Learning and Evolution of Control Systems

John A. Bullinaria & Patricia M. Riddell
Department of Psychology, The University of Reading
Reading, RG6 6AL, UK.

Abstract: The oculomotor control system, like many other systems that are required to respond appropriately under varying conditions to a range of different cues, would be rather difficult to program by hand. A natural solution to modelling such systems, and formulating artificial control systems more generally, is to allow them to learn for themselves how they can perform most effectively. We present results from an extensive series of explicit simulations of neural network models of the development of human oculomotor control, and conclude that lifetime learning alone is not enough. Control systems that learn also benefit from constraints due to evolutionary type factors.

Introduction

For us to see objects clearly at different distances, the human oculomotor control system must produce appropriate eye rotations (vergence) and focus changes (accommodation). These responses are correlated and driven by a range of different cues (blur, disparity, looming, texture, etc.) with degrees of reliability and availability that vary with different viewing conditions. There are also numerous age dependent factors, such as the disparity cue only becoming available after about four months, and the ability to accommodate falling steadily between childhood and old age. The control system also needs to adapt on several timescales, initially to compensate for maturational factors such as eyeball growth, later to correct for eye damage or deterioration, and also to reduce the strain under conditions such as repeated near work.

To program such a system would be a formidable task, yet evolution has resulted in an oculomotor control system that learns efficiently to organise itself during childhood to perform appropriately. We have been attempting to formulate a series of neural network models of the oculomotor control system, and to validate them by comparing the models' development against that of humans from birth to adulthood. We are particularly interested in the emergence of cross-links between the vergence and accommodation systems, and the factors responsible for the significant individual differences that are often found. This work should help identify precursors of abnormal development in children and suggest appropriate remedial actions. Also, from a more technological point of view, our endeavours may provide guidance in the development of similar artificial control systems that are required to learn how to organise themselves to make the best use of available inputs. Allowing such systems to learn how to operate is likely to be a more efficient option than trying to program them by hand.

The basic problem, then, is to build a system that learns to take in a range of cues, process them, and output appropriate responses under a range of conditions. The main questions we need to address are :

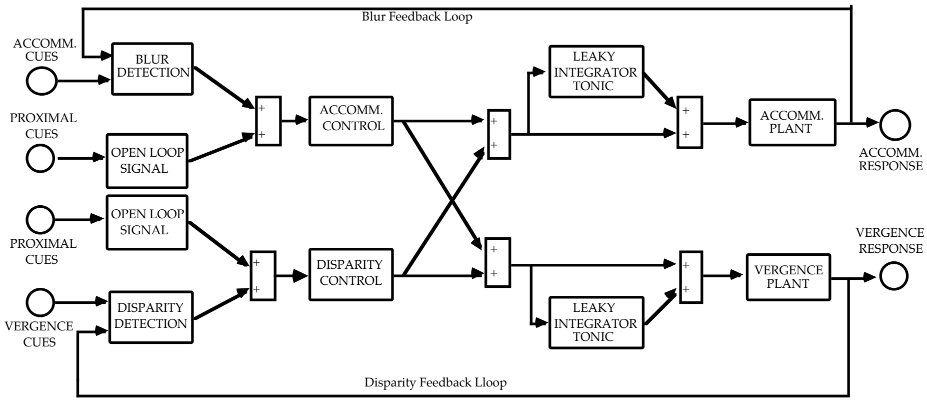

Figure 1: The basic structure of the accommodation and vergence control systems.

1. Do we have to do more than just set up a general architecture and let it learn?
2. How does the system organise itself and what properties emerge?
3. How human-like are the resulting models?
4. Could we improve upon the human system?

We shall begin by describing the basics of the human oculomotor control system, and our attempts to formulate developmental neural network models of it. The lessons learnt from this case study then lead us on to considerations of the importance of evolutionary factors and suggestions on how best to approach the formulation of self-organising control systems more generally.

Neural Network Models of Oculomotor Control

The basic structure of the oculomotor control system takes the form of two cross-linked feed-back loops with appropriate cues driving the relevant responses via controllers, leaky integrator tonics and plants as shown in Figure 1. Several hand crafted engineering style control systems models already simulate the responses of the adult system to unpredictable target sequences in some detail and have been reviewed by Eadie & Cardin [3]. However, these models fail to take into account various developmental factors and have limited biological/physiological realism. Nevertheless, they do provide a convenient starting point for our neural network developmental models.

Our initial modelling approach is to replace the boxes and connections of the control systems models with leaky integrator neurons linked by connections with modifiable weights, for example as shown in Figure 2. These are still fully dynamical systems, but now the models' free parameters are the weights (WA, WV, ... in the figure) that can be learnt using some form of gradient descent algorithm. In this way the system not only learns an appropriate mode of operation, but we can also compare its development against that of humans. Many of the other network details, such as the neuron time constants, can be determined from the same empirical data as the corresponding traditional control models (e.g. [6]). However, if we want to end up with a realistic model, there are numerous additional developmental complications that need to be taken into account: The Blur and Disparity sub-systems both result in input dead-zones that

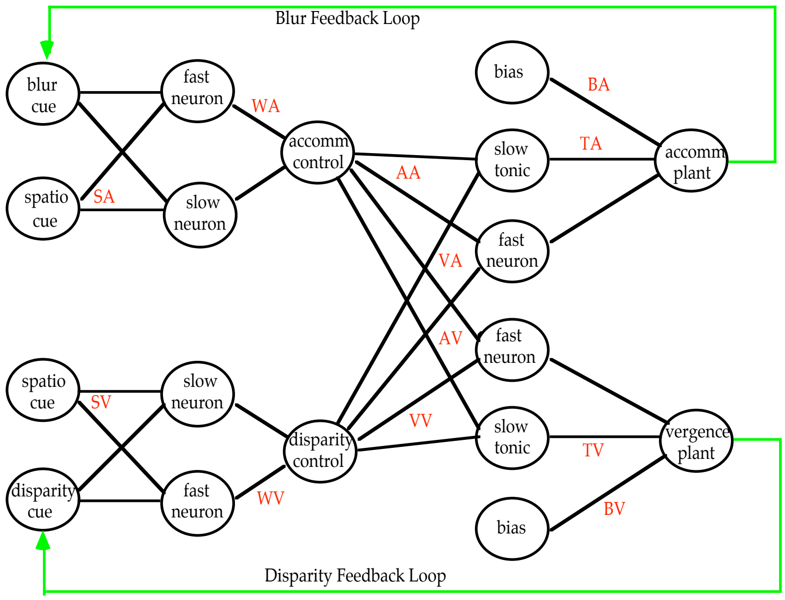

Figure 2: The neural network model based on the systems model of Figure 1.

decrease considerably with age as the eye matures. The growing inter-pupil distance changes the required vergence responses. Some cues (e.g. disparity) become available at later ages than others. The performance of the accommodation plant deteriorates with age. It is not obvious what the cost function for the gradient descent learning should be, or that the learning process should start with all the weights near zero rather than at some larger innate values. Nor is there any real reason why all the weights should have the same learning rates. We also need to determine a reasonably realistic distribution for the training data and how noisy/unreliable it is. Finally, it is important that we coordinate the natural maturational time-scales with the network learning rates.

It turns out that dealing with these developmental complications appropriately is crucial for the system to learn a human-like sequence of network weights. Fortunately, existing empirical data is adequate to determine most of the important details [5]. However, there remain three factors that are particularly difficult to constrain sufficiently accurately without explicit network simulation:

1. The fitness function we use with the gradient descent learning algorithm.
2. The choice of learning rates for the different network weights.
3. The initial (pre-learning) values chosen for each network weight.

Whilst blur and disparity are fairly obvious error components for the fitness function, choosing the regularization (i.e. smoothing) component, and setting the trade-off, is less easy. However, we have already presented an extensive study elsewhere showing that plausible changes to the details of the fitness function lead to relatively minor variations in the patterns of learned weights compared to those resulting from differences in the initial weights and learning rates [2], so we shall concentrate on points 2 and 3 here.

Figure 3 shows the human like responses we get from our models if we manage to get all the developmental details right and train the network weights to asymptote. If we

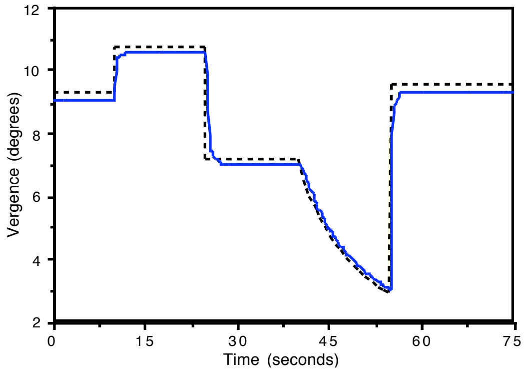

Figure 3: Vergence responses of a model that has learned to perform like a human.

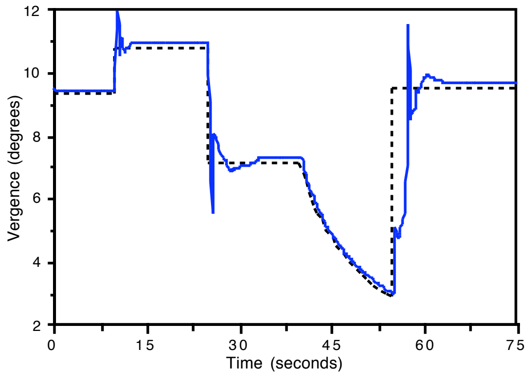

Figure 4: Typical poor vergence responses learned by a rather un-human like model.

are not careful we can easily end up with poor responses like those shown in Figure 4.

The time courses of the weights during training are not particularly illuminating, but the network responses under normal and open loop conditions are directly comparable with traditional measurements of human performance. Figure 5 shows the response gains (i.e. response/stimulus ratios) during training for a typical network. Like human development, we have initial flat responses followed by a period with gains increasing to the adult values [5]. The final accommodation and vergence gains (A and V) are just slightly below one under normal conditions, but when one of the feedback loops are opened, the corresponding gains (VA or AV) fall to around 0.77 and 0.67. We see that, even if we start the training with independent accommodation and vergence sub-systems, the system develops cross-links (i.e. non-zero VA and AV) that allows it to make vergence and accommodation responses even when the corresponding feedback is not available (though usually with gains somewhat less than one). Humans also tend to have gains of nearly one under normal conditions, but show considerable individual differences in their open loop performances and in their developmental histories. It is important to see if our neural network models exhibit similar individual differences.

First, suppose our typical network's initial accommodation input dead-zone were to decrease at a slightly faster rate. Figure 6 shows that this can cause the development of the A response to precede the V response, and this is reflected in the relative values of the cross-link gains (i.e. AV > VA). Now, of particular interest here are the potential

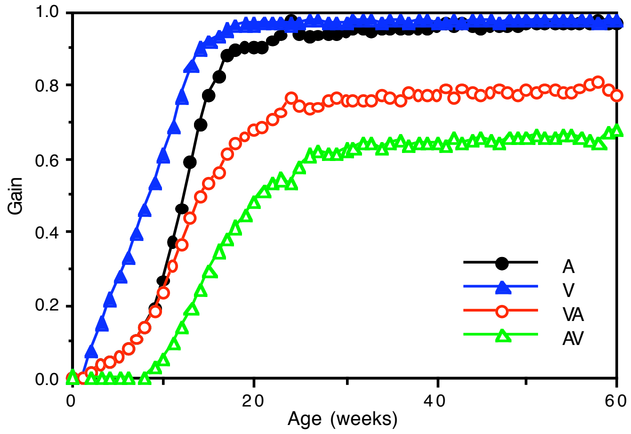

Figure 5: The development of response gains for a typical human-like model.

differences in performance due to dependencies on the initial weights prior to learning and on variations in the pattern of learning rates. To illustrate this, Figure 7 shows that, if our network started learning from large (rather than near zero) innate A weights, we can end up with AV > VA despite the V response developing earlier. In fact, with different plausible choices for each of the initial weights, learning rates and dead-zones, we can generate models with developmental and open loop performances that easily span the whole range of human individual differences. However, the robustness of the learning algorithm and feedback loops ensures that the networks learn good normal adult responses under all but the most extreme pathological conditions.

Clearly then, any initial expectations that all reasonable network variations would lead to the same optimal structure have been proven unfounded. Changing the initial weights and learning rates have a significant effect on the structures (e.g. cross-links) that emerge, with little effect on the normal oculomotor responses. If we also allow the plasticities (i.e. learning rates) to be age dependent, the possibilities are endless.

Evolutionary Factors

Since there are considerable inter-personal differences found within the human population, we should regard the potential variability in our models as a positive feature. The emergence of cross-links between the vergence and accommodation systems with widely varying magnitudes is one classic example of this. However, there are likely to be evolutionary factors which restrict the range of variability of the human system, and these have not yet been included in our current models. We must clearly be very careful about drawing conclusions from any models that learn without taking the constraints of evolution into account.

Given that evolutionary factors will almost certainly have influenced our oculomotor control system, current human performance may simply be insufficient to fully constrain our models. It may be necessary to take into account the whole evolutionary process leading to that performance. This is particularly important given that the goals of evolution and learning within an individual's lifetime will not necessarily be the same. For example, evolution may prevent a young system from learning a good solution which frequently proves to be detrimental in later life. Other constraints may arise from

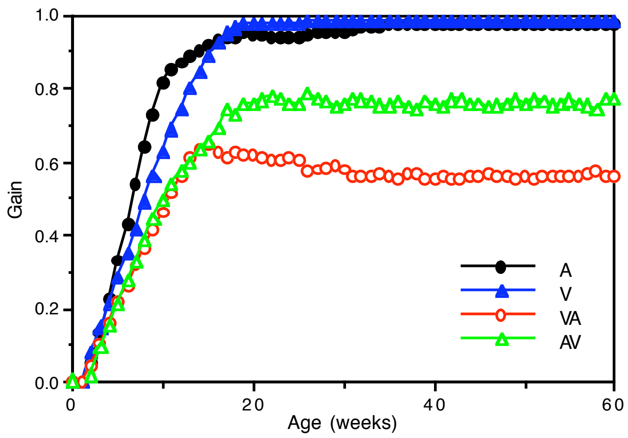

Figure 6: The model's development with faster A dead-zone maturation.

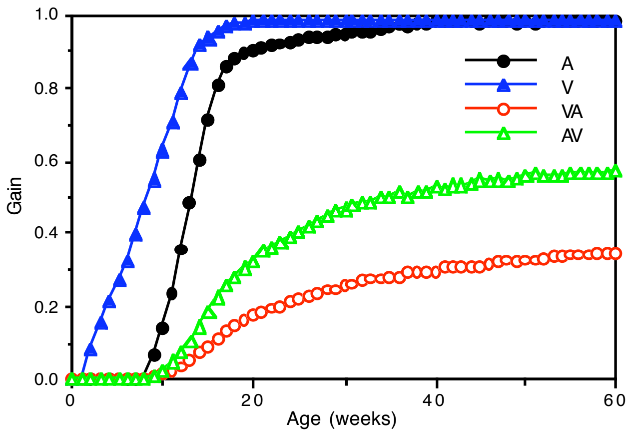

Figure 7: The model's development if we train from larger innate A weights.

accidents of evolutionary history, such as accommodation evolving before vergence.

To see how we might proceed, we need to consider in more detail the potentially conflicting factors that contribute to fitness from the learning and evolutionary points of view. First, the main factors relevant to lifetime learning are:

L1. We must perform the required tasks well (i.e. minimise blur and disparity).
L2. We must smooth/regularize the response (i.e. minimise overshoots).

These are built into our models as the error and regularization terms in the gradient descent cost function. Evolution provides other factors:

E1. Robustness of the learnt system will obviously be advantageous.
E2. Fast learning is advantageous since it minimizes periods of helplessness.
E3. Too much learning, or learning rates too large, can lead to instability.
E4. Too little learning can lead to an inability to adapt.

Our results above indicate that these will also affect what the model learns. It is thus clear that we need to take both learning and evolution into account when building realistic developmental models, since they will both affect what properties emerge within that system. Although useful properties will tend to emerge through learning alone, the best systems may well require evolutionary type processes as well.

The Baldwin Effect

Lifetime learning and evolution are not independent processes – they are tied together by the so-called Baldwin effect [1, 4]. This synergy comes about in two stages:

1. If a mutation (e.g. a change in learning rate or initial weight) that would otherwise be useless can be used by the learning process to allow the system to acquire better properties, then it will tend to proliferate in the population.
2. If the learning has an associated cost (e.g. requires energy and time), then its results will tend to be incorporated into the genotype and the learned behaviours will become innate. In other words, we have genetic assimilation.

However, if the system really does need to retain the ability to learn, for example to adapt to unknown or changing conditions, then we are likely to get only partial assimilation. We can still expect evolution to result in an efficient learning system that has minimal associated cost, but the required presence of a learning process will tend to diminish the genetic assimilation of the final learned behaviour. Moreover, if learning allows individuals with different genotypes to perform equally well, this will reduce the ability of natural selection to discriminate between them, and we will be left with a considerable range of individual differences.

For oculomotor control it seems that some aspects are innate but not tightly constrained (e.g. the initial weights and learning rates) and some are learnt (e.g. weights that change during development). We have a population that all perform well under normal conditions, but can be shown to have considerable underlying individual differences (e.g. a wide range of vergence to accommodation cross-link ratios). In general, a range of emergent properties will arise from a combination of developmental and evolutionary effects. It seems unlikely that we will be able to predict or understand the self-organising process without taking both these effects into account.

Technological Implications

We have seen that modelling complicated systems like oculomotor control is not as straightforward as one might have hoped. Evolution has clearly played an important role in constraining the human system and we can not ignore this in our models. This would seem to add a whole new level of complexity to our modelling endeavours, but from a technological point of view, our results may not be so troublesome. Whilst we should still not ignore them, one may simply regard evolution and the setting of our innate parameters as just another level of cost optimization, or the addition of a few more dimensions to the search space.

Since both evolution and lifetime learning require considerable computational resources, it would be wise to get the balance right. Following biological evolution may not necessarily be the best way to proceed. To produce realistic models of a human system, we should not stray too far from the real evolutionary process, but if we are simply aiming at producing efficient artificial systems, it may be sensible to attempt to improve upon the human system. Given that evolution is effectively a case of reinforcement learning, we may be able to improve on evolution by replacing it with a more supervised learning approach. Factors such as robustness, for example, could be added into the standard learning cost function. Related (almost Lamarckian) avenues for improvement might involve using what gets learnt to bias the evolutionary mutations. It is easy to see how this could assist in choosing more appropriate innate starting weights, but it is not so clear how it could help with other evolved factors such as learning rates. Also, given that evolution and learning often have different (but equally important) goals, this could actually be counter productive. A lot will depend on exactly how we rank the various fitness factors. It seems clear that more explicit simulation and experimentation will be required to determine a good general approach.

Conclusions

We have presented a promising approach for modelling the development of the accommodation and vergence control systems using neural networks. We have built the relevant maturational features into the developmental process, shown how our models can learn to behave in a manner similar to humans, seen how easy it is for the empirical individual differences to arise, and can now begin to relate the correlations between factors in our models with the situation in normal and abnormal human development.

Our main result from the technological implementational point of view is the demonstration that we should not ignore the influence of evolution and the Baldwin effect. To understand the emergence of properties from connectionist learning principles, we also have to understand the emergence of the learning principles and the innate learning starting points from the process of evolution. The minima obtained for the fitness function of lifetime learning will be constrained by the properties of the minima resulting from optimizing the fitness function of evolution. This lesson learnt from our oculomotor control models should be kept in mind when formulating other complex control systems that are expected to learn how to organise themselves.

References

1. Baldwin, J.M. (1896). A new factor in evolution. *The American Naturalist*, **30**, 441-451.
2. Bullinaria, J.A. & Riddell, P.M. (2000). Regularization in oculomotor control. In *Proceedings of the European Symposium on Artificial Neural Networks,* 207-212. Brussels: D-Facto.
3. Eadie, A.S. & Cardin, P.J. (1995). Evolution of control system models of ocular accommodation, vergence and their interaction. *Medical & Biological Engineering & Computing*, **33**, 517-524.
4. Hinton, G.E. & Nowlan, S.J. (1987). How learning can guide evolution. *Complex Systems*, **1**, 495-502.
5. Riddell, P.M. & Bullinaria, J.A. (1999). Incorporating developmental factors into models of accommodation and vergence. *Submitted for publication*.
6. Schor, C.M., Alexander, J., Cormack, L. & Stevenson, S. (1992). Negative feedback control model of proximal convergence and accommodation. *Ophthalmic and Physiological Optics*, **12**, 307-318.