

Cognitive Development in Context: Learning to Pay Attention

Petra Björne
petra.bjorne@lucs.lu.se

Christian Balkenius
christian.balkenius@lucs.lu.se

Lund University Cognitive Science
Kungshuset, Lundagård, SE-222 22 Sweden

Abstract

A developing system must be able to learn new things without forgetting what it has learned before. It should be capable of reacting in different way to the same stimuli in different contexts. Context sensitive reinforcement learning, which parallels some of the functions of the basal ganglia, is a learning algorithm that fulfills this requirement when the context is explicitly given. Here, we extend the algorithm with the ability to identify the relevant features of the environment that defines the context. It is suggested that this is a critical component of an architecture for cognitive development and we present simulation results that illustrate the operation of the system.

1. Introduction

A developing system must be able to learn new things without forgetting what it has learned before. It should also be capable of reacting in new ways in new contexts. Ideally, what has already been learned should be generalized to new situations, while new learning should not interfere in a negative way with previously learned behavior that is still appropriate.

For this to be possible, it is essential that new and old contexts can be distinguished in an efficient way. Otherwise, it would not be possible for the system to know when behaviors should be modified to fit the new situation or simply forgotten and relearned.

However, in most cases, there is no individual stimulus in the environment that signals that the situation is new or different. This can only be determined by examining several cues in combination and even when a single stimulus indicates a new context, this stimulus has to be attended to influence learning.

Here, we want to develop a computational system that can automatically learn to attend to relevant aspects of the environment and use these aspects to determine when they should be used as indications of a changed context. The

system takes its starting point in behavioral data on the role of context in learning and relearning. In addition, the architecture of the system is constrained by the anatomy and physiology of specific brain regions.

It is yet not possible to build a large scale model of the brain at a detailed physiological level. Too many details are simply unknown. Instead, we strive for a model that parallels the brain at a system level where the different components of the model functionally correspond to different brain regions, but we do not attempt to model how these regions work at a neuronal level. Below, we focus on the basal ganglia and its role in the production of context dependent action and in the selection of contextual cues.

A task such as reaching for an object will involve several serially connected chains of specialized motor structures. At the same time, information will be processed in a parallel organization of multiple cortical, basal ganglia, thalamic and cerebellar structures (Salinas et al., 2000). Thus, a model of the functional role of the basal ganglia in a motor task must be consistent with a model of the functional role of other specialized cortical areas which the basal ganglia interact such as the thalamus, cortex and the cerebellum. This is not to say that at every stage of modeling we need to have a complete model of the brain, but that it is necessary to be aware of the fact that no brain structure works in isolation from other structures. A model of any single structure must aim at integration with models of the other structures.

In the case of modeling the functions of the basal ganglia, besides accounting for their involvement in motor tasks, such a model should strive to be consistent with their involvement in non-motor tasks, such as sensory decision making, motivation, attentional and volitional modulation of other neural structures.

Further, a model of the basal ganglia would need to be consistent with data on impairments caused by degenerative illnesses afflicting the basal ganglia, such as Parkinson's disease and Huntington's chorea. A model of this kind would be able to provide valuable information on the interaction of cognitive and motor impairments of these patients.

There is always the possibility that we will find a mis-

match between a neurocognitive model and experimental data, human or animal. This, however, need not signal a drawback. Instead, a mismatch can guide further empirical research and help improve the model (Hyland, 2000). A system model of neurocognitive functioning, if based on neurophysiological and behavioral data, will provide a powerful instrument for analyzing experimental data and develop hypotheses for further research. Both the results from simulations and new experimental data will, in turn, improve the model.

In the following subsections we describe data and models from a different perspectives that have something to say about how learning and attention interacts with context. A more detailed review can be found in Balkenius (2000).

1.1. Learning in Context

In most models of learning, context does not play a part at all. In those models where context have a role, it is often in the form of a dedicated input. Although such a solution is a step forward, it neglects that in the real world stimuli and contexts labeled indicating their role in the learning experiment.

Another problem with many learning models where context appears is that it acts in the same way as any additional stimuli. While this acknowledges that a distinction between stimulus and context is not always easy to make, it ignores that in many learning experiments, the roles of stimulus and context are very different. While initial learning appears to be mainly context insensitive, relearning makes behavior context dependent. This has been shown in an important experiment by Bouton (1991) where learning in one context generalizes to other contexts, but extinguished behavior reappears outside the extinction context (see also Balkenius and Morén, 2000). This learning strategy is very powerful as it maximizes the generalization of learned behavior between contexts, while still being able to differentiate behavior in different contexts when needed. These results imply that there is a need to distinguish between stimulus and context.

1.2. Context and Attention

We have earlier proposed that a context code can be constructed from a sequence of attentional fixations (Balkenius, 2000). Balkenius and Morén (2000) describe a computational model that can automatically generate context codes from sequences of attentional fixations of features in the environment. The model binds each environmental feature to its location before all features are combined into a context code. The selection of features was controlled by a fixed mechanism that would scan the features of the environment in a sequential manner.

In many cases, it would be useful if the selection of the features that makes up the context could be put under reinforcement control. This would potentially allow the system to select the critical features that define the context or task to be accomplished. The initial steps towards such a mechanism were described by Balkenius (2000), where it was suggested that attentional shifts should be considered as any other action and learned in the same way. This principle was called *attention-as-action*.

An important consequence of this principle is that learned attention shifts will become context dependent in the same way as other actions. Since attention controls what stimuli are treated as parts of the context, this will make the contexts codes themselves context dependent.

1.3. The Basal Ganglia

Traditionally, the basal ganglia have been considered to be important for voluntary control and planning of body movements (Middleton and Strick 1994; Hikosaka, Takikawa and Kawagoe, 2000). However, through studies of persons with impairments of the basal ganglia, such as Parkinson's disease and Huntington's chorea, increasing insight into the cognitive functions of the basal ganglia has emerged. Along with the above mentioned neurodegenerative disorders, research into neurodevelopmental disorders, such as ADHD, autism and obsessive compulsive disorders (Bradshaw 2001) has further highlighted the importance of the basal ganglia in higher cognitive functions.

The basal ganglia operate by exerting tonic inhibition with phasic disinhibition (Kimura 1995; Hikosaka, Takikawa and Kawagoe, 2000), i.e., they select appropriate behaviors rather than controlling their detailed execution. This is probably true for both motor and non-motor functions controlled by the basal ganglia. An example would be the orienting response, which requires integration of information from several sensory modalities. From this integrated information an appropriate signal is selected, probably by processes in the basal ganglia (Redgrave, Prescott and Gurney 1999; Hikosaka, Takikawa and Kawagoe, 2000). The actual motor response is controlled by the superior colliculus (SC), which receives input from the frontal eye fields (FEF) and areas of the parietal cortex constituting the neural correlates for selection of saccades or attention. The role of the basal ganglia is to inhibit the SC (Hikosaka, Takikawa and Kawagoe, 2000). This is done through the substantia nigra (SN), which projects to the intermediate layer of the SC. The SN, in turn, is inhibited by the caudate nucleus. Occasionally the SN releases the inhibition of the SC, which results in a saccade to the contralateral side. Here, the basal ganglia select to produce the response, but the specific target of the orientation is controlled by the cortical input to the superior colliculus and

not by the basal ganglia.

The output neurons of the SN or GPi show very high spontaneous activities. In contrast, the projection neurons of the striatum become active only when the animal performs an appropriate task, and whereas the neurons of the putamen can be activated by simple motor tasks, complex behavioral tasks are needed to activate the caudate. This suggests that the neurons in this region are sensitive to the behavioral context in which an action should or should not be selected.

The responses of the neurons in the caudate resembles the responses of those in the SN but with opposite signs, changing their activity when the location of the stimuli must be remembered or attended, or when the saccade uses the working memory.

The function of the basal ganglia has been linked to behavioral learning that is sensitive to reward. The responses of the dopamine cells appear to code for the temporal difference error between the expected and actual reward received. Recently, the responses of dopaminergic cells in the basal ganglia have been shown to react in a context dependent way.

1.4. Working Memory

The basal ganglia are also involved in the manipulation of working memory. In patients with Parkinson's disease, the degeneration of dopaminergic neurons projecting to the basal ganglia leads to a difficulty in manipulating information that is stored in working memory (Lewis, Cools et al. 2003; Lewis, Dove et al. 2004). These patients seem to be able to maintain information over a short time span in a verbal memory task and then retrieve it in an unmodified version. However, they seem to have difficulties in manipulating the same information. According to the authors, this would correspond to the visuospatial tasks of executive functioning that are also considered to be particularly difficult for the PD patients, namely tasks that involve manipulation of spatial information. One way to investigate working memory is through continuous performance tests.

Continuous performance tests are frequently used in assessments of sustained attention (Lin 1999; Oades 2000). The CPT-AX is a continuous performance test that puts high demands on working memory. Frank et al. (2001) have developed an even more demanding version of the task (Fig. 1). In the original version, the subject is presented with a sequence of letters and is expected to respond to the letter X if the previous stimulus was an A. In the extended version, the subject has to respond to the X preceded by an A within a context defined by the number 1. If the number 2 instead defines the context, the subject has to respond to the letter Y if preceded by a B.

This calls for rapid updating of working memory, i.e., an

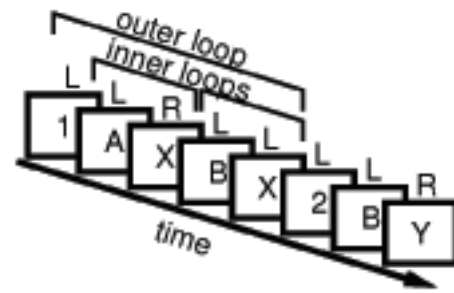


Figure 1. The 1-2-AX task.

incoming stimulus has to be encoded. Furthermore, the context 1 or 2 has to be maintained stably while interference occurs from processing of targets and distracters. Finally, the task calls for selective updating of working memory, where the context 1 or 2 remains stable, while the sequence of letters is continuously updated.

The thalamus is tonically inhibited by the GPi/SNr and phasically disinhibited by the firing of striatal neurons. This functions as a gating mechanism, enabling but not causing other functions to occur, though as mentioned earlier, the context of the action is not defined by the disinhibition by the striatum. Frontal neurons react momentarily to irrelevant stimuli, returning to the task-relevant stimuli and maintaining these after the irrelevant stimuli have disappeared. This intrinsic maintenance is important for working memory and robust maintenance of task-relevant stimuli.

Disinhibition by striatal firing will modulate the intracellular switch of the frontal neurons, leading to an update of current and maintained information. Thus, according to Frank et al. (2001) stimuli will activate corresponding frontal representations and they will be maintained if they trigger the intrinsic maintenance switch. Those stimuli that do not have this intracellular switch activated will decay quickly, but will be maintained by recurrent excitation until other stimuli are presented. This latter function is important for learning what will be relevant to maintain.

Striatal neurons fire for a specific conjunction of environmental stimuli and internal context representations through descending projections from the cortex. Thus, striatal neurons would fire in response to the encoding of a frontal representation of the task 1-2-AX together with the incoming of some stimulus (1 or 2) and the encoding of the sequence of letters, enabling the response to the letters x and y when appropriate.

2. Toward a Model

Taken together, the data presented above suggest that the basal ganglia is a central structure in the learning of con-

text sensitive behavior dependent on reward contingencies whether the actions be external, such as orienting movements, or internal, such as manipulations of working memory. We now turn to a computational model that attempts to cover the central ideas described above. These include the need for a context code that can be adapted to the task the ability to put attention under reinforcement control.

Below we describe

2.1. Context Sensitive Reinforcement Learning

Like most on-line learning algorithms, the standard reinforcement learning algorithms are sensitive to catastrophic forgetting (cf. French, 1999). If it first learns one task and then another, the second learning experience is likely to interfere with the first. This is especially the case when a look-up table is used to store the value of each stimulus-response association.

For a developing system, it is essential that new tasks can be learned without erasing older ones. Balkenius and Winberg (2004) developed a novel context sensitive reinforcement learning algorithm, ContextQ, that overcomes this problem by using an additional input that codes for the context in addition to the input that codes for the current stimulus or state. The algorithm is an extension of the popular Q-learning algorithm (Watkins and Dayan, 1992) and uses a function approximator to estimate the function,

$$Q(c, s, a),$$

which assigns a value to each action a in state s and context c . The algorithm starts out with a zero value for all actions and as long as the received reinforcement is larger than predicted by the Q-function, learning increases the weight of a linear mapping between the stimulus and response. This learning is not influenced by the context and will allow the system to automatically generalize all learning to new contexts. This part of the algorithm is identical to what Sutton and Barto (1998) called LAQ. The difference compared the LAQ algorithm occurs during extinction, i. e. when the received reinforcement is lower than expected. In this case, the linear associator is unaffected. Instead, a shunting inhibition from the context to the active stimulus-response association increases. This will make behavior learned during extinction context sensitive. Optionally, the two modes of learning can be mixed such that both acquisition and extinction involves both the stimulus and the context but to a different extent.

A detailed description of ContextQ can be found in Balkenius and Winberg (2004), where it was shown how the algorithm could learn a number of cognitive experiments including task-switching, a version of the Wisconsin Card Sorting Test, and context sensitive categorization. In Balke-

nus and Björne (2004), it was applied to an attention switch task to model impairments in attention switch in autism.

ContextQ, like ordinary Q-learning, is an off-policy algorithm, which means that it does not need to follow its own policy during learning (Sutton and Barto, 1998). This is a very important property when the reinforcement learning system is used as a component in a larger architecture. Since many different subsystems can suggest actions, it is not always the output of the reinforcement learning system that is used, but it should still learn the consequences of such actions. For example, reflex actions can be triggered directly by external stimuli and the reinforcement learning system could listen-in to these associations and learn to produce them voluntarily.

2.2. Learning to Attend

ContextQ as described above requires that the current context is explicitly given as input. Here, we extend the model with the ability to learn when to select a stimulus for attentive processing (Fig. 2). In order to achieve this, we add a single output to the ContextQ model that tells the sensory system to attend to the stimulus. When this happens, the currently attended stimulus will be included as a part of the context code.

We now need to consider what properties the combined stimulus selection and context system should have. A simple example will show some of the fundamental requirements of the system. Assume that the system encounters a stimulus sequence A, B, B. The task is to learn that a B preceded by an A is to be ignored, while when preceded by another B, a B should evoke a response. Using only A and B as inputs, a reinforcement system would behave as if responses to B were rewarded half of the time. With ContextQ, the correct behavior would be learned if the foregoing stimulus is used as context. In this case the reaction to B would be extinguished in the context of A. The solution to use the previous stimulus as context is clearly *ad hoc* and something more is needed in the general case.

Now assume that the sequence also contains other stimuli but the same rule holds: X Y Z A X Z B Y X Z B X. In this case, to store the preceding stimulus will no longer work. We now put the context system under reinforcement control by allowing ContextQ to select a stimulus. It is clear that if A is attended, and hence becomes a part of the context, it can be used to influence whether the system reacts to B or not regardless of how many stimuli occur in between. Since the engagement of attention is controlled in the same way as other actions, it too is under contextual control. It thus becomes possible to learn more complicated relations between stimuli.

Since the system must select A to be stored in working memory several time steps before the reinforcement is re-

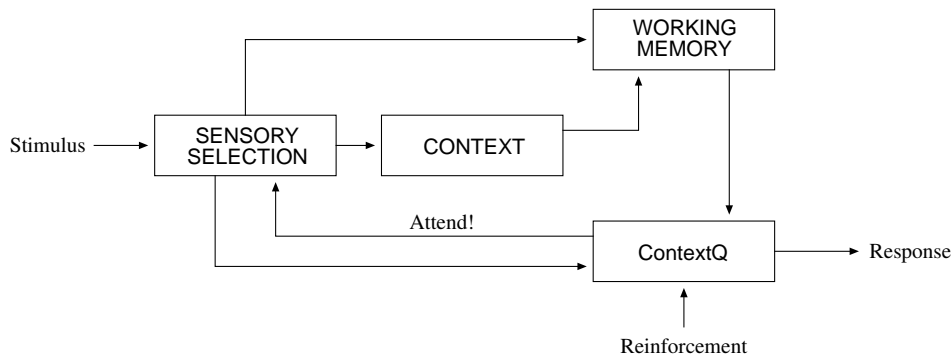


Figure 2. The Model

ceived, it is necessary that the reinforcement can influence behaviors that occurred earlier in the chain. To make this possible, the ContextQ algorithm was extended with an eligibility trace that makes earlier responses eligible for reinforcement. This extension is done in the same way as the ordinary Q-learning algorithm.

So far, we have assumed that the context consists of a single stimulus that is remembered. This is obviously a simplification. A straight forward extension is to allow several stimuli to make up the context, but this raises a new problem. When should stimuli be removed from the context? There appear to exist two solutions. The first, which was used in Balkenius and Morén (2000) is to reset the context code when something unexpected occurs. The second is to let each stimulus that is a part of the context decay over time, possibly as a consequence of new stimuli being made part of the context. Here we opted for the second alternative.

3. Preliminary Simulations

To test the ability of the model to learn to pay attention to stimuli that would change the reward contingencies of actions, we run three simulations with different conditions. In the first simulation, the system was required to learn to respond to an B except when preceded by an A. As described above, the previous stimulus was used as context. As expected, the model quickly learned this task and began to distinguish between the two contexts.

In the second simulation, different distractor stimuli were placed between the A and B. In this case, it is necessary to keep the A in memory even when the distractor appeared in the input. We first tested this task with a tapped delay-line as context. Although the model could easily learn this task it is sensitive to the exact timing of the stimulus A and B. If the inter-stimulus interval changes, it has to re-learn the task again. Another drawback of using a tapped delay-line is that in a more realistic situation, the amount of

stimulus data to store becomes intractable.

Finally, we added a working memory system that could be controlled by a second output from the ContextQ module. In this case, learning progresses much more slowly, but the system is eventually learns to attend to the A and use it as a contextual stimulus. The increased learning time is a consequence of all the initial attentional shifts that disrupts the learning in the initial phase.

4. Discussion

We have extended an earlier model of context sensitive reinforcement learning with the ability to control attention and working memory. Preliminary simulations shows that the extended architecture is able to use its control of attention to explicitly store the appropriate stimuli in working memory and use them as contextual cues.

The main component of the system was inspired by the function of the basal ganglia in working memory and one future goal is to bring the model closer to the actual physiology of the basal ganglia. Another goal is to apply the model to developmental disorders (cf. Balkenius and Björne, 2001, 2004).

Acknowledgments

The code for the simulations is available as part of the Ikaros project: <http://www.lu.se/Ikaros/>

References

- Balkenius, C. (2000). Attention, habituation and conditioning: Toward a computational model, *Cognitive Science Quarterly*, 1, 2, 171-214.
- Balkenius, C., and Björne, P. (2001). Toward a robot model of attention-deficit hyperactivity disorder (ADHD). In Balkenius, C., Zlatev, J., Kozima, H., Dautenhahn, K.,

- and Breazeal, C. (Eds.), *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies, 85.
- Balkenius, C., and Björne, P. (2004). First steps toward a computational theory of autism. In Berthouze, L. et al. (Eds.), *Proceedings of the Fourth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies.
- Balkenius, C. and Hulth, N. (1999). Attention as selection-for-action: a scheme for active perception. In Schweitzer, G., Burgard, W., Nehmzow, U., and Vestli, S. J. (Eds.), *Proceedings of EUROBOT '99* (pp. 113-119). IEEE Press.
- Balkenius, C., and Morén, J. (2000). A computational model of context processing. In Meyer, J.-A., Berthoz, A., Floreano, D., Roitblat, H. L., Wilson, S. W. (Eds.) *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, Cambridge, MA: MIT Press.
- Balkenius, C., and Winberg, S. (2004). Cognitive modeling with context sensitive reinforcement learning. In Malec, J. et al. (Eds.), *Proceedings of AILS 2004*.
- Bouton, M. E. (1991). Context and retrieval in extinction and in other examples of interference in simple associative learning. In L. W. Dachowski and C. F. Flaherty (Eds.), *Current topics in animal learning: Brain, emotion, and cognition* (pp. 25-53). Hillsdale, NJ: Erlbaum.
- Bradshaw, J. L. (2001). Developmental disorders of the frontostriatal system: neuropsychological, neuropsychiatric and evolutionary perspectives. Hove: Psychology Press.
- de Haan, M., Pascalis, O., and Johnson, M.H. (2002). Specialization of neural mechanisms underlying face recognition in human infants, *Journal of Cognitive Neuroscience*, 14, 199-209.
- Frank, M. J., Loughry, B., and O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, & Behavioral Neuroscience*, 1, 137-160.
- French, R. M. (1999). Catastrophic Forgetting in Connectionist Networks. *Trends in Cognitive Sciences*, 3, 128-135.
- Gilbert, P. F. C. (2001). An outline of brain function. *Cognitive Brain Research*, 12, 61-74.
- Hikosaka, O., Miyashita, K., Miyachi, S., Sakai, K., and Lu, X. (1998). Differential roles of the frontal cortex, basal ganglia, and cerebellum in visuomotor sequence learning. *Neurobiology of learning and memory*, 70, 137-149.
- Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reviews*, 80, 953-978.
- Hyland, B. I. (2000). Relationships of substantia nigra dopamine neurone activity to behaviour. In Miller, R. and Wickens, J.R. (Eds.), *Brain dynamics and the striatal complex* (pp. 1-27). Amsterdam: Harwood Academic Publishers.
- Johnson, M.H. (1998). Developing an attentive brain, In Parasuraman, R. (Ed.), *The attentive brain* (pp. 428-443), Cambridge, MA: MIT Press.
- Kimura, M. (1995). Role of basal ganglia in behavioral learning. *Neuroscience Research*, 22, 353-358.
- Lewis, S. J., Cools, R., Robbins, T. W., Dove, A., Barker, R. A., and Owen, A. M. (2003). Using executive heterogeneity to explore the nature of working memory deficits in Parkinson's disease. *Neuropsychologia*, 41, 645-654.
- Lewis, S. J., Dove, A., Robbins, T. W., Barker, R. A., and Owen, A. M. (2004). Striatal contributions to working memory: a functional magnetic resonance imaging study in humans. *European Journal of Neuroscience*, 19, 755-760.
- Lin, C. C. H., Hsiao, Chuhsing K., Chen, Wei J. (1999). Development of Sustained Attention Assessed Using the Continuous Performance Test among Children 6-15 Years of Age. *Journal of Abnormal Child Psychology*, 27, 403-412.
- Middleton, F. A., and Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, 266, 458-461.
- Miller, R. and Wickens, J.R. (Eds.) (2000). *Brain dynamics and the striatal complex* Amsterdam : Harwood Academic Publishers.
- Niebur, E. and Koch, C. (1998). Computational architectures for attention, In Parasuraman, R. (Ed.), *The attentive brain* (pp. 163-186), Cambridge, MA: MIT Press.
- Oades, R. D. (2000). Differential measures of "sustained attention" in children with attention-deficit/hyperactivity or tic disorders: relations to monoamine metabolism. *Psychiatry Research*, 93, 165-178.
- Parasuraman, R (Ed.). (1998). *The attentive brain*, MIT, Cambridge.

- Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*, 89, 1009–1023.
- Ruff, H.A. and Capozzoli, M.C. (2003). Development of attention and distractibility in the first 4 years of life, *Developmental Psychology*, 39, 5, 877–890.
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Watkins, C. J. C. H., and Dayan, P. (1992). Q-learning. *Machine Learning*, 9, 279–292.