First steps towards a computational theory of autism

Balkenius, Christian; Zücker Björne, Petra

First Steps Toward a Computational Theory of Autism

Christian Balkenius
christian.balkenius@lucs.lu.se

Petra Björne
petra.bjorne@lucs.lu.se

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

Abstract

A computational model with three interacting components for context sensitive reinforcement learning, context processing and automation can autonomously learn a focus attention and a shift attention task. The performance of the model is similar to that of normal children, and when a single parameter is changed, the performance on the two tasks approaches that of autistic children.

1. Introduction

Can computational models of cognitive development help understand developmental disorders? A developmental architecture attempts to explicitly formulate the computational processes required for cognitive development. The goal of such architectures is either to better understand cognitive development or to construct developmental control systems for robots (Weng and Zhang, 2002). To the extent that such an architecture claims to describe or parallel human development, it should apply to developmental disorders such as autism. This is the goal of the research reported here were we describe a minimal implementation of a developmental architecture.

Our aim is to design an architecture that can account for normal as well as abnormal development and we think it is a mistake to attempt to formulate a model of autism as such. Instead, we believe that a satisfactory model of autism must have its basis in a model of normal cognitive development.

In addition, it should be possible to perform formal experiments with the implemented model and compare its performance with both that of normal children and that of children with developmental disorders. Below, we show how an architecture incorporating systems for reinforcement learning, context processing and automation can learn an attentional task that has been tested on normal as well as autistic children. The parameters of the system can be altered in such a way that it behaves as an autistic child in an attention task.

The architecture we present is based on biological principles in the sense that the different components parallel the functional roles of different brain regions, but we do not claim to model those regions in any detailed way. Instead, it is the overall interactions between the components that we strive to understand.

1.1 Autism

Autism is a pervasive developmental disorder, with qualitative impairments within three main areas: social interaction, communication and imagination (American Psychiatric Association, 2000). Persons with autism often engage in stereotyped and elaborate ritualistic behaviors and they have difficulties in initiating and ending activities. They also qualitatively differ in their use and understanding of nonverbal activities that influence the interaction with others, such as eye-contact, posture, facial expression and gestures. This makes it difficult for persons with autism to initiate and sustain peer relations, to share interests and to cooperate with others. Persons with autism will not have access to some of the information provided by a social context. Functional knowledge of the world will be further hampered by deviant sensory input, attention and poor integration of different modalities.

Whenever children with autism do not develop language or only a poor one, they will not compensate for this by use of alternative communication, such as gestures or facial expressions. When language develops, it will still leave the person with a difficulty to communicate with others in a proper way, and the language will mostly consist of imperative requests rather than declarative interaction. Persons with high-functioning autism or Asperger's syndrome who have a quantitatively high verbal performance, still have difficulties with the pragmatic use of language.

Persons with autism have a restricted range of interests, their thinking and behavior is rigid. It seems as if they have difficulties in integrating pieces of knowledge into a functional whole. He or she will

lack in the ability to understand complex information, in particular if the information changes from time to time, which in turn will lead to impaired generalization of knowledge. Knowledge will be too closely connected with a certain place, person or time and will thereby be restricted in its application. As a result, persons with autism have difficulties using alternative strategies for solving everyday problems.

1.2 Attention Deficits

Attention is probably one of the most basic abilities necessary to learn about the world and a deficit in this area would to some extent explain the cognitive developmental pathway characteristic for persons with autism.

It might not be the case that persons with autism avoid eye-contact with others, but it seems as if they do not actively search for information or reinforcement from the social environment to the same extent as others. Nor do they seem to attend to a specific location, expecting relevant information to appear there. The attention of persons with autism seems to be quite unpredictable (Damasio and Maurer, 1978). It is also known that persons with autism find it easier to understand visually presented information than auditory.

Attention is not a unitary function, but consists of several components, such as disengaging attention from current focus, orienting to new focus and engaging it there, as well as selecting relevant stimuli and sustaining attention over time in the presence of distracting stimuli. Different components of attention are subserved by different functional systems of the brain, which have to interact successfully for attention to work in an integrated manner (Hopfinger, Buonocore and Mangun, 2000, Hopfinger, Woldorff, Fletcher and Mangun, 2001, Burack et al., 1997).

Persons with autism have difficulties in moving their attention from one focus to another in a smooth manner, thereby integrating different areas of interest into a whole. However, it is not clear what the cause is of this deficit, though there are some indications as to which parts of the brain could be involved (Townsend and Courchesne, 1994, Townsend, Courchesne and Egaas, 1996, Burack et al, 1997, Minshew, Sweeney and Bauman, 1997).

Akshoomoff and Courchesne (1992) and Courchesne et al. (1994) tested 8 children with autism diagnosed with cerebellar abnormalities. The object of the study was to test the hypothesis that the cerebellum contributes to the smooth control of mental focus of attention.

Visual stimuli were green distracters and red target flashes of light. Auditory stimuli were high (target) and low tones, 2-kHz and 1-kHz respectively. A red flash required the person to press a button and it also served as an incitement to change focus to the auditory modality. In the same manner the high tone, the target, required the tested person to press a button and then shift the attention from the auditory to the visual modality, ignoring the previous target modality.

Both visual and auditory stimuli were 50 ms in duration, and the inter-stimulus intervals varied between 450 and 1,450 ms. The target within the visual modality served as a cue to shift attention to the auditory modality and vice versa. A hit was scored when the response occurred between 200 and 1,400 ms after the presentation of the target. Failure to respond within this time window was a miss and an erroneous response was scored as a false alarm given that it occurred before the target stimulus (Akshoomoff and Courchesne, 1992).

A similar setup tested focused attention. In this case, the participants were required to press the button only following the red flash, while ignoring the auditory stimuli completely. Children with autism performed similarly to normal mental age matched children in the focus attention task, but they were impaired in their ability to shift attention accurately and rapidly. When a target was presented at 2.5 s or less after the previous correct target detection, the children with autism only detected 58.9 % of the targets, compared with 78% targets detected by the normal children. When more time elapsed the children with autism did not perform significantly worse than did the controls or from their own performance levels in the focus attention task.

The striking difference in the responses of children with and without autism in this experiment makes it an ideal test case for a computational model of autism. The task is simple enough to be implemented completely in a computer simulation, but still illustrates many of the difficulties people with autism have with attention.

1.3 Earlier Models

There have been several previous attempts to formulate computational models of autism. To our knowledge the first was developed by Cohen (1994) who argues that an inadequate number of hidden units in a back-propagation network can explain the deficits in autism in that a too high number of hidden units compared to the task complexity leads to rote learning and poor generalization.

Gustafsson (1997) proposed a model based on Kohonen's (1995) self-organizing maps where excessive inhibition results in the inadequate formation of cortical feature maps. A surprisingly similar model was independently developed by de Carvalho, de Carvalho Ferreira and Fiszman (1999) who also suggest that autism results from ill-developed and highly discriminative cortical maps, without overlap between different concepts. In this model, the initial amount

of nerve-growth factor is assumed to influence the map formation rather than the level of inhibition as in Gustafsson's model.

A later model based on cortical feature maps was put forward by Papliński and Gustafsson (2002). They present a simulation in which deficient feature maps result from a familiarity preference in stimulus selection rather than low probability of attentional shifts between stimuli. Their conclusion that a familiarity preference (i.e. increased selection of a small set of already learned stimuli) will result in more specialized feature maps is obviously correct, but it is not clear whether this should be considered the cause of the disorder or a consequence. Also, to rule out attentional deficits in autism based on a single very simplified attention model appears to be too hasty.

A drawback with these models is that their operation is not compared in any way to experimental data. Since no behavior is generated by either of the models, it not possible to compare the models with any behavioral data. Nor is the suggestion that inadequate feature-maps would result in the autistic syndrome substantiated in any solid way.

O'Loughlin and Thagard (2000) describe another type of computational model that sets out to explain the lack of central coherence in autism (Frith, 1989). The connectionist model is based on a constraint satisfaction network (McClelland and Rumelhart, 1989), and the authors suggest that the lack of central coherence results from a too high level of inhibition compared to the level of excitation. The model is attractive in that it captures both normal and autistic cognition. It can also be applied to tests that are often used with autistic children, such as a false-belief task (Sally-Ann), and a homograph task, without postulating a specific 'theory of mind' module.

Kamawar, Garfield and de Villiers (2002) pointed out that this model predicts that children with autism would have problems with the false photo task, which is structurally isomorphic to the Sally-Ann task. This is not the case however, and this highlights the main problem with this model: it is not grounded in learning or perception. There are nodes with content like "Sally thinks marble is in basket (but it's not)", but there is no explanation of how these nodes come about. Propositions are treated as atoms without content.

2. A Computational Model

In this section, we describe the developmental model that was used in the simulations of the attention task. This architecture is not intended as a complete model of cognitive development. However, we think it contains some very important components that will also be included in a more complete architecture.
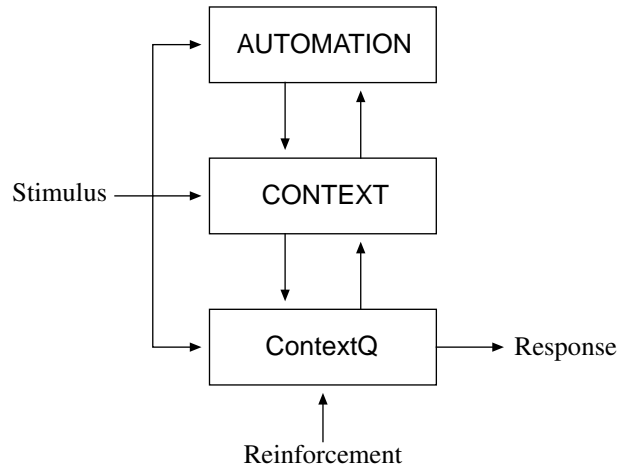

Figure 1: Overview of the Model. A ContextQ system learns associations between stimuli and responses based on the reinforcement. The CONTEXT system controls in what context each stimulus-response association should be used. The AUTOMATION system learns to produce stimulus-triggered contextual shifts.

The model consists of three main parts (Fig. 1). A Q-learning system learns associations between stimuli and responses based on the reinforcement it receives (Watkins and Dayan, 1992). The ContextQ algorithm used here is the result of adding contextual inputs to the Q-learning algorithm (Balkenius and Winberg, 2004).

The reinforcement learning is here under the influence of a context system (Balkenius and Morén, 2000). The function of the context system is to integrate sensory input over time to create a code for the current context (Balkenius, 2000). Here, it operates as a working memory for the last potential target that the system reacted to. This allows the contextual Q-learning module to shift between different behavioral strategies depending on the current context. Unlike other reinforcement learning methods that learn a single policy, this allows for the gradual development of many different behaviors over time, which is a central property of a developmental architecture.

Finally, an automation system learns associations from stimuli to context changes and makes it possible for the context system to change state much quicker when a target appears. The role of the automation system is here to produce conditioned context shifts as a result of the presentation of a target stimulus which will speed up the shift between different tasks. The three components are described in detail below.

2.1 Context Processing

When a location x is attended internally or in the environment, the current input state of the atten-

tional system, the *focus of attention*, is given by $s(x) = \langle s_0, s_1, \ldots, s_n \rangle$, which is a sensory code for the attended stimulus at x, and $L(x)$, which is a code for the attended location x itself. Given a set of attentional fixations $X \subset V$, a binding code $d(X)$ is calculated as the sum of the outer product of the two vectors $s(x)$ and $L(x)$ for each location $x \in X$,

$$d(X) = \sum_{x \in X} s(x) \otimes L(x).$$

The binding d is a form of tensor coding of conjunctions as suggested by Smolensky (1999). Each binding code represents a context or a partial context. The output from the binding stage is integrated over time in a binding vector $b = \langle b_0, b_1, \ldots, b_p \rangle$. This vector is given by,

$$b_i(t+1) = \frac{b_i(t) + \delta d_i + E_i}{\sum_{j=0}^{p} (b_j(t) + \delta d_j + E_j)},$$

that is, the values b_i integrates their input over time and the result is normalized to achieve a form of soft competition between the binding elements. The term E_i is an additional input that will be described further below.

In our earlier models of context processing (e.g. Balkenius and Morén, 2000), individual binding codes where recognized by an additional processing stage that integrated these into a context code. In the current implementation, this stage was not included. Instead, the context output c was set to the output of the binding nodes, $c_i = b_i$.

Also, the location code was here implemented in the most trivial way using only a single location that corresponds to attending to the stimulus location that triggered a response, i. e., $L(x) = a_0$, where a_0 is the response of the system as explained below.

2.2 Contextual Reinforcement Learning

To learn associations between stimuli and responses in a context dependent way, we use ContextQ (Balkenius and Winberg, 2004). This algorithm can learn a large number of distinct behaviors in different contexts and also generalizes from one context to the next in a very efficient way. Here we only give a minimal description of the algorithm.

Let each state be represented by a state vector $s = \langle s_0, s_1, \ldots, s_n \rangle$ and let $\{a_0, a_1, \ldots, a_m\}$ be a discrete set of actions. The Q-function can be estimated in the standard way as,

$$Q(s, a_j) = \sum_{i=0}^{n} s_i w_{ij},$$

where the update rule is

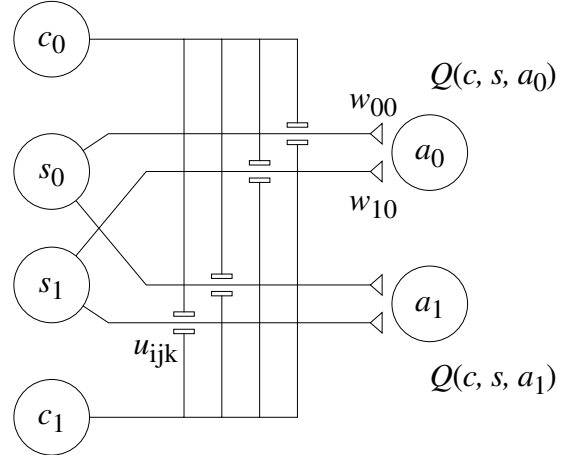$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \alpha s_i \Delta Q_t,$$

Figure 2: The approximation of $Q(c, s, a_j)$ as an artificial neural network with shunting inhibition from the context nodes c_k to the association between a state node s_i and an action node a_j.

j is the index of the last action and,

$$\Delta Q_t = \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right].$$

Each weight is updated according to the error in the Q-function multiplied with the value of the state component s_i. This means that only components of the state that contributed to the selected action will be updated. It is clear that the linear approximator will generalize learning to states that are similar to each other (See Sutton and Barto, 1998). This is so far the original Q-learning algorithm described by Watkins and Dayan (1992) to which we now add the contextual input.

Let the context be described by a vector $c = \langle c_0, c_1, \ldots, c_p \rangle$ received from the context system. We include the context in the linear estimator with additional weights u_{ijk}, which relates each association w_{ij} to the context c_k:

$$Q(c, s, a_j) = \sum_{i=0}^{n} s_i w_{ij} I_{ij},$$

where,

$$I_{ij} = \prod_{k=0}^{p} (1 - c_k u_{ijk}).$$

In neural network terms, I_{ij} can be seen as shunting inhibition from the context of the association from the state to the action (Fig. 2). We now need to consider how the learning rule should be changed to reflect the new context sensitive estimator.

The learning rule for the contextual weights is,

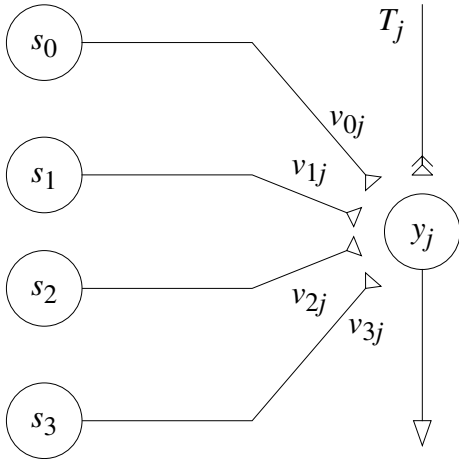$$u_{ijk}^{(t+1)} = u_{ijk}^{(t)} - \beta s_i c_k \Delta Q_t.$$

Figure 3: The automation network associates its input with the target output.

Depending on the sign of ΔQ_t, two different set of learning constants are used. When $\Delta Q_t > 0$, $\alpha = \alpha^+$ and $\beta = \beta^+$. Otherwise, $\alpha = \alpha^-$ and $\beta = \beta^-$. Typically, $\alpha^+ > \alpha^-$ and $\beta^+ < \beta^-$. In other words, the inhibition from the current context will primarily increase the association between the current state and the selected action when the actual reinforcement is lower that the expected reinforcement. For a discussion of the psychological motivation for this learning rule, see Balkenius (2000) and Balkenius and Morén (2000).

In the simulation below, only a single action was used, and the Q-value was treated as the probability that a response would be generated. The single action was also used as an input to the context system, as described above.

2.3 Automation

Automation is the learning of stimulus-response associations by observing the actions of another system in such a way that they can be automatically produced later without the aid of the system that originally produced them. This process is similar to some forms of classical conditioning, for example as it reveals itself in the conditioned eye-blink response (Yeo and Hesslow, 1998).

We model automation in a very basic way using two types of inputs and one output. The input $s = \langle s_0, s_1, \ldots, s_q \rangle$ is associated with the output y through the associative weights v_{ij} as,

$$y_j = \sum_{i=0}^{q} s_i v_{ij}$$

The delta rule (Widrow and Hoff, 1960) is used for learning the target output T:

$$v_{ij}(t+1) = v_{ij}(t) + \epsilon[T_j(t) - y_j(t-\tau)]s_i(t-\tau).$$

The time difference τ is used to let the automation system produce its output before the system it taps into and ϵ is the learning rate. In conditioning terms, τ is the optimal inter-stimulus interval for conditioning.

The target for the automation system is the change in the output of the context system (cf. Klopf, 1988):

$$T_j(t) = c_j(t) - c_j(t-1),$$

and the output from the automation module is sent to the context system through the extra input $E_j(t) = \varphi y_j(t)$.

3. Simulations

We simulated the focus attention and shift attention tasks described above in section 1.2 (Akshoomoff and Courchesne, 1992, Courchesne et al., 1994). In the simulations, the following constants where used: $\alpha^+ = 0.2$, $\alpha^- = 0.1$, $\beta^+ = 0.1$, $\beta^- = 0.2$, $\gamma = 0$, $\delta = 0.035$, $\epsilon = 0.01$, $\varphi = 50$ for the normal group and 0 for the autistic group, and $\tau = 2$. A simulated time step corresponded to 50 ms. Of these parameters only two are critical for the result. The constant δ, which governs the speed of a context shift, was selected to reproduce the response characteristics of the autistic persons as closely as possible. The constant φ, which describes the influence of the automation system on the context, was subsequently set to a value that would produce a close fit to the normal group. The model received a reinforcement of 1 when it responded correctly and a reinforcement of -1 for a false alarm. Reinforcement was given throughout the test.

Fig. 4 shows the activation of the different contexts at different times during training of the tasks. The curves are symmetrical since the total context activation was normalized to 1 and there were two main contexts that were used when the task was learned. For the autistic parameters which would disconnect the automation system (i. e. $\beta = 0$), the context shifts are slow throughout the simulations as shown in Fig. 4A. For the normal simulation ($\beta = 50$), the automation system learns to generate context shifts when appropriate. Fig. 4B, C and D shows the development of the conditioned context shifts during learning.

Fig. 5 shows the experimental data from Akshoomoff and Courchesne (1992) and Courchesne et al. (1994) together with the simulation results. In the focus attention task, the model correctly learns what stimuli to attend to and what to ignore. The model learns the task perfectly both for the autistic and the normal case, which differs from the experimental data where subjects only reach a level of approximately 90% correct hits.

In the shift attention task, both the simulation and experimental results show an increasing hit rate
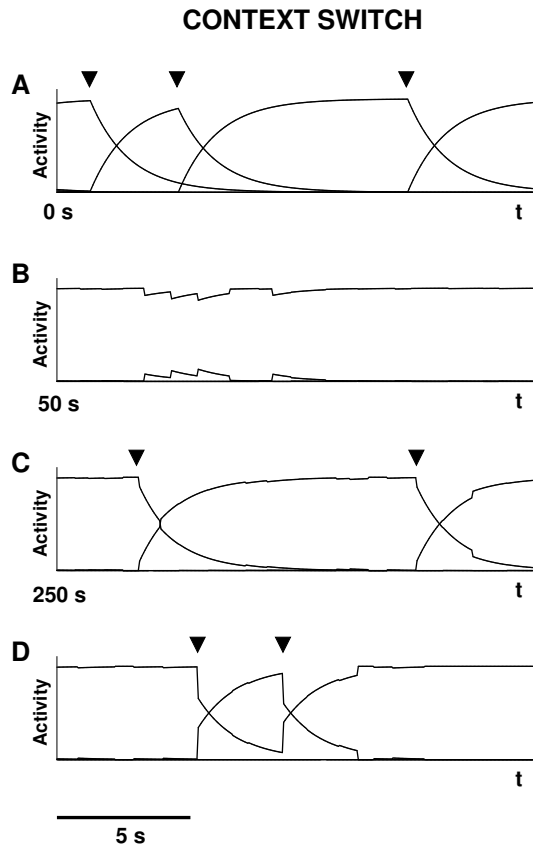
CONTEXT SWITCH

Figure 4: Context shifts at different times during training. The curves represent the activation level of the different contexts. The black triangles indicate when a new context cue occurs in the input. A. Initial shifts without automation input. Slow shifts occur early in training both in the normal and autistic condition. Since no learning has yet occurred, more than two contexts are simultaneously active. B. Initial automation input which is not yet correctly correlated with the correct context shifts causes small but incorrect changes in the context activation. C. Shifts after 50 s of training. The automation system starts to correctly influence the context shifts at the correct time. D. Shifts after complete training. The automation system now almost completely shift the context on its own.

starting close to 80% for the short interval and reaching a level around 90% for the longer intervals. For the autistic group, both the model and the subjects respond at a level of approximately 60% and approaches 85% when the interval increases. The main difference is in the interval 2.5–4.5 seconds, where the simulations shows a lower level than the experiment.

4. Discussion

The simulations show that the computational model can reproduce the main behavior of normal and autistic persons in the focus attention and shift attention tasks.

In the focus attention task, the model outperforms both the normal and autistic person. The reason for this is most likely that there are no distracting stimuli outside the experimental set up in the simulations. In this sense, the simulations differ from a natural experiment where all kinds of potentially distracting stimuli are simultaneously present.

In the shift attention task, the behavior of the model more closely resembles that of the normal and autistic subjects. There are two differences however. First, the model performs worse than the autistic persons in the intermediate range of 2.5–4.5 seconds where autistic persons in the experiment perform as well as normals. Second, the hit rate does not decrease as quickly as for the experimental subjects in the autistic group.

Since we have not had access to the original data by Akshoomoff and Courchesne (1992) and Courchesne et al. (1994) however, we have not been able to statistically compare the model with the experimental data. It is thus unclear whether these small discrepancies in the means are significant or not.

Another aspect of the model is that it illustrates how a computational system can learn to participate in a complex cognitive experiment based only on the reinforcement it receives. This is a type of learning situation that has traditionally been seen as outside the range of reinforcement-based models.

The three subsystems of the model can be linked to the functional roles of different brain structures (cf. Doya, 2000). The reinforcement learning system corresponds to the basal ganglia. It learns reinforcement-sensitive stimulus-response associations. The context sensitivity in the model presented here is however a step beyond most current computational models of the basal ganglia which typically do not include this extra input.

The automation system has a function similar to the cerebellum in that it learns stimulus-response association on the repetition of a response after the occurrence of a certain stimulus independently of reinforcement. The context system corresponds to the hippocampus and prefrontal cortex (Balkenius, 2000, 2003).

As a consequence of these identifications with brain structures, the model links together autism as a problem in context sensitivity and attention shift with the suggested role of cerebellar dysfunction in autism. This is in line with the hypothesis put forward by Skoyles (2001) that autism is caused by a disconnection of the cerebral cortex from the cerebellum.

The current model includes only the minimal implementation of each subsystem required to produce the desired results. There is also a minimum of interaction between the subsystems. In the future, the model will be extended to a more complete develop-
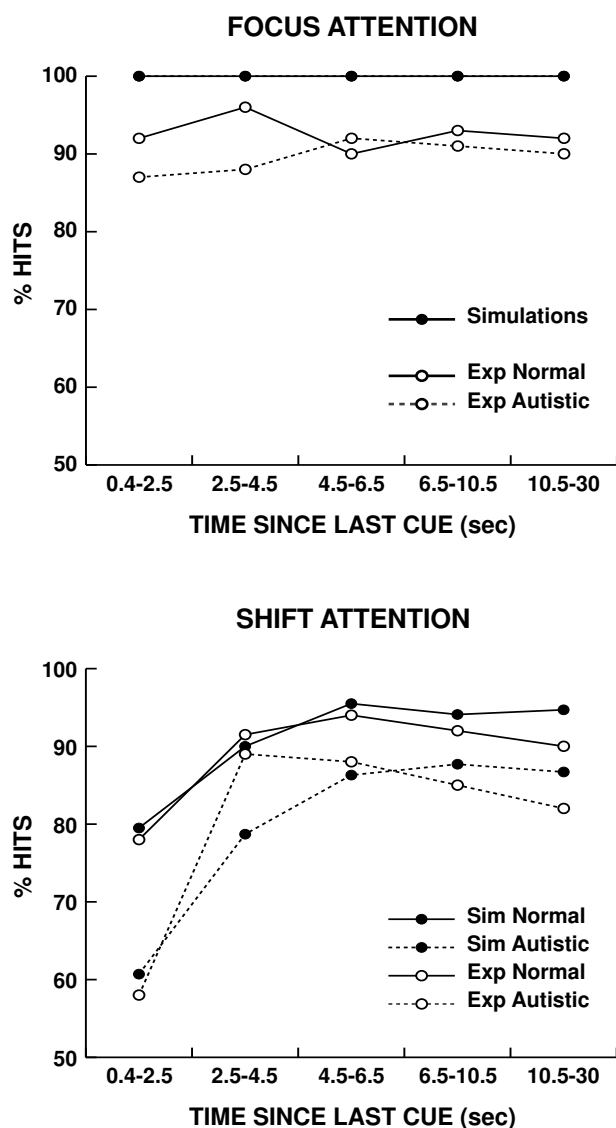
FOCUS ATTENTION

SHIFT ATTENTION

Figure 5: Simulations of the focus attention and shift attention tasks compared to experimental data from Akshoomoff and Courchesne, 1992 and Courchesne et al. 1994.

mental architecture.

First, the context module will be extended with contextual categorization that was not included in the simulation here (cf. Balkenius and Morén, 2000, Balkenius, 2003). Second, the control of attention will be under reinforcement control and a separate stage with sensory selection will be added (cf. Balkenius, 2000). Third, there will be full interaction between the subsystems that will allow, for example, automation of responses, reinforcement control of attention, context shifts and working memory. Finally, subsystems for involuntary attention will be included (cf. Balkenius, 2003). Another future goal will be to apply the model to a wider range of experiments.

To conclude, we have presented a computational model of autism that can be directly tested against experimental data from autistic persons, and have shown that it can reproduce the main differences between normal and autistic children in an experimental task. Furthermore, we have demonstrated that the model can learn the experimental task without being specifically designed for this task. We have finally suggested that the components of the model map onto different brain structures and that disconnection of brain structures suggested to be responsible for autism do in fact produce autistic behavior in the model.

Acknowledgements

References

Akshoomoff, N. A. and Courchesne, E. (1992). A new role for the cerebellum in cognitive operations. *Behavioral Neuroscience*, 106, 5, 731–738.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV)*.

Balkenius, C. (2000). Attention, habituation and conditioning: Toward a computational model, *Cognitive Science Quarterly*, 1, 2, 171-214.

Balkenius, C. (2003). Cognitive processes in contextual cueing. In Schmalhofer, F., Young, R. M., and Katz, G. (Eds.), *Proceedings of the European Cognitive Science Conference 2003* (pp. 43-47). Mahwah, NJ: Lawrence Erlbaum Associates.

Balkenius, C., and Morén, J. (2000). A computational model of context processing, In Jean-Arcady Meyer, Alain Berthoz, Dario Floreano, Herbert L. Roitblat, Stewart W. Wilson, (Eds.) *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behaviour*, Cambridge, Mass., 2000. The 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, in press*.

Burack, J. A., Jenns, E. T., Stauder, J. E. A., Mottron, L., and Randolph, B. (1997). Attention and autism; behavioral and electrophysical evidence. In Cohen, D. J. and Volkmar, F. R. (Eds.), Handbook of autism and pervasive developmental disorders (pp. 226–247), New York: John Wiley & Sons, Inc.

de Carvalho, L. A. V., de Carvalho Ferreira, N., and Fiszman, A. (1999). A Neurocomputational

Model for Autism, In *Proceedings of IV Brazilian Conference on Neural Networks - IV* (pp. 888–999).

Cohen, I. L. (1994). An artificial neural network analogue of learning in autism. *Biological Psychiatry*, 36, 5–20.

Courchesne, E., Townsend, J., Akshoomoff, N. A., Saioh, O., Yeung-Courchesne, R., Lincoln, A. J., James, E. H., Haas, R. H., Schreibman, L., and Lau, L. (1994). Impairment in shifting attention in autistic and cerebellar patients. *Behavioral Neuroscience*, 108, 5, 848–865.

Damasio, A. R., and Maurer, R.G. (1978). A neurological model for childhood autism, *Archives of Neurology*, 35, 777–786.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, 10, 732–739.

Frith, U. (1989). *Autism: Explaining the enigma.* Oxford: Blackwell.

Gustafsson, L. (1997). Inadequate cortical feature maps: a neural circuit theory of autism. *Biological Psychiatry*, 42, 1138–1147.

Hopfinger, J. B., Buonocore, M. H., and Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, 3, 3, 284–291.

Hopfinger, J. B., Woldorff, M. G., Fletcher, E. M., Mangun, G. R. (2001). Dissociating top-down attentional control from selection perception and action, *Neuropsychologia*, 39, 1277–1291.

Kamawar, R., Garfield, J. L., and de Villiers, J. (2002). Coherence as an Explanation for Theory of Mind Task failure in Autism, *Mind and Language*, 17, 3, 266–272.

Kohonen, T. (1995). *Self-organizning maps.* Berlin: Springer-Verlag.

Klopf, A. H. (1988). A neuronal model of classical conditioning. *Psychobiology*, 16, 85–125.

McClelland, J. L., and Rumelhart, D. E. (1989). *Explorations in parallel distributed processing.* Cambridge, MA: MIT Press.

Minshew, N. J., Sweeney, J. A., and Bauman, M. L. (1997). Neurological aspects of autism, In Cohen, D. J. and Volkmar, F. R. (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 344–369), New York: John Wiley & Sons, Inc.

O'Loughlin, C., and Thagard, P. (2000). Autism and Coherence: A Computational Model, *Mind and Language*, 15, 4, 375–392.

Papliński, A. P., and Gustafsson, L. (2002). An attempt in modelling autism using self-organizing maps. In Wang, L., Rajapakse, J. C., Fukushima, K., Lee, S-Y, and Yao, X. (Eds). *Proceedings of the 9th International Conference on Neural Information Processing*, 4.

Skoyles, J. R. (2001). Is autism due to a cerebral–cerebellum disconnection? *Medical Hypotheses*, 58, 4, 332–336.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence, 46*, 159–216.

Townsend, J., and Courchesne, E. (1994). Parietal damage and narrow "spotlight" spatial attention. *Journal of Cognitive Neuroscience*, 6, 3, 220–232.

Townsend, J., Courchesne, E., and Egaas, B. (1996). Slowed orienting of covert visual-spatial attention in autism: specific deficits associated with cerebellar and parietal abnormality, *Development and Psychopathology*, 8, 563–584.

Watkins, C. J. C. H., and Dayan, P. (1992). Q-learning. *Machine Learning*, 9, 279–292.

Weng, J., and Zhang, Y. (2002). Developmental robots – a new paradigm. In Prince, C. G., Demiris, Y., Marom, Y., Kozima, H., and Balkenius, C. (Eds). *Proceedings of the second international workshop on epigenetic robotics: modeling cognitive development in robotic systems.* Lund University Cognitive Studies, 94.

Widrow, B. and Hoff, M. E. (1960). Adaptive switching circuits. *IRE Western Electric Show and Convention Record*, part 4, 96–104.

Yeo, C., and Hesslow, G. (1998). Cerebellum and conditioned reflexes. *Trends in Cognitive Science*, 2, 9, 322–331.