Modeling Cognitive Development in the Human Brain

L. Andrew Coward
Department of Computer Science
Australian National University, ACT 0200, Australia
landrewcoward@shaw.ca

Abstract

Any architecture for modeling cognitive development must have several general characteristics. It must be possible to learn complex combinations of interacting cognitive capabilities using information derived from the same experience stream. Learning must be bootstrapped from experience with minimal a priori guidance and limited external guidance during learning, but in such a way that later learning does not interfere with earlier learning. Learning must be possible from single experiences. The architecture must provide an account for the observed dissociations between the various types of memory including semantic, episodic and procedural memory. A connectionist architecture with these characteristics is described.

1. Introduction

Much recent computational modeling of human cognition concentrates on accurate modeling of the phenomenology without regard to plausible matching between the detailed information processes employed and neural capabilities. An example is ACT [1], which provides quantitative modeling of experimental psychological data in terms of cognitive categories and processes which are defined by psychological observation but implemented by computer without regard to physiology. Another approach which focuses on practical applications is expert systems [e.g. 12] which capture the skills of human experts in a computer. These approaches make no effort to realistically model development of cognitive skills. Connectionist modeling is the only approach which makes claims to plausible modeling both of cognitive learning and of information processes at the neuron level [2]. However, the types of high level learning which have been successfully modeled tend to use restricted input and behavioural domains. For example, Roy et al [17] model the learning of words from visual and auditory inputs, but comment that "…the model is limited in its ability to deal with complex scenes…". Although Roy et al argue that their model has the potential for modeling learning of real cognitive processes, an issue often encountered with connectionist learning of more complex domains is the catastrophic forgetting problem [13] in which later learning sometimes overwrites and obliterates prior learning.

This paper argues that to be a plausible approach to modeling cognitive development, a model must demonstrate the potential to achieve a number of general characteristics exhibited by human development. A connectionist architecture which appears capable of these characteristics is then described. This architecture employs neural device algorithms with qualitative differences from conventional connectionist algorithms. The ability of the architecture to meet the required general characteristics and the strong dependence of this ability on the different device algorithms is described. In particular, the general ability to bootstrap memory and behaviour from experience and to use the same information recorded during experience to support episodic, semantic and procedural memory are described.

2. Criteria for Effective Modeling of Cognitive Development

There are a number of sometimes overlooked characteristics of human cognition which must be effectively addressed by any cognitive architecture which aims to model human development. Firstly, human beings learn a complex combination of different types of behaviour making use of the same experiences. For example, experimental psychology distinguishes between episodic, semantic, procedural and working memory and priming [19]. However, perceptual processing on the same stream of experience must generate information to support all these memory types, and information initially available to one memory type must over time become available in suitable form to others, while still remaining available to the original type. Thus episodic memories can result in semantic and procedural memories while still being accessible to episodic memory. Secondly, humans can learn new behaviour types with minimal interference with existing behaviours. This capability poses problems for conventional connectionist models, which tend to exhibit the catastrophic forgetting problem [13]. Thirdly, humans can bootstrap their cognitive capabilities from experience with minimal a priori guidance. For example, genetic guidance would not be able to specify categories of visual objects, but could perhaps provide preliminary and general associations between types of sensory input and types of behaviour which would need to be corrected and made much more specific by experience. The feedback available following behaviour can be reward or punishment, but not supervision in the sense of explicit indication of targets in terms of internal brain information structures as is required for connectionist supervised learning algorithms. However, a genetically defined tendency to imitate can make general reward and punishment feedback more efficient. Fourthly, humans are capable of significant, permanent learning from single experiences. For example, given a few seconds to examine each photograph in a set of 2500, subjects can later pick the familiar photograph from pairs in which only one came from the examined set at an accuracy level of 90% [20].

3. The Recommendation Architecture

Any system which must learn to perform a complex combination of interacting features with limited information handling resources in such a way that new learning does not interfere with prior learning tends to be constrained within a set of architectural bounds called the recommendation architecture [8]. These bounds define how the operations of the system are separated into modules, the ways in which modules interact, and the type of learning algorithms available to modules and devices. For a detailed description of the design of an electronic system implemented within these bounds, see [9].

In the recommendation architecture there is a primary architectural separation between a modular hierarchy called clustering and a subsystem called competition. Clustering defines a population of conditions within the available sensory information space and detects the occurrence of any defined condition. A subset of the conditions detected at any point in time is communicated to competition. Competition interprets each such condition as a set of recommendations in favour of a range of different behaviours, each with a different weight. Competition adds the weights of each recommended behaviour across all currently detected conditions, and implements the most strongly recommended behaviours. Consequence feedback following a behaviour can change the recommendation weights of recently active conditions into recent behaviours, but cannot change the definition of the conditions.

Devices in clustering learn and respond in radically different ways from conventional connectionist device algorithms. A clustering device permanently records a set of similar conditions. To be recorded, a condition must actually occur within the information available to the device, be similar in an information sense to conditions already recorded on the device, and at the time it occurs the device must also be receiving signals encouraging it to record conditions. These signals come from devices in other modules within clustering. The device is activated by the recording of a condition or by any subsequent repetition of the condition. On activation a device produces an output which is a series of activity spikes. The average rate of spike production indicates the number of its programmed conditions which are currently present, and a frequency modulation of the spike rate (i.e. bunching of spikes close to peaks in a regular modulation frequency) indicates the input population within which the condition was detected. In general, conditions will be detected by a device within a group of inputs much more strongly if a frequency modulation is present at the same phase on all the inputs, because otherwise fewer activity spikes will arrive within the time interval over which the device integrates its inputs. As an example of this frequency modulation in practice, if modulation is imposed on a subset of visual inputs corresponding with an area within a closed boundary (i.e. a visual object), only conditions within the object will be detected and recommended behaviours will be in response to this object. Frequency modulation also makes it possible to detect separate populations of conditions within two different objects in the same physical set of devices, if different phases of frequency modulation are imposed on inputs from the two objects. For a more detailed discussion of the frequency modulation mechanism, see [10].

A device in clustering thus records a set of similar conditions and indicates any repetition of a previously recorded condition. This device algorithm is in strong contrast with conventional connectionist device algorithms, in which devices have inputs with different weights which can be constantly adjusted, with no guarantee of response to an exact repetition of a condition that previously generated a response.

Devices in clustering are arranged in layers in which the condition defining inputs to one layer come from just one preceding layer. The first layer receives raw sensory inputs. This arrangement ensures that all the conditions detected within one layer are within the same range of complexity, where the complexity of a condition is the number of raw sensory inputs (including duplicates) that contribute to the condition either directly or via intermediate conditions. The layering also means that all the conditions detected at one time within a layer tend to be present within a system input state at one time (such as one visual object). Conditions within one range of complexity may be more appropriate for a particular behavioural function than conditions in other ranges.

The clustering device algorithm means that tight management is required over when and where additional conditions will be recorded. This management of change is a major role of the modular hierarchy in clustering. The first level of module above the device is a small area on one device layer. The next level is a column made up of a sequence of such areas across several layers. The next level is an array of such columns and the next level is a sequential block of such arrays. Each module detects a set of conditions made up of the sum of the sets detected by each device in the module. However, most of these conditions are only communicated within the module and used for change management within the module, only a small subset are communicated to other modules. A column module manages when conditions will be recorded within the column and within other columns in the same array. An array module ensures that some conditions are detected in every input state from a specific input domain. A block module ensures that the conditions detected within its constituent arrays are consistent with each other as indicated by a tendency for conditions in different arrays to have been active and recorded at the same time in the past. A block module then generates outputs in behaviourally useful ranges of complexity to competition.

Only some arrays target outputs on to competition, but any column within such an array can target competition. However, only devices within a column which detect conditions within a specific range of complexity (i.e. are located within a specific deviec layer) can target competition. These devices have sets of conditions which they detect, and the sum of these sets for a column is called the portfolio of the column. Portfolios are important in understanding the processes which lead to cognition.

4. Definition of Conditions and Portfolios

A vast range of raw inputs containing information about the external environment and the internal state of both the brain and the body are available to the brain from the senses. A somewhat oversimplified way of understanding the definition of information conditions is that one condition corresponds with a specific set of these inputs each being present to an individually specified degree. Because conditions cannot be specified a priori, there is a random element to the definition of conditions, and because conditions are not changed after being recorded, any one condition or portfolio is perceptually, cognitively and behaviourally ambiguous. Unambiguous meanings are only achieved in competition across populations of conditions. However, conditions with complexities of the same order of magnitude as visual object perceptions will tend to be less ambiguous with respect to categorization of visual objects than conditions on other levels of complexity, athough no conditions on any level correlate unambiguously with such categories.

This simple view of conditions is made considerably more complex because a column portfolio can be activated not only by the presence of its conditions within sensory inputs, but also indirectly by two other types of mechanism. One mechanism is that it can be activated if a number of other columns are already active which have often been active in the past at the same time as the column. The other mechanism is that columns can be activated if a number of other columns are already active which have recorded conditions at the same time in the past as the column. These indirect activations are behaviours which must be recommended by the already active columns into competition and accepted. When there is simultaneous activity or condition recording in two columns, there is a strong recommendation weight created in competition in favour of the activity of one column activating the other, but this recommendation weight declines fairly rapidly with time. However, if an indirect activation actually occurs in the course of generating a behaviour which is followed by positive consequences, the decline is reduced, and frequent such occurences stabilize or increase the weight.

These indirect activation mechanisms can be viewed as supplementing the conditions present in current sensory inputs with other conditions which have a significant probability of being relevant to determining the most appropriate current behaviour. For example, conditions which have been active in the past at the same time as currently present conditions may contain information about the current environment which cannot currently be observed [7, 11].

A newly recorded condition is made up of a set of currently active component conditions. Some of these component conditions may be combinations of currently present sensory inputs, and some could have been activated by one of the two indirect mechanisms. Both the definition of conditions in terms of sensory inputs and the relationship between sensory inputs and the resultant pattern of condition activation can therefore become very complex.

Learning occurs by permanent addition of conditions to modules on many levels including device, column, array and block. Conditions are defined heuristically, and there is no a priori knowledge of which higher level modules such as arrays will require many column portfolios, or which column portfolios will need many device level portfolios and which devices will require inputs from which other devices etc. Hence assignment of column and device resources must be performed heuristically on the basis of need. A resource management function must therefore assign provisional conditions to devices, devices to columns and columns to arrays on the basis of current need.

Resource management requires two components. One is a map of resources identifying which are unassigned, the other is a process for identifying appropriate connectivity [6]. Resource management is then a periodic process during which requirements for new resources are identified, resources are assigned, and appropriate provisional connectivity provided. Connectivity to support indirect portfolio activation on the basis of simultaneous condition recording could be efficiently provided via the resource map, at least initially. Connectivity to support indirect portfolio activation on the basis of prior or subsequent condition recording would for efficiency reasons tend to continue to be dependent on the map.

5. Behavioural Interpretation of Portfolios

Competition is made up of devices which total the excitatory and inhibitory weights of currently active inputs from a range of sources, and produce an output if the total exceeds a threshold. The devices adjust their input weights in response to consequence feedback. Unlike the device algorithms used in clustering, these algorithms are generally similar to the perceptron type algorithms used in conventional connectionist networks.

The competition system is made up of components corresponding on a one-for-one basis with all possible system behaviours. Each component is a device or group of devices. Some components correspond with behaviours which are “atomic” in the sense that the system can implement the behaviour or not implement it, can vary the speed and perhaps the degree with which the behaviour is performed, but cannot change the nature of the behaviour. An atomic behaviour could be the contraction of an individual muscle, or genetically programmed groups and sequences of such contractions. Other components correspond with higher level behaviours such as groups and sequences of atomic behaviours, and yet higher behaviours which are groups and sequences of such groups. At the highest cognitive levels, behaviour is achieved by outputs from clustering driving a sequence of competition components which in turn activate more specific competition components. Any very frequently occurring sequence or set of behaviours will tend to result in a new component in competition which receives most of the inputs from clustering and drives behaviour more directly into atomic behaviours.

The outputs of a component are the outputs of specific devices within the component. Because of the use of consequence feedback within competition, such outputs cannot have operationally complex meanings. Only two types of operational meaning are possible. One is a recommendation to perform the behaviour corresponding with the component. If such an output exits competition, it becomes a command to perform the corresponding behaviour. Otherwise it is directed at a range of components corresponding with more detailed or specific behaviours within the recommended type, and increases the probability of such behaviours being accepted. These more detailed components could also receive inputs directly from clustering.

The other type of operational meaning is a recommendation against performing any behaviour other than the component behaviour. Such outputs are directed at competition components corresponding with different behaviours. A high proportion of these outputs are directed at peer components, in other words components corresponding with different behaviours on roughly the same level of detail.

When a condition is recorded in clustering, it can immediately acquire a range of different behavioural meanings either directly through recommendation weights in competition or indirectly by incorporation in other conditions with such recommendation weights. Any subsequent change to the condition would therefore result in a wide range of uncontrolled behavioral side effects. The restrictions that conditions cannot change, and devices can only add similar conditions, limits these side effect much more effectively than perceptron type algorithms [9].

6. An electronic system with the recommendation architecture

An electronic system with the recommendation architecture has been implemented, and demonstrated the capability to define portfolios from experience with no a priori guidance, and to associate different combinations of portfolios with different behaviours using only reward and punishment feedback. The ability to learn with minimal interference with prior learning has also been demonstrated [8, 9]. Processes within the electronic implementation which strongly resemble cognitive processes including category learning, learning to activate appropriate visual information in response to words, and activation of mental images have been observed [9].

7. The Recommendation Architecture cognitive model

In the recommendation architecture, information can be accessed by four qualitatively different mechanisms. Firstly, the actual presence of a condition within current sensory inputs activates the substrate on which the portfolio containing the condition is recorded. Secondly, an activated portfolio can recommend activation of other portfolios which have often been active at the same time in the past, and the recommended portfolio will activate if adequate recommendation strength is present. A variant of this mechanism is activation of portfolios which were often active somewhat before or somewhat after activity in the active portfolio.

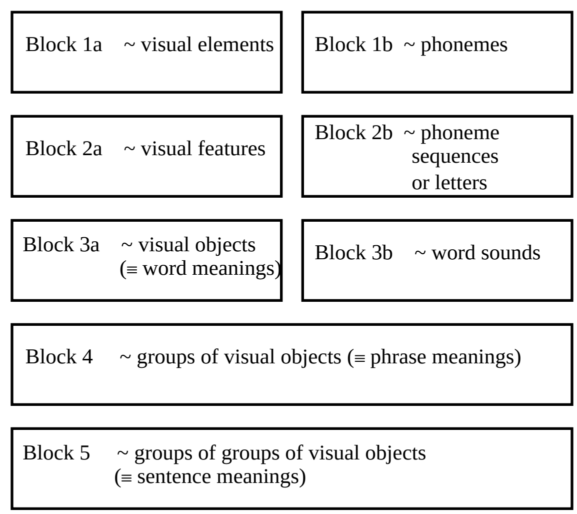

Figure 1. Architecture to support cognitive processes. Block modules detect conditions on five different levels of complexity, with condition defining information passing sequentially from top to bottom. The outputs of a block indicate the detection of conditions within the same range of complexity as the indicated cognitive category (features, objects, groups of objects etc.) but conditions do not correlate unambiguously with such categories. The subdivision of levels 1 through 3 reflects different input domains within which conditions are detected.

Thirdly, an activated portfolio can recommend activation of other portfolios which recorded conditions at the same time in the past. A variant of this mechanism is activation of portfolios which recorded conditions somewhat before or somewhat after an episode of condition recording in the active portfolio. Indirectly activated portfolios can in turn recommend activation of yet other portfolios. The fourth mechanism is comparison of recommendation weights. The weights of all active behaviours into each recommended behaviour are totaled, and the behaviours with the strongest weights are implemented.

The simplest arrangement of clustering blocks able to support complex cognitive behaviour is illustrated in figure 1, and examples of competition subsystems associated with one block are illustrated in figure 2. In figure 1, outputs from block 2a to competition indicate the detection of portfolios with a complexity comparable with visual features. Outputs from block 3a indicate portfolios comparable with visual objects, outputs from block 4 indicate portfolios comparable with groups of objects, and outputs from block 5 indicate portfolios comparable with groups of groups of objects. Block 4 will therefore detect portfolios in a sequence of perceptual objects, and block 5 will detects portfolios incorporating information derived from several area 4 outputs, in other words portfolios containing information derived from all members of the sequence.

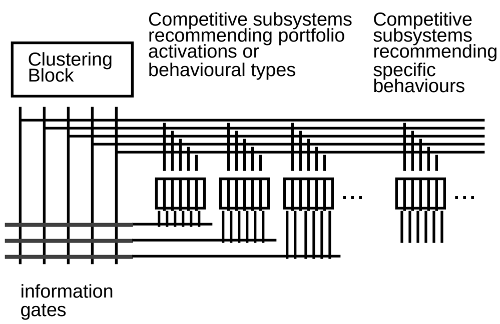

Figure 2 Competitive components receiving outputs from one sequence module. Different behavioural interpretations are placed upon the same clustering outputs by different components. There is competitive inhibition between and within competitive components to limit selected behaviours to a small, consistent set. In some cases the behaviour accepted by a competitive subsystem is release of the outputs from clustering which correspond with the behaviour to either the next clustering level or to a more detailed competition subsystem. This release behaviour is indicated by the information gates. In other cases competition outputs drive their corresponding individual behaviours, either external (e.g. eye movements) or internal (e.g. prolonging the activity of clustering neurons in specific modules).

Ten types of competition component corresponding with ten types of behaviour which could be recommended by a clustering area are as follows: prolong the activity of some currently active portfolios (for example of a group recommending a sequence of behaviours until the sequence is complete); activate portfolios active at the same time in the past as the currently active portfolio; activate portfolios containing conditions recorded at the same time in the past as some conditions in the currently active portfolio; activate portfolios containing conditions recorded just before or just after some conditions in the currently active portfolio; synchronize the activity (i.e. phase of modulation frequency) of several different groups of currently active portfolios; perform a general sequence of attention behaviours; perform a specific sequences of attention behaviours; perform an individual attention behaviour; speak a word; and say a phrase.

Competition receives outputs from portfolios currently being detected by clustering. If a portfolio has been present in the past at the same time as the performance of a number of different behaviours, it will have acquired recommendation weights in favour of or against those behaviours in the component corresponding with the behavior, depending on the consequence feedback from those behaviours. Competition adds the weights of all currently recommended behaviours and selects the behaviour with the largest weight. New portfolios are given an initial weight similar to the weights of the most similar previously existing portfolios, or genetically defined initial weights. If portfolios are new and no significant recommendation strength has been assigned in these ways, a behaviour can be selected randomly. Such random selection can be limited to behaviours within a behaviour type which has already been selected. Alternatively, a behaviour can be selected by imitation of an externally observed behaviour.

8. Bootstrapping of memory and behaviour

Behaviours can be defined heuristically with limited a priori guidance. Such definition will be illustrated by describing a possible process for acquisition of simple speech, with careful attention to the nature of the a priori (genetic) guidance needed. Learning goes through a series of steps generally consistent with observation of how humans learn to speak and understand words, but the purpose of this section is not to offer a formal model for speech acquisition but rather to demonstrate that speech can be acquired in the recommendation architecture model with only limited and plausible genetic guidance.

Genetic information specifies creation of a set of detailed competition components which drive muscle movements contributing to sound generation. Every possible such movement has a corresponding genetically specified competition component, and activation of the component results in the movement. Genetic information also specifies the existence of intermediate competition components which activate randomly selected sequences of detailed components and therefore generate sounds. Learning proceeds in a series of partially overlapping steps.

The first step is creation of an array of portfolios in clustering blocks 1b and 2b of figure 1 in response to hearing sounds, at different levels of condition complexity. Because speech is somewhat different from other sounds, there will be a tendency for speech related portfolios to be somewhat separate from the portfolios created in response to other sounds. This tendency could be reinforced by genetically determined connectivity biases within clustering.

The next step is generation of sounds using intermediate competition components. Initially a component is randomly selected after any activation of a portfolio population in the array of sound related portfolios in levels 1 and 2. A positive consequence feedback is genetically programmed to be generated if the portfolios activated in response to hearing an external sound (i.e. not self generated) are similar to the portfolios activated shortly afterwards in response to hearing a self generated sound. One effect of this consequence feedback is that the sequence of detailed components activated by the intermediate component is fixed long term. In other words, the presence of a sound in the environment results in production of that sound becoming instantiated in an intermediate component. If there is no feedback within some period of time, the intermediate component is reconfigured with a different randomly selected sequence of sound generating muscle movements, or deleted. The second effect of the consequence feedback is that the activated portfolios acquire recommendation strength in favour of activating the intermediate component. In other words, the behaviour of imitating sounds which are heard is acquired.

The next step in learning is that portfolios are created in area 3b of figure 1 in response to sequences of sequences of sounds which are heard. These portfolios will correlate partially (or ambiguously) with frequently heard sequences, and therefore with words which are heard. Higher level competition components are defined which activate randomly selected sequences of intermediate components. If a self generated sound sequence activates a portfolio population in area 3b similar to an immediately prior population activated by an external sound sequence, a genetically defined consequence feedback results in the higher level component being fixed and the active portfolios acquiring recommendation strength in favour of activating it. Thus the behaviour of imitating words which are heard is acquired.

The next step utilizes a genetically programmed tendency for the portfolios created in level 3b in response to sequences of sounds to have recommendation strength in favour of activation of portfolios created in level 3a in response to visual experiences, if the visual experience portfolios are often active at the same time as the sound sequence portfolios. The effect is that hearing the word will tend to activate a partial visual image of the type of object often seen when the word was present in the past. The portfolios making up a visual image will also recommend any other behaviours which have become associated with the object. In addition, the visual experience portfolios acquire recommendation strength in favour of activating the higher level component which tends to be activated by the sound sequence portfolios. The effect is that seeing the object will tend to result in speaking the corresponding word. Consequence feedback associated with the perceived behavior of adults in response to activating a higher level component (i.e. speaking a word) will affect the recommendation strengths of active portfolios in favour of the word just spoken.

Thus a set of genetically defined tendencies result in relatively efficient acquisition of simple speech behaviours. Learning does not require a priori internal definition of cognitive categories. Genetic information provided three types of information. Firstly, it indicates the available range of detailed muscle movements. Secondly, it biases initial connectivity in favour of the types of sensory inputs and portfolio condition complexity ranges which will most effectively drive those behaviours. Thirdly, it defines in general terms the circumstances in which consequence feedback will be generated and the effects of such feedback.

9. Different Types of Memory

It has been argued that there are a number of different memory systems in the brain, based on the observed dissociations between different memory phenomena [19]. These sytems include semantic, episodic, procedural and working memory. The following discussion will focus on semantic, episodic, and procedural memory, working memory is discussed in detail in [10].

There are two mechanisms by which information can be recorded in the recommendation architecture. One is permanent recording of conditions in clustering, the other is adjustment to recommendation weights in competition. The permanent recording of conditions means that the system has the capability to learn from single experiences. In the model there will be a level of condition recording in response to every experience, with higher levels for experiences with higher levels of novelty [6]. This higher level of recording in response to novelty accounts for the high human capability to detect the novelty or otherwise of an experience. For example, subjects exposed briefly to a set of several thousand photographs could a few days later distinguish between photographs in or not in the set with 90% accuracy[20].

As discussed earlier, there are four mechanisms by which information can be accessed in the recommendation architecture. The use of different combinations of these mechanisms can account for the phenomena and dissociations between semantic, episodic and procedural memory.

10. Episodic Memory

Episodic memory is memory of the past with a context of what else happened at the time, in contrast with semantic memory in which memories of facts are detached from memories of where those facts were learned [21]. Various types of recall experiments measure episodic memory.

In targeted recall subjects are asked to recall particular past events [14]. In cued recall, subjects are given a cue which may be a word [16] or a type such as "vivid memories" [18]. In involuntary recall, some specific environmental stimulus such as a smell or a taste brings a memory to mind unsought [3].

The starting point for targeted recall in the recommendation architecture model is hearing words which describe an event. Portfolios are activated which contain conditions within the sounds of the words. Secondary populations of portfolios containing conditions which occurred within visual and other sensory inputs are activated on the basis that the secondary portfolios have often been active in the past at the same time as the primary "auditory" portfolios. A significant proportion of the portfolios in these secondary populations were also active during the event, and a somewhat smaller proportion recorded conditions during that event. Because of the words used, the proportions are larger for the target event than for any other event.

All active portfolios have recommendation strengths in favour of activating other portfolios which recorded conditions at similar times in the past. Active populations at higher levels derived from the presence of words like "recall" have recommendation strength in favour of accepting these types of recommendations. Because the target event has the highest proportion of activated portfolios in the secondary population, acceptance of such recommendations will tend to result in an active tertiary population with an even higher proportion of portfolios which recorded conditions during the target event. This process is self reinforcing, especially if a large number of conditions were recorded during the target event. The resultant population will be experienced as a general re-perception of the original event, although in general the portfolios closest to input from the senses are not reactivated. The activated portfolios in this population have recommendation strength in favour of, for example, generating verbal descriptions of the event. Use of recommendation strength in favour of activating portfolios which recorded conditions somewhat before or somewhat after condition recording in currently active portfolios allows the re-perception to be set at the beginning of the event and to be moved through the event.

Recommendation strengths will always also be present in favour of activating portfolios on the basis of simultaneous past activity and simultaneous past recording during other events. The activated population is therefore unlikely to be an exact match for the original, although in general the higher the level of condition recording during the event, the greater the probability of a close match.

Cued recall operates in a very similar fashion. However, the initial secondary population may contain portfolios which recorded conditions during a number of past events. In the absence of specific indication of one event in the verbal cue, the tertiary population will evolve towards the event which happened to be represented by the highest degree of condition recording in the initial population. Events which resulted in a high degree of condition recording across many portfolios will tend to be the end points of this process.

Involuntary recall is the result of strong condition recording in portfolios activated in response to a sensory stimulus (for example, a novel smell or taste) at the same time as strong condition recording in other portfolios in response to some event. A later repetition of the sensory stimulus activates the portfolios which originally responded to that stimulus. These portfolios in turn activate portfolios which recorded conditions at the same time in the past, resulting in a re-perception of the event.

11. Semantic Memory

A typical way of measuring semantic memory in the laboratory is sentence and category verification. Sentence verification experiments measure the time for subjects to respond with the correctness of sentences like “Is a robin a bird” or “Is a penguin a bird”. Category verification experiments are essentially equivalent and measure such times for simple category-exemplar pairs like bird-robin or bird-tree. It is found that for different members of the same category paired with the correct category, responses are faster for more typical category exemplars. For example, the response to bird-robin is faster than to bird-chicken. However, responses to clearly incorrect category-exemplar pairs like “Is canary an animal” is also fast [15].

The recommendation architecture model for category verification experiments can be understood by considering the portfolio populations activated in response to the words indicating category and exemplar. The portfolios activated in response to hearing the name of the category are portfolios which have often been active in the past when the category name was also present. Hence they will be portfolios also present when exemplars of the category were present, since this is the way the category is learned. Hence the active population is the set of portfolios which have most often been present when different category exemplars have been present. There will therefore be an overlap between the population activated in response to the name of the category and that activated in response to the name of the exemplar. This overlap will be greater for more typical exemplars, and very small if the exemplar is not a member of the category. The degree of overlap is itself a condition which can be detected and recommends for or against identifying the exemplar as a category member. If the exemplar is typical, overlap is substantial and its detection is rapid. Atypical exemplars have more moderate overlap, and more time may be required to expand the portfolio populations to include portfolios active at the same time in the past but slightly less often to achieve an overlap adequate to generate the appropriate verbal response. However, objects which are not in any way members of the category will have negligible overlap which again is detected rapidly. The model is therefore in agreement with the observations of [15].

In contrast with the spreading activation model of Collins & Loftus [4], there are no units which correspond with concepts like categories or the features of categories. Portfolios are groups of conditions in which there has been a degree of randomness in the definition of individual conditions, but conditions within one portfolio have some similarity with each other and have tended to occur at similar times in the past. A portfolio may therefore have a probabilistic correlation with many different features and categories, with the probabilities expressed, for example, as recommendation weights into naming the features or categories.

12. Procedural Memory

Procedural memory is defined as the ability to acquire skills. Observations of amnesics indicate that such memory is at least partially dissociated from semantic memory, since amnesics can acquire such skills at apparently normal rates. Thus amnesics can acquire motor skills such as mirror tracing tasks [5].

In the recommendation architecture model for learning a skill, portfolios activated within clustering in environments where the skill is relevant must acquire weights in competition associated with skilled behaviours. These portfolios will generally include both new information elements resulting from novelty in the environments and information elements recorded in prior experiences. Although the new elements may be particularly useful for recommending the new behaviours, some skill learning would be possible using only previously recorded elements which happen to occur in the new environments. Thus skill acquisition could proceed in the absence of condition recording.

13. Conclusions

The recommendation architecture cognitive model demonstrates the general capabilities to learn complex combinations of capabilities utilizing information drawn from a common experience stream, to bootstrap learning from experience with minimal and genetically plausible a priori guidance and limited and plausible external guidance during experience. The catastrophic interference between new and prior learning found in other connectionist architectures can be avoided. Significant, permanent learning is possible in response to single experiences. The information recording and access mechanisms of the recommendation architecture can provide an account for the phenomena of and disociations between semantic, episodic and procedural memory. The recommendation architecture thus has considerable advantages as a starting point for modeling cognitive development.

References

[1] Anderson, J.R. and Lebiere, C. (1998). *The Atomic Components of Thought*. NJ: Erlbaum.Coward, L. A. (1990). *Pattern Thinking*. New York: Praeger.
[2] Ballard, D. H. (1997). Introduction to Natural Computing. MIT Press.
[3] Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology* 10, 435 - 454.
[4] Collins, A. & Loftus, E. (1975). A spreading-activation theory of semantic processing. *Psychological Review* 82(6), 407 – 428.
[5] Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia* 6, 225 - 264.
[6] Coward, L. A. (1990). *Pattern Thinking*. New York: Praeger.
[7] Coward, L. A. (1999). A physiologically based theory of consciousness, in S. Jordan (Ed.). *Modeling Consciousness Across the Discipline*. Maryland: UPA.
[8] Coward, L.A. (2000). A Functional Architecture Approach to Neural Systems. *International Journal of Systems Research and Information Systems*, 9, 69-120.
[9] Coward, L. A. (2001). The recommendation architecture: lessons from the design of large scale electronic systems for cognitive science. *Journal of Cognitive Systems Research*, 2, 2, 111-156.
[10] Coward, L. A. (2004). The Recommendation Architecture Model for Human Cognition.Proceedins of the 2004 Conference on Brain Inspired Cognitive Systems, Stirling, Scotland.
[11] Coward, L.A. and Sun, R. (2004). Some Criteria for an Effective Scientific Theory of Consciousness and Examples of Preliminary Attempts at Such a Theory. Consciousness and Cognition in press.
[12] Firebaugh, Morris W. (1993). Artificial Intelligence, A Knowledge-Based Approach. PWS-Kent Publishing Company.
[13] French, R. M. (1999). Catastrophic Forgetting in Connectionist Networks. *Trends in Cognitive Science* 3(4), 128-135.
[14] Neisser, U., & Libby, L. K. (2000). Remembering life experiences. In E. Tulving and F. I. M. Craik (Eds.).*The Oxford Handbook of Memory*. Oxford: Oxford University Press.
[15] Rips, L., Shoben, J. & Smith, E. Semantic Distance and Verification of Semantic Relations. *Journal of Verbal Learning and Verbal Behaviour* 12, 1 – 20.
[16] Robinson, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology* 8, 578 - 595.
[17] Roy, D. K. and Pentland, A. P. (2002). Learning words from sights and sounds: a computational model. Cognitive Science 26, 113 - 146.
[18] Rubin, D. C. & Kozin, M. (1984). Vivid memories. *Cognition* 16, 81 - 95.
[19] Schacter, J.C., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In E. Tulving & F. I. M. Craik (Eds.). *The Oxford Handbook of Memory*. New York: Oxford University Press.
[20] Standing, L., Conexio, J., & Haber, R.N. (1970). Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychonomic Science* 19(2), 73 - 74.
[21] Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Reviews of Psychology* 53, 1 – 25.