbitrary, thus freeing the system to respond to non-trivial concepts. Furthermore, the model has been designed to exhibit the classical conditioning features of blocking[1] and second order conditioning[2].

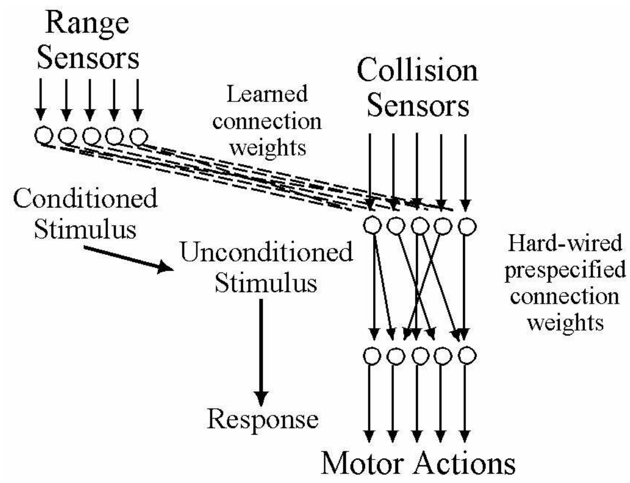

Figure 1: DAC2 Architecture

The actual DAC2 learning algorithm is an approximation to the theoretical model, so as to be implementable in a physical robot.[3] However, it does seem to inherit the major properties of the theoretical models.

In the final DAC2 architecture, there is a clear distinction between unconditioned sensors, which are hard-wired to elicit a particular response, and conditioned sensors, which become associated with particular unconditioned stimuli. In the standard implementation, the unconditioned sensors detect physical collisions between the agent and its environment, and the conditioned sensors detect the range from the robot to solid obstacles in various directions. Initially, the robot uses only its collision sensors to perform a hard-wired movement task by moving forward if there is no current collision, and turning away if there is a collision.

At the outset, the system does not use its conditioned sensors, as no associations have been formed. However, as the agent explores its environment, associations are slowly formed. In particular, if the agent forms an association between nearby obstacles and the activation of its collision sensors, then it should (and, indeed, does) turn when near an obstacle, even though the collision sensor has not been directly activated. The pattern in the conditioned sensor creates the associated pattern in the unconditioned sensor, leading the robot to turn. The model is also capable of handling multiple types of CS and US sensors.

This system was re-implemented and validated, and a series of experiments were performed to compare its associative learning characteristics to those of classical conditioning (Stewart, 2000b).

[1]When multiple new stimuli are being learned, the learned association strengths between the new stimuli and the unconditioned stimuli are not independent of each other. Instead, the one new stimulus with the most predictive power (i.e. the one that is experienced most often with the unconditioned stimulus) keeps its strong association with the unconditioned stimulus, and the other stimuli get a lower association than they would normally.

[2]The association of a second CS with the first CS resulting in the extension of all the properties of acquisition to the second CS

[3]A complete description of the actual DAC2 architecture and its more recent extensions, with particular focus on its connection to classical conditioning, can be found in (Verschure and Voegtlin, 1999) and (Voegtlin and Verschure, 1999).

The key question at hand was to determine if DAC2 is an effective model which explains associative learning in a manner consistent with classical conditioning. The original DAC2 experiments had shown only that DAC2 was capable of a particular aspect of classical conditioning. This aspect, the ability to form an association between a new stimulus pattern and an old stimulus pattern (*acquisition*), is at the core of classical conditioning, but it is not the full story. A hundred years of psychology experiments have uncovered many other quirks and characteristics of classical conditioning, each of which may well be vital to explaining the effectiveness of this style of learning. For example, the ability to acquire associations may not be very useful without the ability to have these associations extinguish over time if the learned pattern is no longer applicable.

These tests can be divided into two parts. The first part explores DAC2's concept formation abilities: the sorts of concepts (i.e. regularities) it is able to develop. The second part examines the more standard aspects of classical conditioning: acquisition, extinction, generalization, and discrimination.

Concept Formation

The study of concept formation is not particularly common in associative models. In most models of classical conditioning, the concepts do not need to be extracted from the sensor data, since the sensor data directly presents those concepts that the agent is supposed to learn about.

Also, the very process of determining if a concept is formed is problematic. How can we evaluate whether the agent has formed a particular concept? This is a question about mental states and the internal workings of the agent's 'mind'. Even with DAC2's simple architecture, it is not easy to analyse the resulting neural network to determine the presence of various concepts. Indeed, according to the dynamical systems view (*cf* van Gelder, 1998), these concepts may not be found by looking solely at the organism's internals; instead, the outside environment must be examined as well. Thus, we can only answer this question by looking at the actions of the agent within its environment.

The concept formation experiments revolve around a test which operationalizes the definition of a 'concept' in a manner that is based on potential utility to the agent. In other words, it is based on the agent being able to make predictions about its environment. Thus, the determination as to whether a certain concept exists is done by examining the behaviour of the organism. This side-steps the issue of how concepts are represented in the brain by providing a functional test for the concept. In DAC2 we do this by seeing if the agent behaves as if it has a certain concept. To start with a simple concept, how can we test to see if it can learn the concept of "being near a wall"?

One way to do this would be to see if it could form an association between "being near a wall" and some unconditioned stimulus. We can thus change the question from "does it have this particular concept?" to "can it perform some consistent action whenever confronted with a situation that we would identify as a particular concept?" This is a question which can be directly tested through an experiment.

We do this by taking an unconditioned stimulus-response pair. This is a very simplistic S-R configuration, with one neuron for the stimulus, which is directly wired to the one neuron response. This can be thought of in a similar manner to the presence of food causing drooling, or any other standard unconditioned situation. Now, we allow the robot to explore its environment, and whenever it encounters whatever concept we wish it to learn, we activate this new stimulus. For example, whenever the robot comes "near a wall", we can trigger the unconditioned stimulus, which fires the unconditioned response, which we can record. We now run our experiment to see if the robot can learn associations based on this concept. In other words, can it use its other sensors (i.e. the range sensors) to predict this concept of being near a wall?

Of course, the associations it can learn can only be based on the sensor system available to the robot; however, if a regularity in the sensory pattern, consistent with a concept, can be detected by the robot, it can nonetheless be said to have learned that concept. For example, the robot might be able to learn the concept of 'being near blue objects' even without being able to detect colour. This could happen if there is some other regularity in the environment that correlates with the blue objects. If the blue objects are generally set apart from other objects, the robot may be able to seize upon that feature instead.

This observation is not a problem for this set of experiments. Indeed, it merely points out a truth in all experimentation based on concepts: there is no guarantee that the concept the experimentor is using matches perfectly with the concept that the organism is using. This is simply something which should be kept in mind while performing and interpreting these experiments.

To analytically determine if a given concept can be learned, we allow the agent to encounter the concept in the manner described above for a given period of time. Then, we measure the agent's predictive abilities for that concept. Two separate measures are needed here: a positive accuracy (how accurate the agent is at correctly identifying the concept when it should be there) and negative accuracy (how accurate the agent is at determining the concept is not present). If both of these measures are above 0.5, then we can reasonably say that the concept has been learned.

Experiments were carried out for the concepts of "near a wall", "in a corner", and "in a corridor". For the first two concepts, the distance theshold was varied, and for the third concept, the width of the corridor was varied. In all cases other than for very wide corridors, DAC2 was able to successfully learn the concept. However, DAC2 was unable to learn the inverse of these concepts. This difficulty is not uncommon to learning systems of this type, however. DAC2 is closely related to Hebbian learning, and inherits its property of associating large values with other large values. In other words, it cannot form an association between one value being large and another being small. In statistical terms, this style of learning works solely on positive correlations, not negative ones. This is an important limitation on DAC2's

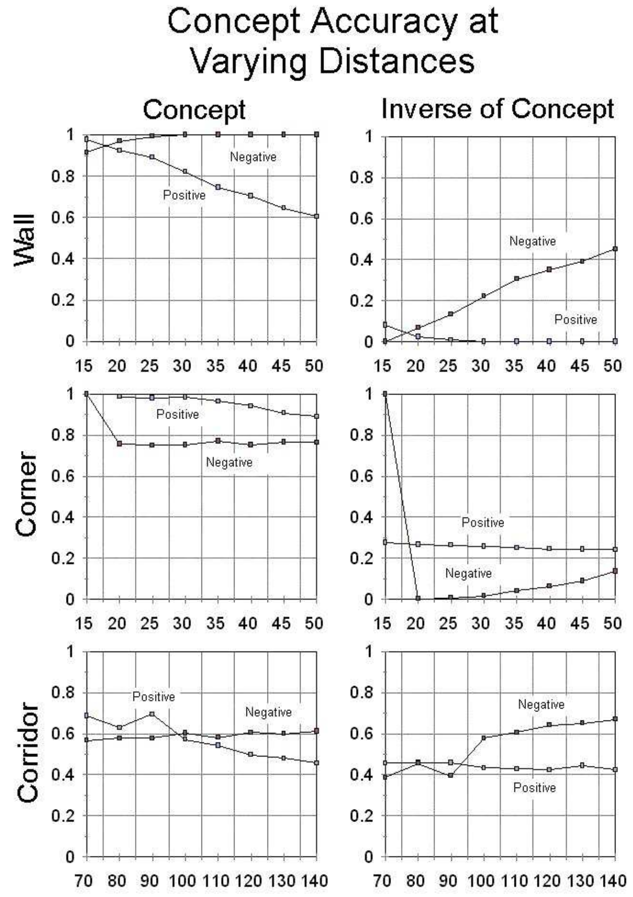

Figure 2: Accuracy Results

abilities, and one not pointed out by the initial research.

In an attempt to determine if this is the only problem with learning these negative concepts, an experiment was done exactly as above, but with the range sensors reversed. That is, their output was set to be 1 minus their normal output (bounded by 0 and 1). This resulted in a vastly improved positive accuracy (approximately 0.95), but poor negative accuracy (no more than 0.3). It seems clear from these results that the sorts of patterns DAC2 is capable of finding are highly dependent on the details of the raw sensory data, and will thus function very differently given different representations of the same environmental regularities.

Features of Conditioning

The further experiments which look at various standard features of classical conditioning are very typical experiments, based directly on those found in undergraduate psychology textbooks (*cf* Atkinson and Hilgard, 1996). The only difference here is that they are being applied to a simulated organism, rather than an organic one. The results are thus directly comparable to the expected results in behavioural science. Deviations from the 'natural' results indicate the areas in which DAC2 does not match well with real classical conditioning.

It is hoped that these tests are not solely applicable to DAC2. It should be possible to run the same set of tests on a different learning algorithm. This could provide a solid basis of comparison for researchers trying to emulate classical conditioning.

Since these tests are to be directly compared to results derived from experiments on living creatures, it is important to make the DAC2 situation as close as possible to the real-life situation. For this reason, some of these experiments do not use stimuli that are patterns to the agent, but rather use a local representation. That is, associations are formed between single sensor nodes. This is because the live experiments do not consider the concept formation requirements of classical conditioning in any way. They consider something like "hearing the sound of a bell ringing" to be a single sensory input.

The acquisition experiments simply presented DAC2 with a single CS and US. By varying the temporal overlap, we were able to examine how the system deals with simultaneous, delayed, and trace conditioning[4]. Unsurprisingly, associations were successfully formed for the first two types, but failed on trace conditioning. In trace conditioning, the CS and US are never present at the same time, and DAC2 was unable to form the connection. Also, the response to delayed conditioning, while successful, was atypical of classical conditioning; the agent would, in fact, respond more stongly to the CS by itself, as compared to the CS and the US together.

A single explanation underlies the results from the simultaneous and trace conditions. In both cases, the strength of the learned association is based on the amount of time the agent has been exposed to the relationship, not the number of times the relationship has been shown to it. For instance, an association can be fully formed by presenting the US and the CS together once for a long period of time, rather than presenting them for short periods of time over and over again. How long this period of time has to be can be tuned via the learning rate. Hence under trace conditioning, the agent is exposed not at all to the relationship, so fails to learn, and under simultaneous conditioning associative learning takes place according to the learning rate.

We can also explain the delayed conditioning result purely by appealing to the mathematics of the learning algorithm. In the limit, the strength of the connection between the CS and US settles to a value of 1 if the US is not present, and only 0.618 if both are present. Note that this analysis is only valid for situations where the CS and US are represented by single sensory nodes.

Extinction experiments involve determining how long a response continues after the CS is no longer followed by the US. The results on this test show that DAC2 is completely unable to lose an association. This is in stark contrast with natural learning.

[4]In simultaneous conditioning, the CS and US are presented concurrently. In delayed conditioning the CS precedes the onset of the US, but their activation overlaps. In trace conditioning there is a delay between activation of the CS which ceases prior to activation of the US.

To study generalization, DAC2 was first trained to recognize the concept of a corridor of one particular width, and was then exposed to corridors of varying widths. In this case, there was none of the expected generalization drop-off. That is, the system trained on corridors of width 100 units responded to corridors of width 70 identically to a system initially trained on a width of 70 units.

In natural generalization, the creature would respond most strongly to the particular situation it was trained on, and less strongly to situations similar to that one. This was not the case with this test on DAC2.

Extending the generalization experiments were the discrimination experiments. Here, the system was alternately exposed to corridors of differing widths. The concept sensor was activated for the first width, but not for the second. In classical conditioning, it is expected that the creature would initially respond to both widths, and then over time learn to respond to just one. This effect was not seen in DAC2.

Owing to the system's lack of extinction abilities, these results are not unexpected. In order to specialize, it is necessary for the system to lose the previous related association.

Conclusion

The final result is that while Distributed Adaptive Control may show the surface capabilities of classical conditioning (the ability to have a conditioned stimulus act as a predictor for an unconditioned stimulus), it does not have the deeper abilities of classical conditioning. In particular, the types of concepts it can extract are limited by having to associate large values of one sensor type with large values in another sensor. Also, it does not exhibit the property of extinction, meaning that it cannot lose associations which are no longer valid.

Without extinction the agent cannot adapt to a changing environment, nor can it adjust and fine-tune its learned associations to become more appropriate. Adding this to DAC2 is non-trivial: it is not sufficient to merely implement a decay parameter that causes associations to disappear slowly over time. A true extinction system could perhaps be based on looking at when the predictions of the system are wrong, not purely based on how often they are used. Such a model may even use a learned inhibition to handle the extinction, and then a decay in that inhibition could explain spontaneous recovery[5]. How such a system would interact with the concept-formation characteristics embedded in this view of classical conditioning is unknown.

The second major issue is the limitations in the sorts of 'concepts' that can be formed. We have seen how DAC2 is only capable of forming associations between fairly simple patterns, and that it has severe difficulties in forming associations based on negative correlations. To combat this problem, the reader is directed to (Stewart, 2000a), which presents many approaches which seem suitable for finding more complex patterns. In particular, the sparse representation networks may be a richer source for forming concepts.

Also, there has been an underlying theme in this paper about combining concept formation with conditioning. We have seen that these two aspects of learning may be tightly related, and are, at the very least, dependent on one another. What use is conditioning without found regularities in the environment on which to do the conditioning? And what use is concept formation if you are not doing anything with those concepts? It seems fruitful to investigate these two issues together.

[5]Having extinguished an association, and following a period of absence for either the CS or the US, natural organisms display a sudden recovery in the association between CS and response (which in the absence of the US quickly falls off again).

References

Anderson, J. 1983 *The Architecture of Cognition*. Cambridge, Mass: Harvard University Press.

Atkinson, R.L., and Hilgard, E.R. eds. 1996. *Hilgard's Introduction to Psychology*. 12th ed. San Diego: Harcourt Brace.

Edelman, G.E. 1987. *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.

Hinton, G.E., and Ghahramani, Z. 1997. Generative Models for Discovering Sparse Distributed Representations. *Philosophical Transactions of the Royal Society*, B, 352: 17-1190.

Stewart, T.C. (2000a) Neural Models of Concept Formation and Conditioning: A Literature Review. Technical Report, CSRP-524, School of Cognitive and Computing Sciences, Univ of Sussex.

Stewart, T.C. (2000b). Learning in Artificial Life: Conditioning, Concept Formation, and Sensorimotor Loops. MPhil. diss., School of Cognitive and Computing Sciences, Univ of Sussex. Forthcoming.

Thelen, E. and Smith, L. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

van Gelder, T. 1998. The Dynamical Hypothesis in Cognitive Science. *Behavioural and Brain Sciences* 21:1-14.

Veogtlin, T. and Verschure, P.F.M.J. 1999. What can robots tell us about brains? A synthetic approach towards the study of learning and problem solving. *Reviews in the Neurosciences* 10 (3-4): 291-310.

Verschure, P.F.M.J. 1998. Distributed Adaptive Control: Explorations in robotics and the biology of learning. *Informatik-Informatique* 1, 25-29.

Verschure, P. and Coolen, A. 1991. Adaptive Fields: Distributed representations of classically conditioned associations. *Network* 2: 189-206.

Verschure, P. and Voegtlin, T. 1999. A bottom up approach towards the acquisition and expression of sequential representations applied to a behaving real-world device: Distributed Adaptive Control III. *Neural Networks* 11:1531-1549.