Exploring the Baldwin Effect in Evolving Adaptable Control Systems

John A. Bullinaria
Department of Psychology, The University of Reading
Reading, RG6 6AL, UK

Abstract

A neural network model is presented which is an abstraction of many real world adaptable control systems that seems to be sufficiently complex to provide interesting results, yet simple enough that computer simulations of its evolution can be carried out in weeks rather than years. Some preliminary explorations of the interaction between learning and evolution in this system are described, together with some suggestions for future research in this area.

1. Introduction

The idea of evolution by survival of the fittest is now widely accepted. However, whilst it is clear that many human abilities have become innate as a result of evolution, others still need to be learned or modified during an individual's lifetime, so the nature-nurture debate rages on in many areas. It has been known for some time that there will be an interaction between learning and evolution (commonly called the Baldwin Effect [1, 2, 4]), but realistic explicit simulation is difficult due to the enormous computational resources required. Training a non-trivial fully dynamical neural network takes time, and consequently, training a changing population of such neural networks over many generations with reasonable procedures for procreation, mutation and survival is still barely feasible. In this paper I shall present a neural network model that is an abstraction of many real control systems (e.g. reaching, pointing, oculomotor control) which seems to be sufficiently complex to provide interesting results, yet simple enough that computer simulations of its evolution can be carried out in weeks rather than years. A range of preliminary explorations of the Baldwin effect in this system will be described, together with some suggestions for future research in this area.

2. The Simplified Control Model

The simplified generic control system that will form the basis of the study presented in this paper is shown in Figure 1. Its inputs are fairly accurate near cues and less accurate far cues. The network's response to these cues is generated by an initial approximate open loop signal based on the far cues, followed by a more accurate closed loop signal based on the near cues. These signals feed into integral and proportional controllers, the outputs of which are added to bias and tonic signals, and

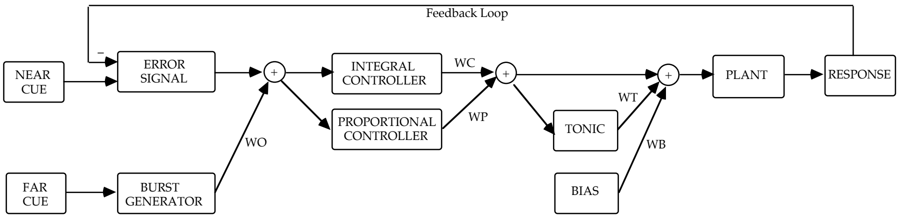

Figure 1: Simplified control model with five learnable parameters.

fed into the plant to produce the required response. The bias provides an appropriate resting state, and the tonic allows short time-scale adaptation of the resting state during periods of constant demand. The whole system can be regarded as a fully dynamical network of leaky integrator neurons. In the human accommodation (eye focussing) system, for example, we have blur and proximal cues being processed to generate appropriate signals for the ciliary muscles in the eye [3]. The model has five adjustable parameters (weights WO, WC, WP, WT, and bias WB) which are learnt by a simple gradient descent algorithm that minimizes a cost function consisting of response error and regularization (smoothing) components which will be readily available to the system. Corresponding to these learnable weights, each instantiation of the model has five fixed initial weights (iWO, iWC, iWP, iWT, iWB) and five fixed learning rates (eWO, eWC, eWP, eWT, eWB). The model also has various other parameters (neuron time constants, plant characteristics, feedback time delay, and so on) which we take to be the same for all instantiations. Such a system that has learnt/evolved a good set of parameters will produce appropriate damped responses to arbitrary discontinuous output requirements (steps) and smooth pursuit of arbitrary continuous output changes (ramps) as illustrated in Figure 2.

3. Evolving the Model

To simulate an evolutionary process for our model we take a whole population of individual instantiations of the model and allow them to learn, procreate and die in a manner approximating these processes in real (living) systems. The genotype of each individual will depend on the genotypes of its two parents, and contain the initial weights and learning rates. Then during its life the individual will learn from its environment how best to adjust its weights to perform most effectively. Each individual will eventually die, perhaps after producing a number of children.

In realistic situations, the ability of an individual to survive or reproduce will rely on a number of factors which can depend in a complicated manner on that individual's performance on a range of related tasks (food gathering, fighting, running, and so on). For the purposes of our simplified model, we shall consider it to be a sufficiently good approximation to assume a simple linear relation between our single task fitness function and the survival or procreation fitness. In fact, any monotonic relation will result in similar evolutionary trends.

We shall follow a more natural approach to procreation, mutation and survival than many evolutionary simulations [2]. Rather than training each member of the

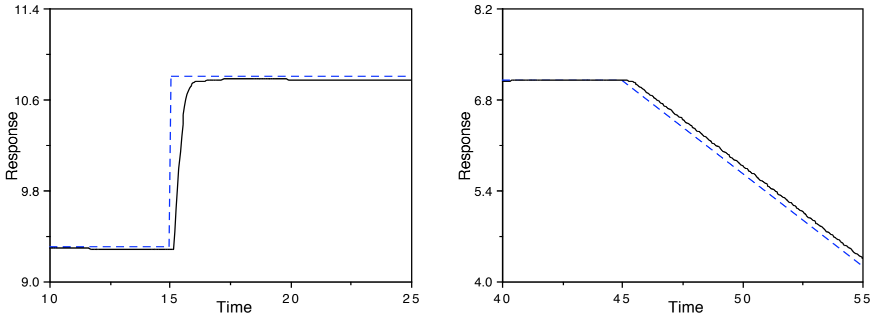

Figure 2: Typical system responses to arbitrary step and ramp stimuli.

whole population for a fixed time and picking the fittest to breed and form the next generation, our populations contain competing learning individuals of all ages, each with the potential for dying or procreation at each stage. During each simulated year, each individual learns from their own experience with a new randomly generated common environment (i.e. set of training/testing data) and has their fitness measured. A biased random subset of the least fit individuals, together with a flat random subset of the oldest individuals, then die. These are replaced by children, each having one parent chosen randomly from the fittest half of the population who randomly chooses their mate from the rest of whole population. Each child inherits characteristics from both parents such that each innate free parameter is chosen at random somewhere between the values of its parents, with sufficient noise (or mutation) that there is a reasonable possibility of the parameter falling outside the range spanned by the parents. Ultimately, our simulations might benefit from more realistic encodings of the parameters, concepts such as recessive and dominant genes, learning and procreation costs, different inheritance and mutation details, different survival and procreation criteria, more restrictive mate selection regimes, offspring protection, different learning algorithms and fitness functions, and so on, but for the purposes of this paper, our simplified approach seems adequate.

4. The Baldwin Effect

Since the Lamarckian idea of inheriting acquired characteristics is now known not to happen in real biological systems, it is commonly assumed that lifetime learning and evolution are independent processes. However, they are actually tied together by the so-called Baldwin effect [1]. This synergy comes about in two stages:

1. If a mutation (e.g. a change in learning rate or initial weight) 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 time or energy), then its results will tend to be incorporated into the genotype and the learned behaviours will become innate.

Put another way, evolution first creates a population that can learn good properties, and then removes the need for learning whenever it is possible to do so. In this way we have genetic assimilation of the learnt behaviour. Although this "new factor in evolution" dates back to the nineteenth century, it was not until the work of Hinton and Nowlan [3] that it became widely known to the connectionist community [2]. They demonstrated explicitly, for a particular simplified system, that the ability to learn a behaviour was able to improve the rate of evolution of that behaviour, without the learnt characteristics being passed between generations in the genotype. Their task, however, involved a very localized fitness function that was particularly difficult for evolution to search on its own. They noted that "for biologists who believe that evolutionary search space contains nice hills (even without the restructuring caused by adaptive processes) the Baldwin effect is of little interest". This, of course, depends on what interests you. It might be of little interest if you just want to demonstrate that learning can speed evolution, but there remains much of interest if you wish to explore genetic assimilation, the nature-nurture debate, individual differences, and such like.

In this paper I wish to begin to explore the Baldwin effect in its broadest sense. As with any modelling endeavor, there are numerous system design choices that can potentially have crucial consequences for the results, and in any complete study these will have to be explored systematically. Many of these details will, as in real systems, ultimately be determined as a result of the evolutionary process itself. For the purposes of this preliminary study, however, we shall concentrate on exploring the interaction of learning and evolution for the particular hand-crafted system described above. It is clearly important for any realistic developmental model to know how much of the system's behaviour is innate and how much has to be learnt by interaction with its environment. Assuming that learning does have a significant cost, a fixed system in a stable environment can be expected to evolve so that its optimal behaviour is completely innate. However, if the system really does need to retain the ability to learn, for example to adapt to unknown or changing conditions, or to compensate for aspects of its own natural maturation, then we may only get partial assimilation. We can still expect evolution to result in an efficient learning system that has minimal associated cost, but the appropriate innate properties may no longer correspond to a 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. The remainder of this paper explores these effects by explicit simulations of the model presented above.

5. Simulation Results

Even after we have made the basic system design decisions discussed above, there are still a number of parameter choices to make. Clearly much further work will be required to justify particular choices for realistic models of human evolution. For concreteness here, all the fixed network parameters (neuron time constants, plant characteristics, environment characteristics, and so on) were chosen to match those of the human oculomotor control system [3]. The main evolutionary parameter values in this study were largely forced on us by limited computational resources. Choosing a fixed population size of 200 was a trade-off between maintaining genetic

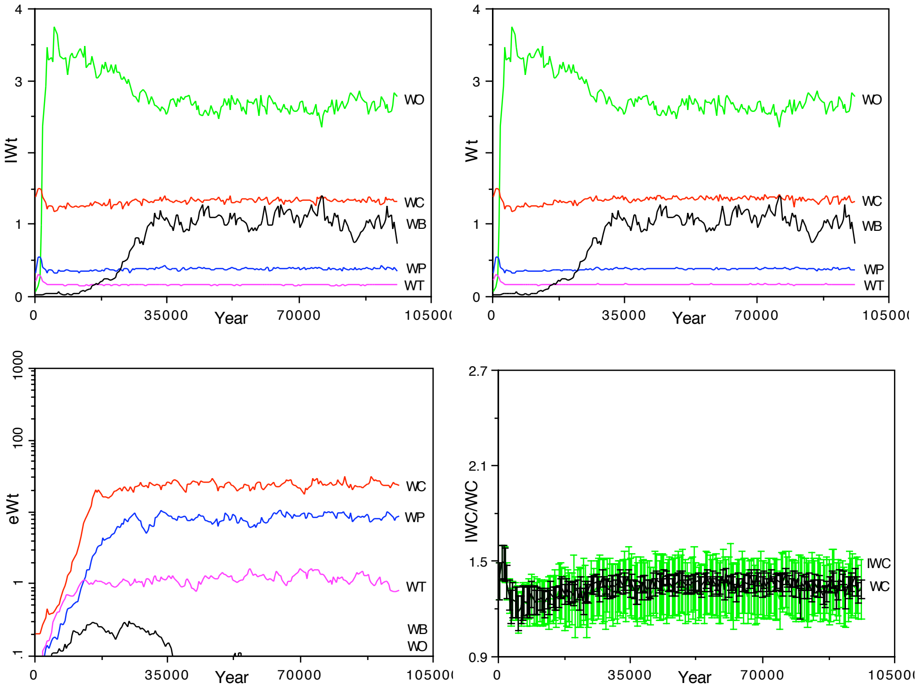

Figure 3: Evolution of weights and learning rates from zero.

diversity and running the simulations reasonably quickly. The death rates were set to produce reasonable age distributions. This meant about 5 deaths per year due to competition, and another 5 individuals over the age of 30 dying each year due to old age. The fixed population size meant each individual then produced on average two children. The mutation parameters were chosen to speed the evolution as much as possible without introducing too much noise into the process. The amount of training data presented to each individual per year also had to be restricted to minimise the computation times. All these compromises led to coarser simulations than one would like, but otherwise the simulations would still be running.

The natural starting point is to begin with a population consisting of individuals that have all their weights and learning rates zero, and see if they can evolve into a population that can perform well at their given task. Figure 3 shows how such a population's mean initial weights (iWt) and learning rates (eWt) evolve in this case. It also shows the mean actual weights (Wt) across all the individuals, which appear little different to the mean initial weights. The difference is clear, however, in the fourth graph which shows the means and standard deviations of the iWC and WC weights for mature individuals (i.e. those at least 10 years old). We see that the learning process has the effect of tightening the weight distribution from that generated by the inaccurate procreation process towards the optimal values.

As noted above, there is a crucial distinction between systems that really do need to learn, for example to adapt to changing conditions during their lifetime, and those which could perform perfectly well without learning if given appropriate innate

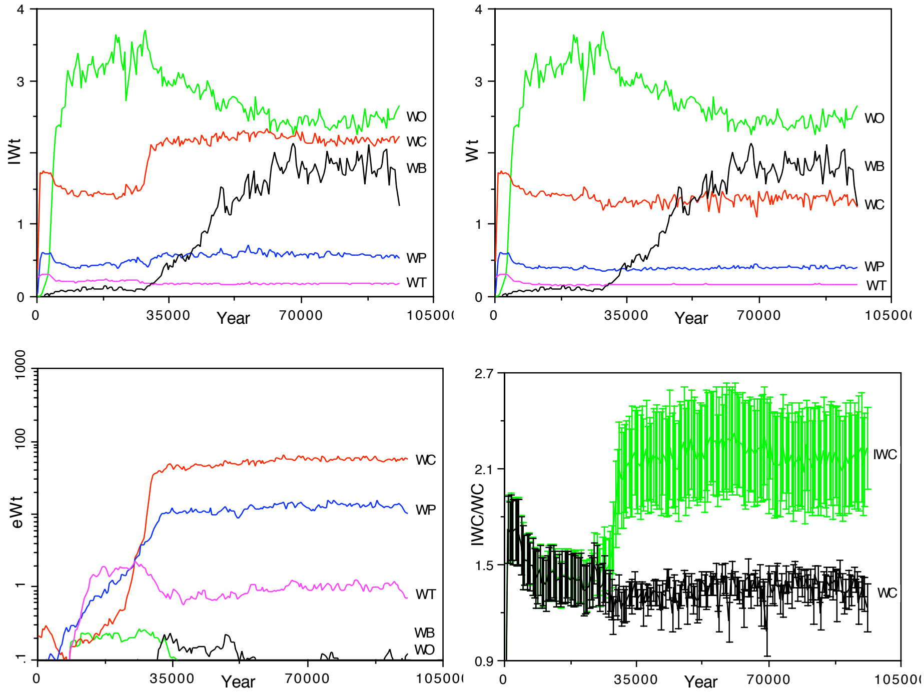

Figure 4: Evolution when the system has to cope with maturation.

values for all the necessary parameters. We have seen that there is a certain necessity for our models to adapt due to the variability that is built into the procreation process, but we also need to investigate the effect of needing to adapt to changes that take place during an individuals lifetime, in which case good adult weights are not necessarily appropriate initial weights. For example, in oculomotor control, the relationship between the eye rotations and the object's distance will need to vary as the inter pupil distance grows during childhood [3]. To incorporate such an effect into the simulations, a simple maturational scale factor was applied to the output, varying linearly from 0.5 to 1.0 between the ages of 0 and 10 years. The evolution of this new system is shown in Figure 4, and it does indeed differ significantly from that seen in Figure 3. In particular, we see larger evolved learning rates and, once a significant amount of learning has evolved, there is a clear difference between the mean initial and mature weights *WC* and *WP*. Comparing the distributions of the initial and mature *WC* weights we see that, once learning has evolved, there is a clear separation between the initial weight distribution and the tighter mature weight distribution (cf. the corresponding graph of Figure 3).

To understand this system better, Figure 5 presents two further graphs which show important general characteristics of our models. On the left we see how the mean, minimum and maximum individual fitnesses evolve, with clear improvements corresponding to the onset of learning seen in Figure 4. Notice that the maximum fitness often decreases. This is because even the fittest individuals eventually die, and they will not necessarily produce children as good as themselves, and even when

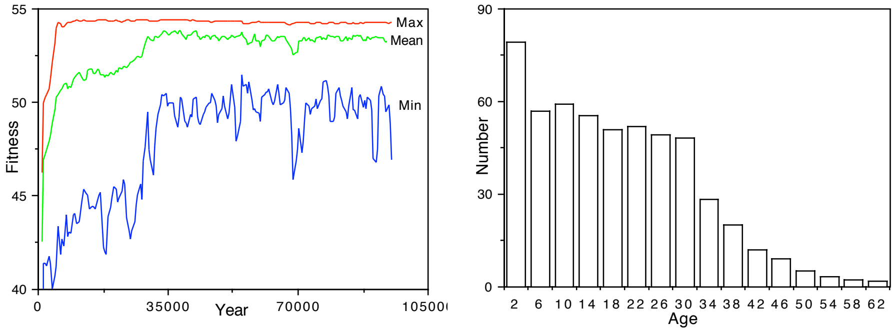

Figure 5: Evolution of fitness and evolved age distribution.

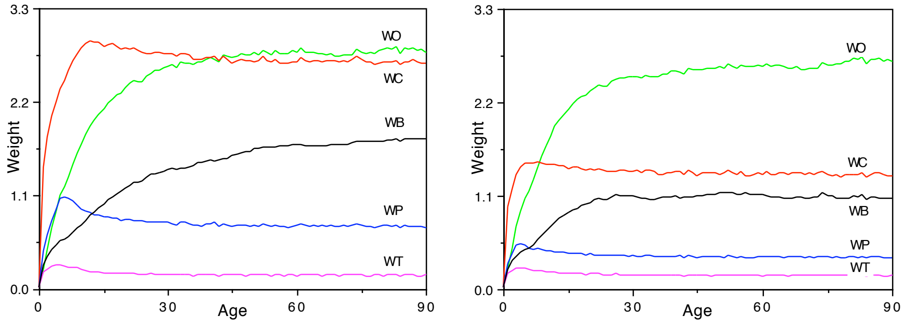

Figure 6: Learning weights appropriate for a newborn and an adult.

they do, those children are not necessarily going to experience learning environments as good as their parents'. As a result, the evolution is slower, but it is the price we pay for realism. This is one area where we might want to improve upon the human system when building artificial systems. On the right of Figure 5 we have the age distribution of our evolved population. We see an initial sharp drop corresponding to a high infant mortality rate, followed by a slower fall until the age of 30 after which individuals start to die of old age.

To assess the initial weights and learning rates that our systems evolve, we really need to know what the optimal newborn and mature weights would be to produce a good average performance across the whole (infinite) distribution of training data. Fortunately, these can easily be determined sufficiently accurately using the model's own learning algorithm whilst keeping the maturational scale factor fixed at either 0.5 or 1.0. Figure 6 shows how such learning alone leads to appropriate weights for a newborn and an adult. We end up with the same final weights using a range of initial weights and learning rates, indicating that these represent the true global minima which we might expect our evolving populations to reach. Note that, because of the linearity of our model and the way we have chosen to parameterize it, the maturational output scale factor can be compensated by an inverse scaling of WB, WC and WP, whilst leaving WO and WT alone. This is seen clearly in the optimal weights from Figure 6, which are in broad agreement with the mean evolved initial weights in Figures 3 and 4. The expected equality of the evolved mature

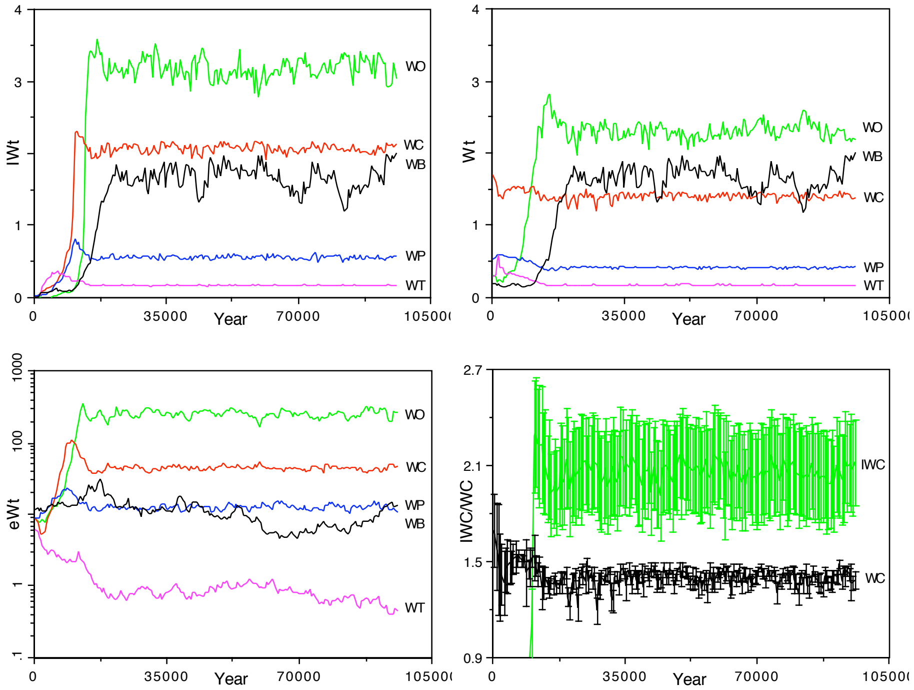

Figure 7: Evolution starting with large learning rates.

weights between Figures 3 and 4 is also found, apart from the bias *WB*, which we would expect to have evolved a significant learning rate so that it could compensate appropriately for the maturation. The reason for this discrepancy is something that will be investigated further later.

The next variation to consider, now that the basic model has been simulated and examined, is the specification of the initial population. Starting all the initial weights and learning rates from zero makes sense, but so does starting them off with a large range of random values. Figure 7 shows the evolution that results when all the learning rates start off randomly distributed in the range 0.0 to 20.0. There are several interesting differences from the corresponding graphs in Figure 4. First, having large learning rates right from the start *delays* the evolution of the innate weights, because individuals are able to *learn* the weights necessary to perform well. In due course however, genetic assimilation does occur, and the *WC* learning rate can be seen to fall dramatically as *iWC* increases. Eventually, the system settles down with mature weights near the optimal values observed before. For some reason, however, the *WO* innate weights and learning rates become unexpectedly large, and the *WB* weight again fails to adjust in line with the maturation.

This leads us to consider the case in which both the initial weights and the learning rates are randomly distributed in the range 0.0 to 20.0. Figure 8 shows how this affects the evolution. Naturally the initial stages are different from those seen in Figures 4 and 7, but the system soon settles down into the same pattern observed for the case of zero starting weights in Figure 7.

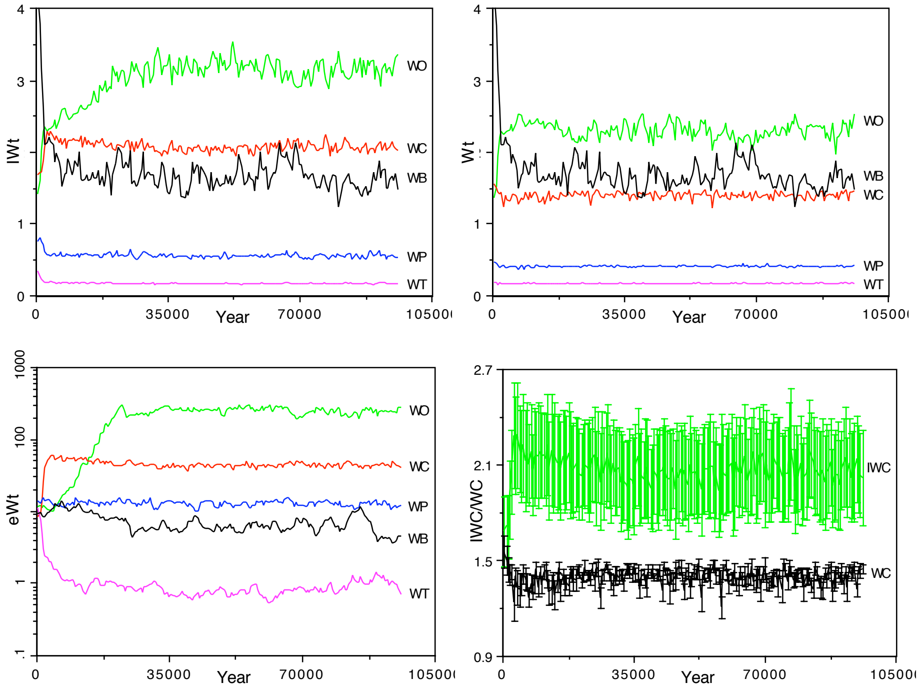

Figure 8: Evolution starting with large initial weights and learning rates.

Finally, to ensure a fair assessment of how learning affects evolution, Figure 9 shows how the system evolves when there is no ability to learn. The unused learning rates are plotted to show the kind of random walks that result when there is no evolutionary pressure. The initial and mature weights in this case are identical and constitute a compromise that allows the individual to survive the immature stage whilst allowing reasonable performance when mature.

6. Discussion

To a large extent our simulations have behaved how we would hope and expect them to, with appropriate evolved values for the innate weights and learning rates which lead to appropriate mature learned behaviour. However, we have noticed that the bias *WB* consistently fails to take on optimal values, and that when we start with large learning rates, the innate values of the weight *WO* appear larger than optimal. Clearly, some further investigation is required here.

One thing that is particularly evident from the simulations is that the weights are not all behaving in the same way – some are evolving faster than others, some are evolving larger learning rates, and so on. A convenient way to explore this in more detail is to test the sensitivity of the cost function to changes in each weight while all the other weights are kept fixed. This will give an indication of the forces acting on each weight during learning and evolution. Figure 10 presents the relevant results corresponding to the simulation shown in Figure 7. On the left we have the mean

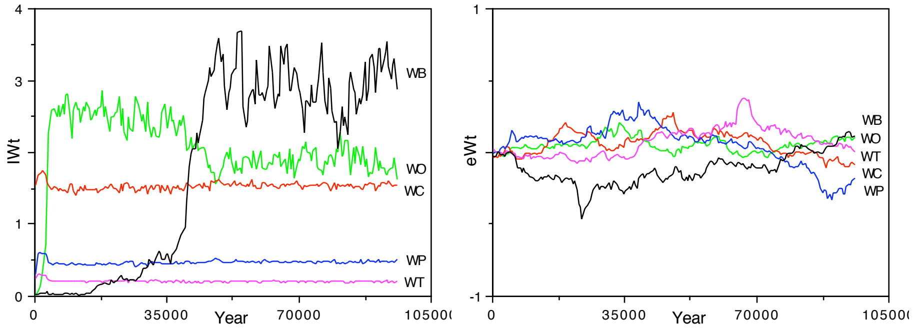

Figure 9: Evolution when there is no ability to learn.

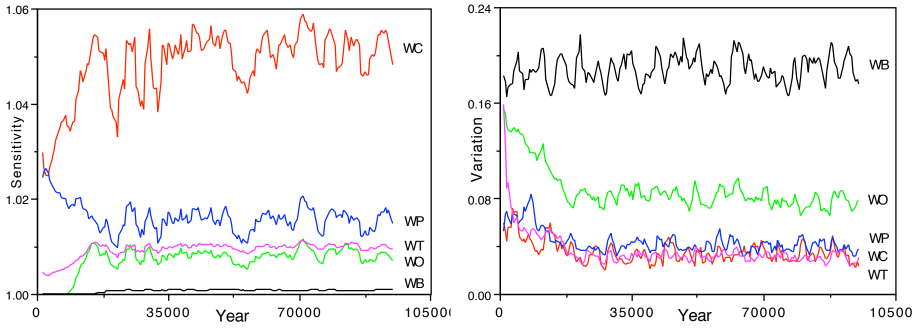

Figure 10: Weight sensitivities and individual variations.

proportional changes in the cost that is caused by increasing or decreasing each weight by 20% at each stage of the system's evolution. From this it is now clear why the bias *WB* is so slow to evolve and take on its optimal values during maturation. The cost function is so insensitive to it that the drive for it to change is being lost in the noise. It is quite possible that simulating a longer evolutionary period, or performing a less noisy simulation (e.g. by using a larger population size or more training data per simulated year), will allow the drive to manifest itself, but for the present we must be satisfied with understanding what is happening.

On the right of Figure 10 is shown the variability (i.e. standard deviation over mean) of each weight across the population during the system's evolution. We see that the bias *WB* stands out as having the particularly large range of individual differences we would expect given its small influence on the fitness. Here it is worth remembering that an important relevant feature of the models is that the weights within each individual are not independent. The optimal value for each weight will depend on the values of the other weights. There will exist an overall optimal set of weights, but if for some reason one weight does not take on its optimal value, then the others can partially compensate. For example, in each of Figures 4, 7, 8 and 9, we can see that before the innate weight *WO* evolves, the values of *WC*, *WP*, *WT* are larger. Similarly, low values of *WB* seem to coincide with higher values of *WO*. This anti-correlation of *WO* with the poor bias *WB* signals may well explain the unexpected behaviour of *WO* in some situations.

A realistic feature of our models is that each individual not only needs to perform well on average, but must also be able to cope if an unusual/extreme environmental condition arises (such as requiring a particularly large response change, or a particularly rapid sequence of response changes). To do this may require a set-up that is less than optimal for normal conditions. So, whilst it is reassuring to see that our evolved individuals are broadly in line with the optimal performance parameters, this is not something we can rely on happening in general. The extent to which this is responsible for the sub-optimal parameters that have evolved in our models is clearly worthy of further investigation in the future.

Having understood the factors underlying what has been happening in our simulations, there are several further issues worth commenting on. First, for the models, survival is determined by competition among the individuals, rather than by competition of the individuals against a hostile environment. Moreover, children are always produced to replace the dead, however young and unfit the parents might be. Consequently, whilst a non-adaptive population appears to evolve faster than their adaptive counterparts (e.g. Figure 4 compared with Figure 7), the poor fitness of individuals during the early stages of evolution (seen in Figure 5) may, under more realistic circumstances, result in the population not surviving long enough for the evolution to happen at all. It is in this sense that learning will assist evolution. The traditional Baldwin Effect [1, 4] corresponds to the comparison of the evolution in Figure 7 not with Figure 4, but with that of a system which has to wait an enormous time for a rare mutation to create a non-adaptable individual that is able to survive long enough in its environment to procreate.

Finally, note that the cost of learning is not encoded explicitly into any of our fitness functions, but it is implicit in the sense that if any unnecessary learning is required, then individuals are at a disadvantage until the learning is completed. The population can compensate for this to a certain degree by having individuals learn quickly, but if the learning rates are constant throughout each individual's lifetime, this may lead to instability. A sensible strategy, that occurs in real systems and will be explored further for our models elsewhere, might be to evolve initially high plasticities that decrease with age. Alternatively, a parent or population may evolve a propensity to protect its offspring until they have acquired the ability to fend for themselves, as many real species do. This is also worthy of a study in its own right, but for the present simulations, we simply protect the offspring until they have completed their first year of learning, and expect them to compete with adults after that. The net effect will be that, where possible, evolution will result in the need to learn being replaced by innate behaviour. For now, we leave open the question of whether a more explicit learning cost might be more beneficial or realistic.

7. Conclusions

We have seen how it is possible to simulate the interaction of learning and evolution in a class of simplified neural network control models. The processes whereby appropriate innate connection weights and learning rates evolve can be understood, and are broadly in line with what we might expect. However, it is also clear how the need to cope in unusual/extreme environmental conditions may result in individuals evolving in such a way that their performance is sub-optimal under normal conditions. We have also seen how some weak effects can easily get lost in the noise and result in the evolution of less than optimal solutions, which in turn may result in various compensatory effects generating unexpected properties.

We already knew that in order to understand real developmental processes, it is important to understand how those developmental systems have been constrained by their evolution [3]. The simulations presented in this paper now indicate that to understand the evolutionary process we also need to take into account the history of the evolutionary environment, as there are dependencies on the evolutionary initial conditions and we cannot rely on an optimal developmental system from evolving within a given time-scale. Moreover, fitness insensitivity to particular parameters can lead to considerable ranges of individual differences that can be difficult to predict without explicit simulation. In the case of oculomotor control, for example, individuals exhibit a wide range of cross-link strengths between accommodation and vergence with little effect on their normal performance [3].

Having seen how fundamental questions about the nature-nurture debate and individual differences may be answered, the next stage of this work will be to check the extent to which our results change as we vary the details of the models and the simplifications we have made to the real evolutionary process. We also need to test the approach against some real systems. This will be complicated by the fact that, in practice, real control systems will evolve alongside their plant, rather than independently with a fully operational plant. It is also quite likely that the initial population will arise as mutations of some other existing system (possibly adaptable, possibly not) and this will surely affect what evolves. There is clearly some way to go to achieve reliable simulations of real human or animal evolution. Alternatively, we could consider the approach to be an appropriate technique for developing efficient artificial control systems for real world problems. In which case, we have some way to go in another direction.

Acknowledgements

This work was begun many years ago while at Edinburgh University supported by the MRC. It is currently supported by the EPSRC as part of a project with Patricia Riddell on models of the development of accommodation and vergence.

References

1. Baldwin, J.M. (1896). A New Factor in Evolution. *The American Naturalist*, *30*, 441-451.
2. Belew, R.K. & Mitchell, M. (Eds) (1996). *Adaptive Individuals in Evolving Populations*. Reading, MA: Addison-Wesley.
3. Bullinaria, J.A. & Riddell, P.M. (2000). Learning and Evolution of Control Systems. *Neural Network World, 10*, 535-544.
4. Hinton, G.E. & Nowlan, S.J. (1987). How Learning Can Guide Evolution. *Complex Systems*, *1*, 495-502.