# The Evolution of Variable Learning Rates

**John A. Bullinaria**
School of Computer Science
The University of Birmingham
Birmingham, B15 2TT, UK
*j.bullinaria@physics.org*

## Abstract

Neural plasticity in humans is well known to be age dependent, with 'critical periods' for the learning of many tasks. It is reasonable to hypothesise that this has some intrinsic advantage over constant plasticity, and that it has arisen as the result of evolution by natural selection. If this is true, then it may also prove useful for building more efficient artificial systems that are required to learn how to perform appropriately. In this paper I explore these ideas with a series of explicit evolutionary simulations of some simplified control systems.

## 1 INTRODUCTION

Evolutionary algorithms have shown much promise for generating artificial neural networks with performance superior to those formulated directly by human researchers. Factors such as network architecture, learning rules and connection weights have all been successfully optimised by evolution (e.g., Yao, 1999). A similar approach can equally well be applied to optimising the adjustable parameters and learning rates in other systems that learn, such as traditional adaptable controllers (e.g., Levine, 1996; Bullinaria, 2001). In this paper I take this work one stage further by considering how an evolutionary approach might lead to more efficient systems by allowing the emergence of non-constant learning rates.

It is well known that human neural plasticity varies considerably with age, and that there are "critical periods" during which learning must take place if the given task is to be mastered successfully (Julesz & Kovacs, 1995). The idea of variable neural plasticity is also quite common in the field of artificial neural networks where modellers have found it beneficial to vary their network learning rates during the course of training (Jacobs, 1988). For example, near the end of training it may be useful to decrease the learning rates to minimise the weight variations seen after each sample in online training, or to increase them to speed the saturation of sigmoids as the errors become small. Alternatively, if the performance of a task depends crucially on some lower level of processing, it may be sensible to delay the learning of that task until the lower level processes have fully developed. It is not clear to what extent factors such as these have been responsible for the evolution of the patterns of plasticity found in humans, or if it has been more a matter of minimizing the physical overheads of the plasticity. In this paper I shall present a series of explicit simulations of the evolution of some simple adaptable control systems. The evolutionary processes will result in efficient patterns of variable learning rates for these artificial systems that can then be used to develop better learning strategies for real world applications, and perhaps also provide some constraints on our explanations of the critical learning periods found in humans. The overall aim will be to see which learning strategies evolve naturally, and to explore how different strategies evolve under different circumstances.

## 2 THE CONTROL MODEL

The control system that will form the basis of the current investigation is shown in Figure 1. It is actually a simplified version of the part of the oculomotor control system that focuses and rotates the human eye (Schor et al., 1992), though similar systems can be applied quite generally (Levine, 1996). The input is a sequence of target responses and a feedback loop allows the determination of an error signal. This signal then feeds into standard simple integral and proportional controllers, the outputs of which are added to bias and

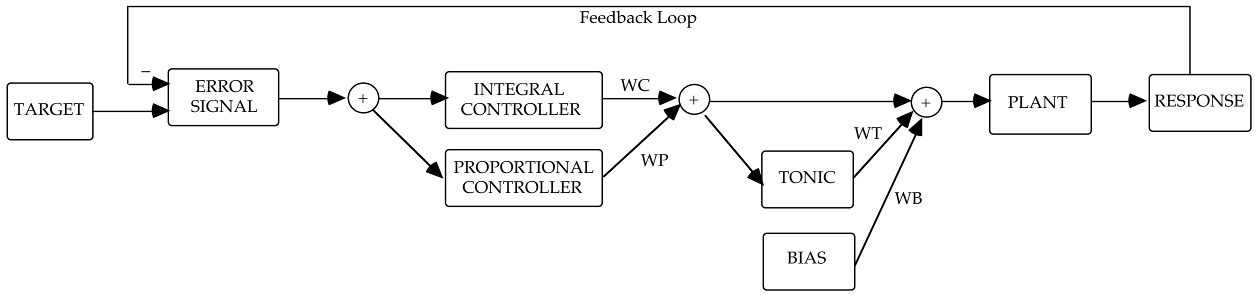

Figure 1: A simplified control model with four learnable parameters: WC, WP, WT, WB.

tonic signals, and fed into the plant to produce the response. The unit bias provides an appropriate resting state, and the leaky integrator tonic allows short timescale adaptation of the resting state during periods of constant demand. In the human eye focusing system, for example, we would have blur being processed to generate signals for the ciliary muscles in the eye appropriate for the distance of the visual target. The system can equally well be regarded as a traditional control system (Levine, 1996), or as a fully dynamical network of leaky integrator neurons.

Simulating the evolution will involve working with a large number of copies of this model, each with four adjustable parameters/connection weights $\mathbf{W}(t) = \{WC(t), WP(t), WT(t), WB(t)\}$ where $t$ is the time/age of the individual model measured in simulated years. These are learned by a simple on-line gradient descent algorithm that minimizes a cost function consisting of response error and regularization (smoothing) components which would be readily available to the system (Bullinaria & Riddell, 2001). Corresponding to the learnable weights, then, each instantiation of the model will have four variable learning rates/plasticities $\mathbf{P}(t) = \{PC(t), PP(t), PT(t), PB(t)\}$. The model will also have various other parameters (time constants, plant characteristics, feedback time delay, and so on) which we take to be the same for all instantiations, with values appropriate for human oculomotor control (Schor et al., 1992). Such a system that has evolved/learned a good set of weights will produce appropriate damped responses to arbitrary discontinuous output requirements such as steps, and smooth pursuit of arbitrary continuous output changes such as ramps (Bullinaria & Riddell, 2001).

For the purposes of this paper, I shall assume that all the learning rates in a given model vary with age in the same manner, and that this variation depends only on the genotype (innate parameters) of the individual, and not on the environment that the individual finds itself in. Naturally, it will be important to relax this condition in the future, but this means that we can write $\mathbf{P}(t) = s(t).\mathbf{P}(0)$, where $\mathbf{P}(0)$ are the four initial learning rates, and $s(t)$ is a simple age dependent scaling factor. Clearly, if there is no plasticity variation, $s(t) = 1$ for all $t$. A convenient parameterization is simply to take $s(t)$ to be piecewise linear with parameters $\mathbf{S} = \{s(t) : t = 1, \ldots, N\}$. The part of the model's genotype that varies between individuals thus represents the $8 + N$ parameters $\{\mathbf{W}(0), \mathbf{P}(0), \mathbf{S}\}$.

## 3 EVOLVING THE MODEL

Simulating an evolutionary process for our model involves taking a whole population of individual instantiations and allowing them to learn, procreate and die in a manner approximating these processes in real (living) systems. The genotype of each new individual will depend only on the genotypes of its two parents and random mutation. Then during their life each individual will learn from their environment how best to adjust their weights to perform most effectively. Eventually, perhaps after producing a number of children, each individual dies. Obviously, in nature, the ability of an individual to survive or reproduce will depend on a number of factors that are related in a complicated manner to that individual's performance on a range of related and unrelated tasks (food gathering, fighting, running, and so on). For the purposes of our simplified model, however, I 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 should result in similar evolutionary trends, but it is easy to lose weak effects in the noise of the rather coarse simulations forced upon us by limited computational resources.

Given that, initially at least, we are aiming to replicate an effect that arises in human evolution, it seems appropriate here to follow a more natural approach to procreation, mutation and survival than has been used in many evolutionary simulations in the past (e.g. in Belew & Mitchell, 1996). Rather than training each member of the 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, every individual learns from their own experience with a new randomly generated common environment (i.e. set of training/testing data) and has its fitness measured. Random pairs of individuals are then forced to compete, with the least fit dying (i.e. being removed from the population). Additionally, a random subset of the oldest individuals die of old age. The dead are replaced by children, each having one parent who is the fittest of a randomly chosen pair from the remaining 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 SIMULATION RESULTS

A previous study (Bullinaria, 2001), employing a slightly more complex control system and a slightly simpler evolutionary regime, has already explored the Baldwin Effect, i.e. the interaction of learning and evolution (Baldwin, 1896; Belew & Mitchell, 1996), in models of the type considered here. This demonstrated explicitly how genetic assimilation of learned behaviour (i.e. learned parameter values) will occur automatically, without Lamarckian inheritance, to reduce the inherent costs of learning (e.g. periods of poor performance). However, even when a good set of innate parameters have evolved, a control system will still benefit from being plastic since that will allow it to fine tune its performance after a noisy procreation process and/or being born into an unpredictable environment. Many biological systems will also need plasticity to compensate for the changes (e.g. growing size) that naturally take place during their own maturation period. For the current study, such a maturation process was simulated by a simple output scale factor that varies linearly from 0.5 to 1.0 over the first ten years of life for each individual. (It turns out that the precise details of this variation are not crucial.) In humans this maturation might correspond to changes in inter-pupilliary distance for the eye rotation system, or changes in arm length for reaching or pointing. The important consequence is that the appropriate innate/newborn weights will not be the same as the adult values. The pattern of plasticities that evolve will allow the system to learn most efficiently how to optimize its weights throughout its life.

Unfortunately, limited computational resources allowed only a rather coarse simulation of the evolutionary process, but for an initial study it proved sufficient to have a fixed population size of only 100, with around 10 deaths per year due to competition, and around 4 individuals over 30 years old dying each year due to old age. (Such a system coded in C typically simulated around 20,000 years per CPU day on an average UNIX workstation.) The procreation and mutation parameters were chosen to speed the evolution as much as possible without introducing too much noise into the process. These evolutionary details were kept constant across all the simulations I shall now present.

Figure 2 shows the simulation results for a typical run of the basic system described above. First we see that the population means of the initial weights $\mathbf{W}(0)$ and learning rates $\mathbf{P}(0)$ quickly evolve to take on appropriate values. (Note the large variation between the learning rates that emerge for the different weights.) These lead to good values for the weights throughout the individuals' life. All the weights will need an initial fine tuning to remove the noise in the procreation process, then some weights ($WC$ and $WP$) need to adjust during the maturation period, while others ($WT$ and $WB$) need little further change. The plots of $WC(t)$ and $WT(t)$ for a typical evolved population show this quite clearly. The plots of the mean weights $\mathbf{W}(t)$ for the whole population show how they differ in magnitude and noise from the initial weights $\mathbf{W}(0)$. Finally, we see how the plasticity scale factor $s(t)$ varies with age $t$. In particular, we see that the plasticity falls drastically between birth and the end of the maturation period, thus confirming that critical periods for learning will arise as a natural consequence of evolution.

The results from the basic system naturally lead to the question of what happens if an individual needs to adapt or learn later in life, after the standard learning period is over. There is a traditional saying that

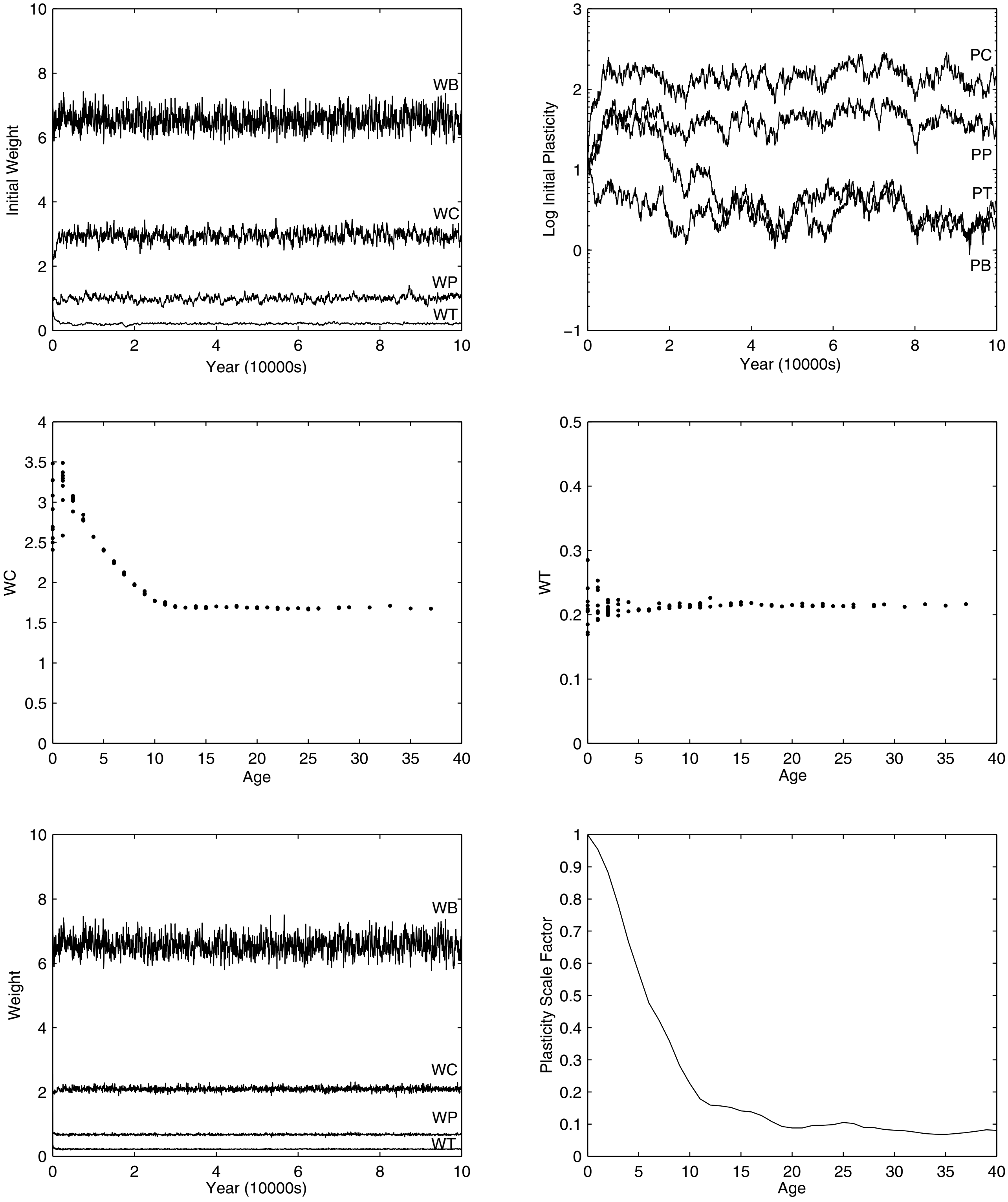

Figure 2: Evolution and learning in a typical simulation of the basic system. Individuals in the evolved population have plasticities that fall rapidly between birth and the end of their maturation period.

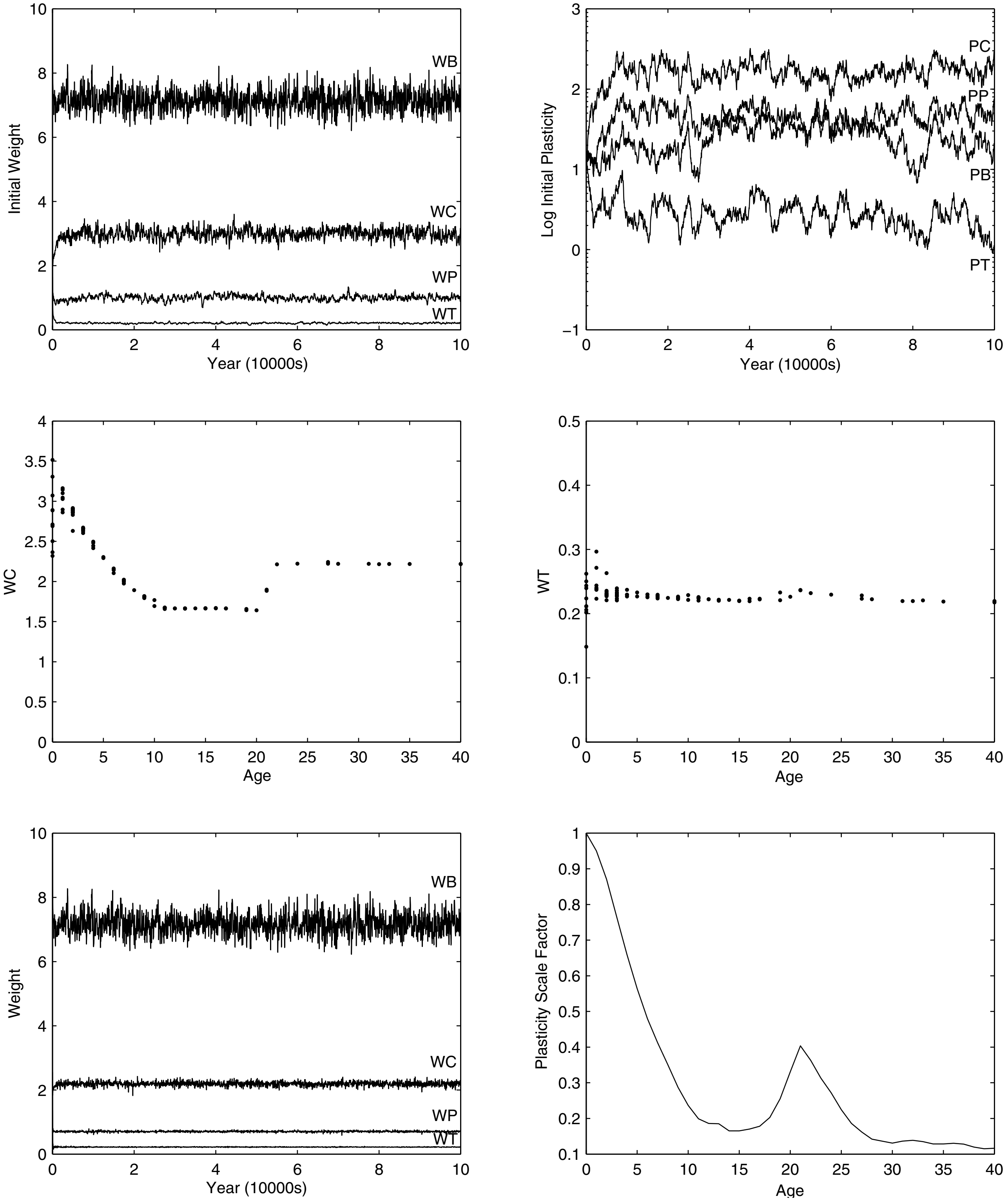

Figure 3: Evolution and learning in a typical simulation when late life adaptation is required. Individuals in the evolved population have plasticities appropriate for the learning or adaptation that is forced upon them.

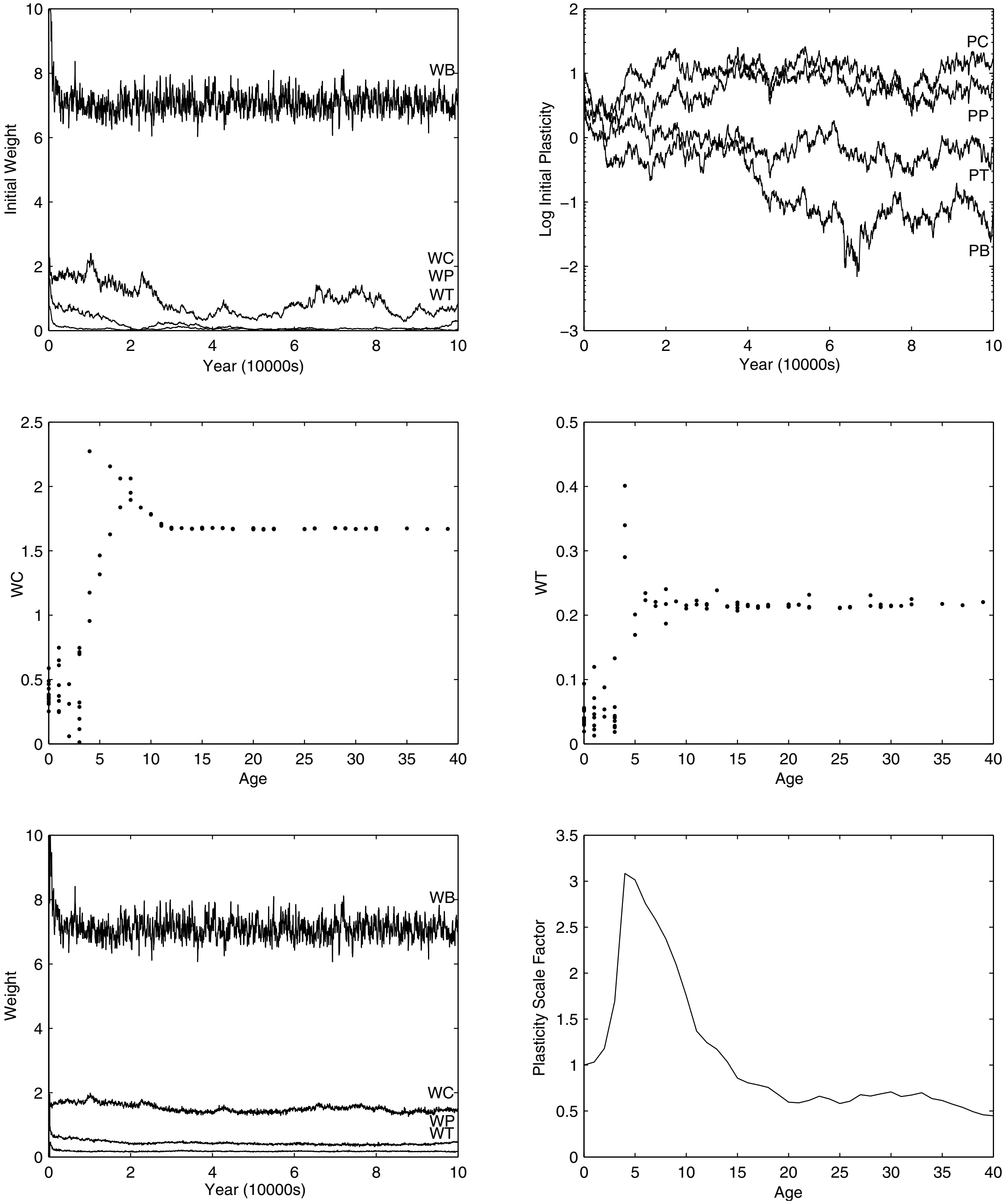

Figure 4: Evolution and learning in a typical simulation when there is a dependency on the development of lower level sub-systems. Individuals in the evolved population have a critical period for learning.

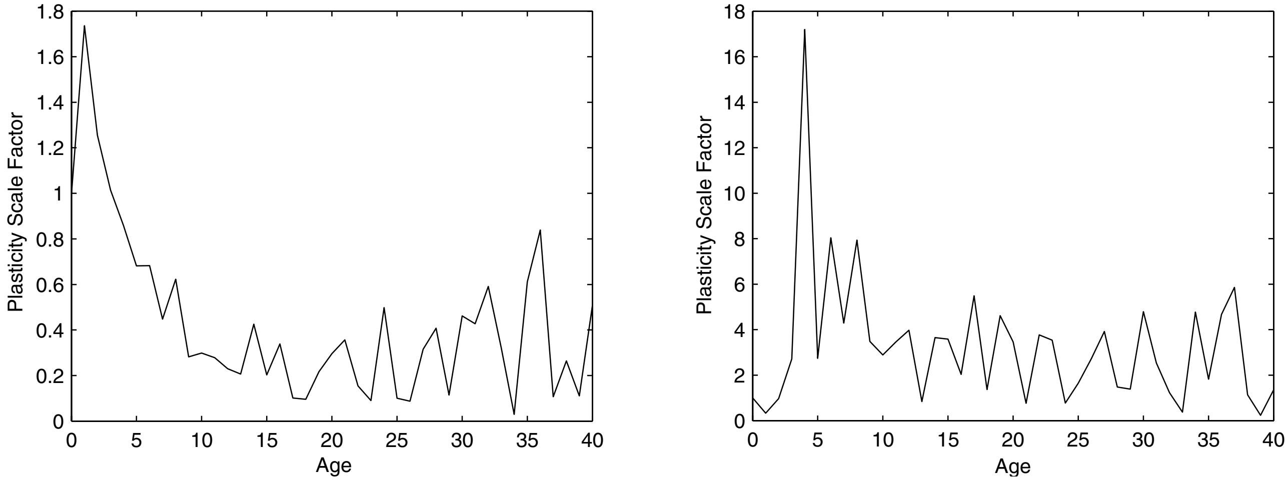

Figure 5: Typical plasticity scale factors arising from a different implementation to that used for Figures 2 and 4. The basic patterns are the same, but considerably noisier.

"old dogs cannot learn new tricks", but it seems unlikely that evolution would allow the plasticities to decay away to small values in situations where late life adaptation is regularly required. To introduce such a requirement, the basic model was modified so that there was a sudden step in the output scale factor from 1.0 to 0.75 at the age of 20. (Again it turns out that the precise details of this variation are not crucial.) There is no need to specify whether this variation corresponds to an internal factor (e.g. compensation for system damage or deterioration) or an external factor (e.g. adaptation to changes in the operating environment), as they will have the same effect. Obviously, the need for real late life adaptation will rarely be so predictable, but the consequences for our model will be similar, and the simplification makes it easier to interpret the results.

Figure 3 shows how this changes the simulation results from those of the basic model in Figure 2. The most obvious difference is in the plot of $WC(t)$ where we see the required step change at age 20 has been learned successfully. The plot of $s(t)$ shows the initial fall as before, but then a sharp rise to give the required increased plasticity at the age of 20. This gives us confidence that our evolutionary simulations really are picking up the requirement for plasticity, and not some confounding factor.

A final situation to consider, that regularly arises in human development, is when one level of processing relies on signals from another system. If the sub-system supplying those signals is not fully developed, it might be sensible to wait until it is before beginning to learn how to use the signals. For example, the adult eye rotation (vergence) system uses an image disparity signal, and humans have to wait until 12-16 weeks of age before this signal relatively suddenly becomes available. To simulate such an effect in our basic model, the error signal was replaced by low level noise for each individual until they reached three years of age.

Figure 4 shows how this affects the standard results of Figure 2. The changes here are rather clear. First, the initial/innate weights $WC$, $WP$ and $WT$ all drop to very low values, leaving the system with an appropriate constant output driven by the bias $WB$, and no interference from the noisy input signal. Naturally, the initial learning rates are also all very low, because learning from noise is not a good strategy, but they quickly rise to coincide with the onset of the input signal at the age of three. By the age of seven, the system has caught up with the performance levels of Figure 2. Once again our simplified evolutionary approach leads to a sensible pattern of plasticity variations.

## 5 SCALE FACTOR MUTATIONS

As with all modelling endeavours, it is important to test the robustness of the results with respect to the implementational details. Naturally, in this case it is the encoding of the plasticity scale factor $s(t)$ that we need to be particularly careful about. If each point $\{s(t) : t = 1, \ldots, N\}$ defining the piece-wise linear function were simply allowed to evolve in isolation in the same manner as the weights and learning rates, we would actually end up with the rather noisy results shown in Figure 5.

The individual performance advantages that would keep the curves smooth, and reduce any unnecessary plasticity, are rather weak and get lost in the noise of our coarse simulations. This is particularly apparent after the age of about 10. The weakness is partly due to the error signals being relatively low after the maturation period is complete, and partly because it will be relatively unimportant if the fitness starts decreasing again after a number of children have already been produced, or if the majority of individuals normally die before reaching that age.

Fortunately, we can compensate for these limitations by variations of the plasticity scale factor mutations. First, we can prevent unnecessary plasticity (which will surely have an intrinsic cost in real systems) by allowing mutations which set random points $s(n)$ to zero. Then, it is unlikely in real systems to be efficient to have $s(t)$ varying wildly with age, so it is reasonable to encourage smoothness of $s(t)$ by allowing mutations which swap the values of random adjacent points $s(n)$ and $s(n+1)$. It was these simple variations that turned the noisy and inefficient results of Figure 5 into the smooth and efficient results of Figures 2, 3 and 4.

## 6 CONCLUSIONS

By simulating evolving populations of simple adaptable control systems we have seen that there is a natural propensity for the evolution of plasticities that vary sensibly with age, quite independently of any physical overheads of the plasticity. This is consistent with the well known "critical periods" of human brain development (Julesz & Kovacs, 1995). It is reasonable to expect that such an evolutionary approach will also be a profitable strategy for obtaining improved performance in adaptable systems for real world applications.

There are two competing effects at play. In order to survive in competition with fitter adults and/or a hostile environment, a newborn needs to adapt as quickly as possible to its environment. It also needs to adapt efficiently to its own maturation. Large plasticities will be beneficial for both. In adults, however, large plasticities can lead to an unstable learning system, in which unusual/extreme experiences can potentially result in a large shift of the systems' parameters with a serious reduction in overall fitness. Lower learning rates in this situation will allow smoother optimal parameter estimation and more consistently good responses in a varied environment. In this paper it has been demonstrated how a process of evolution by natural selection can result in a population of individual systems that deal with these conflicting requirements by having plasticities that vary appropriately with age under normal maturation, when late life adaptation is required, and when there is a dependence on the prior development of other sub-systems. We have also seen how appropriate changes to the implementational details (e.g. the plasticity scale factor mutations) can lead to vastly superior results.

In complex systems, such as the human brain, we can expect each of the various sub-systems to evolve appropriately for its own requirements, so there may well be no single global behaviour. The next stage of this work will be to develop and test larger scale and more realistic simulations of specific human sub-systems, and to explore explicitly how these ideas could be applied to the formulation of more efficient artificial adaptable systems for particular real world engineering applications.

### References

Baldwin, J.M. (1896). A New Factor in Evolution. *The American Naturalist*, **30**, 441-451.

Belew, R.K. & Mitchell, M. (Eds) (1996). *Adaptive Individuals in Evolving Populations.* Reading, MA: Addison-Wesley.

Bullinaria, J.A. (2001). Exploring the Baldwin Effect in Evolving Adaptable Control Systems. In: R.F. French & J.P. Sougne (Eds), *Connectionist Models of Learning, Development and Evolution*, 231-242. London: Springer.

Bullinaria, J.A. & Riddell, P.M. (2001). Neural Network Control Systems that Learn to Perform Appropriately. *International Journal of Neural Systems*, **11**, 79-88.

Jacobs, R.A. (1988). Increased Rates Of Convergence Through Learning Rate Adaptation. *Neural Networks*, **1**, 295-307.

Julesz, B. & Kovacs, I. (Eds) (1995). *Maturational Windows and Adult Cortical Plasticity.* Reading, MA: Addison-Wesley.

Levine, W.S. (1996). *The Control Handbook.* London: CRC Press.

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.

Yao, X. (1999). Evolving Artificial Neural Networks. *Proceedings of the IEEE*, **87**, 1423-1447.