From Biological Models to the Evolution of Robot Control Systems

By John A. Bullinaria

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

Attempts to formulate realistic models of the development of the human oculomotor control system have led to the conclusion that evolutionary factors play a crucial role. Moreover, even rather coarse simulations of the biological evolutionary processes result in adaptable control systems that are considerably more efficient than those designed by human researchers. In this paper I shall describe some of the aspects of these biological models that are likely to be useful for building robot control systems. In particular, I shall consider the evolution of appropriate innate starting points for learning/adaptation, patterns of learning rates that vary across different system components, learning rates that vary during the system's lifetime, and the relevance of individual differences across the evolved populations.

Keywords: Evolution, Control systems, Modelling, Biology, Robots

1. Introduction

There exists a natural concern about the possible adverse effects of using novel viewing devices such as night vision goggles for pilots or virtual reality head-mounted displays in amusement arcades (Sheehy & Wilkinson, 1989; Mon-Williams, Wann & Rushton, 1993; Kotulak & Morse, 1995). Their recreational use by children is particularly worrying given the more plastic nature of their visual systems (Rushton & Riddell, 1999). However, the wide ranges of individual differences found in the oculomotor control systems across the human population, and the possibility that different subsets within that range may be prone to different problems and require different remedial actions (Morse & Jiang, 2000), mean that it is empirically rather difficult to draw reliable conclusions concerning the validity of these concerns. One can certainly begin to quantify potential problems and corrections by examining existing users of such devices, but it would clearly be unethical to set up more widespread tests on children, or to experiment on them with new remedial actions when problems are found. An alternative would be to build sufficiently realistic models of the relevant parts of the human oculomotor control system, and subject those to demands of the kind required of humans using different types of viewing devices. In this way we might be able to identify simple predictors of which individuals are most likely to experience problems, and determine which corrective procedures might be most appropriate for each individual with a particular deficit.

For this reason, I have recently been involved in formulating increasingly sophisticated models of the development of the human oculomotor control system with a view to understanding better the problems that may occur under normal conditions, and also as a result of using novel viewing devices (e.g. Bullinaria, Riddell & Rushton, 1999; Bullinaria & Riddell, 2001). Building a model of the whole visual system, that takes in real images as its inputs, was not really feasible, nor actually necessary. Our main concern was to determine how accommodation (eye focusing) and vergence (eye rotation) are controlled so as to minimize blur and image disparity. As with any system that is required to respond appropriately under varying conditions to a range of different cues of varying degrees of accuracy, reliability and availability, and be able to adapt across various different timescales, designing an efficient oculomotor controller is a difficult task. Nevertheless, there is an enormous literature on linear control systems models based on empirical human data which already provide a good account of the performance of the adult oculomotor control system for unpredictable target sequences (e.g. Schor, Alexander, Cormack & Stevenson, 1992; Eadie & Carlin, 1995). However, they do not provide an account of the development of that system in individuals, nor the large ranges of individual differences that are found in humans. To do this we have had to incorporate the ability to learn from and adapt to typical visual cues, and take careful account of numerous maturational factors, such as the eyes growing further apart.

Our models have been set up with general architectures based on known physiology, and learn for themselves how to perform the given tasks as best they can from realistic sequences of visual demands. These models have been rather successful in that they can simulate adult human performance reasonably well, and also the developmental progression towards those abilities, yet they remain lacking. A recurring feature is that, although the models do learn human-like performance under normal conditions, the internal processes they develop (that can be deduced from their operation under various open loop conditions) depend crucially on such details as their initial conditions and time-dependent learning rates, which are extremely difficult to determine empirically. These factors will have been constrained by evolution in the human systems, and without simulating this evolutionary process it seems unlikely that we will be able to produce realistically constrained models, nor fully understand the normal and abnormal development of those systems (Bullinaria & Riddell, 2000). Particularly important are the innate starting points of the learning process, the advantages of having different learning rates for distinct components of the system, and the relevance of the critical periods of learning that are observed in humans. Some preliminary small scale simulations of human-like evolution of these factors (Bullinaria, 2001a,b) suggest that such an approach will not only result in improved models of human systems, but also result in much more efficient systems than researchers can reasonably be expected to design themselves. This has led to the suggestion that a similar approach might yield improved performance for robotic control systems.

The idea of using evolutionary techniques for creating robotic systems is already well established in the field of *Evolutionary Robotics* (e.g. Nolfi & Floreano, 2000). Evolutionary algorithms have also already 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 simulated evolution (e.g.

Yao, 1999). In this paper, however, I would like to concentrate on the effects of three related factors that appear to be crucial for the human system, and yet have perhaps not received adequate attention in the field of robotics. First, the so-called *Nature-Nurture Debate* and the distinctions between properties that are innate in each new individual, and those that must be learned from the environment during an individual's lifetime. Second, the ranges of *Individual Differences* and how these are constrained by evolution and learning. Third, the advantages and disadvantages of the *Critical Periods for Learning* that are often observed in human development. I shall argue that to study these issues properly, it is important that the evolutionary simulations are more closely aligned with biological evolution than is commonly the case.

2. The Baldwin effect

Initially we were primarily interested in how evolution by natural selection has resulted in the current population of humans, but now we also want to see how it can improve our artificial systems' (robots') abilities to learn from, and act in, their environment. Inevitably, that interaction between learning and evolution known as the *Baldwin Effect* (Baldwin, 1896; Hinton & Nowlan, 1987; Belew & Mitchell, 1996) will be crucial for understanding the processes involved. For present purposes this interaction occurs in two stages: (1) if a mutation improves the ability of the learning process to acquire better properties, then it will tend to proliferate in the population, and (2) if the learning process has an associated cost (e.g. requires time or energy), then its results will tend to get incorporated into the genotype and the learned behaviours will become innate. In effect, natural selection will generate an iterative sequence of adjustments to the innate starting points that reduces the need for learning. This will result in genetic assimilation of the learned characteristics, without Lamarckian inheritance. Then any population in a stable environment, that is able to reproduce accurately, can be expected to evolve so that its optimal behaviour is completely innate, and learning will no longer be required. However, if the system really does need to retain the ability to learn, for example to adapt to the system's own maturation or degradation (as we generally have in biological systems), or to adapt to unknown or changing environmental conditions, or to correct for statistical fluctuations or mutations in the procreation process, then only partial assimilation will occur. We should still end up with an efficient learning system, but the appropriate innate properties will no longer generally correspond to the final learned behaviour. This is where the nature-nurture debate comes in (Elman et al., 1996). Moreover, if natural selection cannot discriminate between different genotypes (i.e. different innate properties), then we will be left with a range of individual differences. I have recently discussed these two issues in more detail elsewhere (Bullinaria, 2001a). They both have a crucial effect on the evolution of efficient adaptable control systems.

3. Variable plasticity and critical periods

The third issue I wish to consider here is how evolution might lead to more efficient systems by allowing the emergence of variable plasticities (i.e. non-constant learning rates). It is certainly well known that human neural plasticity varies considerably
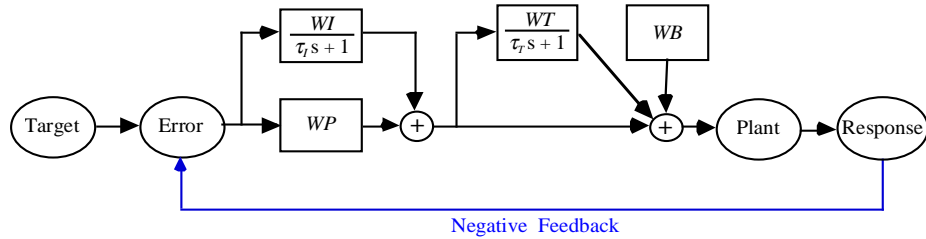
Figure 1. The simple feedback control model to be evolved. Learning results in its four weights/parameters WI, WP, WT, WB being adjusted so that it performs as best it can.

with age, and that there are often 'critical periods' during which learning must take place if the given task is to be mastered successfully (Greenough, Black & Wallace, 1987; 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 minimize the weight variations seen after each sample in online training, or to increase them to speed the saturation of sigmoidal activation functions 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 age-dependent plasticity found in humans, or if it has been more a matter of minimising the physical overheads of the plasticity. Given that robotic systems will generally have overheads somewhat different from biological systems, it is worth exploring this in some detail.

4. The control model

In this paper I shall discuss the issues introduced above in the context of the results from a series of explicit simulations of the evolution of some simple adaptable control systems. The overall aim will be to see what innate systems and learning strategies evolve naturally, and to explore how different strategies evolve under different circumstances. These can be expected to inform the issues for biological systems discussed above, and help us to formulate better adaptable controllers for our robotic systems.

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 feedback control systems can be applied quite generally (Levine, 1996). The input will be a natural sequence of target responses that will generally be supplied by another (probably rather complex) subsystem in the brain/robot. This is combined with the signal from the feedback loop to provide an error signal. Generally this will involve some complex computations with variable time delays, saturations and dead-zones (such as depth of field), although for the current study we shall restrict ourselves to a simple difference calculation with constant feedback latency. The error signal then feeds through a standard integral controller (a leaky inte-

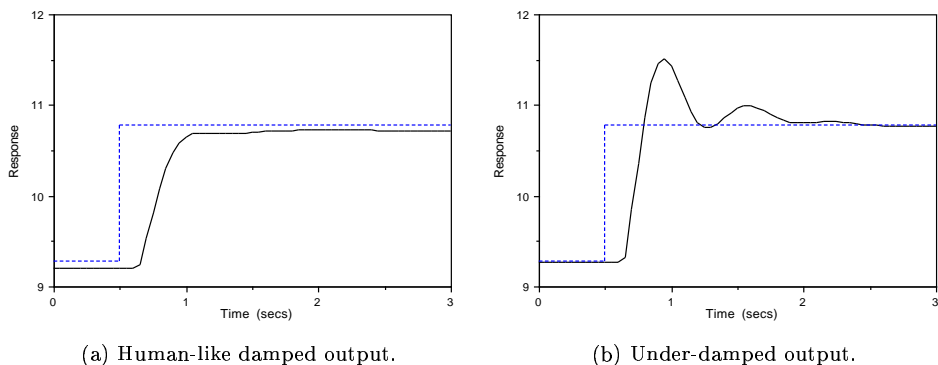(a) Human-like damped output.

(b) Under-damped output.

Figure 2. Responses (solid lines) resulting from our simplified control system, for two different parameter sets, when presented with a step in demand (dashed lines).

grator with gain/weight WI and time constant τ_I) and a standard proportional controller (a simple gain/weight WP). The combined outputs from these are added to a constant bias signal (of strength/weight WB) and a leaky integrator tonic signal (of gain/weight WT and time constant τ_T), and fed into the plant (which is approximated by another leaky integrator) to produce the final response. The bias provides an appropriate general purpose resting state, while the tonic allows short time-scale 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 (Schor et al., 1992). The system can equally well be regarded as a traditional control system (Levine, 1996), or as a dynamical network of leaky integrator neurons (Bullinaria & Riddell, 2001). It is simple enough to render the simulations tractable, yet complex enough to incorporate the essential features of many real control systems.

Simulating the evolution of this system will involve working with a large number of copies of the model, each with their own four adjustable parameters/weights $\mathbf{W}(t) = \{WI(t), WP(t), WT(t), WB(t)\}$ where t is the time/age of the individual model measured in simulated years. These parameters are learned by a simple on-line gradient descent algorithm that minimizes a cost function consisting of response error and smoothness (overshoot minimization) components, which would be readily available to the system, for representative sequences of response targets. The precise formulation of this learning algorithm has been discussed in some detail by Bullinaria & Riddell (2001), and the error versus smoothness trade-off set to match human performance. Corresponding to the learnable weights, then, each instantiation of the model will have four variable learning rates/plasticities $\mathbf{P}(t) = \{PI(t), PP(t), PT(t), PB(t)\}$. The model will also have various other parameters (the time constants τ_I and τ_T, 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 automatically produce appropriate damped responses to arbitrary discontinuous output requirements, as well as smooth pursuit of arbitrary continuous output changes (Bullinaria & Riddell, 2001).

Figure 2 shows typical human-like and under-damped responses to a step change in the target response.

For the purposes of this paper, I shall assume that all the learning rates in an individual 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 for now we are able to 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 were no plasticity variation we would have $s(t) = 1$ for all t. To evolve this function we need a convenient parameterization. Here we shall use one that involves few additional assumptions, namely set $s(t)$ to be piecewise linear with parameters $\mathbf{S} = \{s(t) : t = 1, \ldots, N\}$. This extends my earlier study (Bullinaria, 2001*b*) where I took $s(t)$ to be an exponential function determined by only two evolvable parameters. That parameterization, however, had the advantage of allowing a straightforward implementation of control systems that could evolve appropriate environment dependent plasticity variations.

The part of the current model's genotype that varies between individuals thus represents the $8 + N$ parameters $\{\mathbf{W}(0), \mathbf{P}(0), \mathbf{S}\}$. There is clearly nothing in this approach, apart from the inevitable increase in computational requirements, to prevent straightforward extensions to more complex control systems that involve any number of parameters.

5. Evolving the model

Simulating the 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 those processes in real (biological) 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 attempt to 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, or for complete physical robots, 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. For the purposes of our simplified model here, however, it is reasonable to assume that all other factors are equal across the population, and consider it to be a sufficiently good approximation to take 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 very easy to lose weak effects in the noise of the rather coarse simulations forced upon us by current computational resource limitations.

Given that, initially at least, we were aiming to replicate effects that arise in biological evolution, it was appropriate 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). If, as is often done, we were to train each member of the whole population for a fixed time and pick the fittest to breed and form the next generation, there would be no incentive for individuals to learn as quickly as possible, and efficient learners would not evolve. The natural alterna-

Table 1. *Summary of the evolutionary process*

A population of individual models each born with:
An architecture (fixed): As in figure 1
Four initial weights (evolved): $\mathbf{W}(0) = \{WI(0), WP(0), WT(0), WB(0)\}$
Four initial plasticities (evolved): $\mathbf{P}(0) = \{PI(0), PP(0), PT(0), PB(0)\}$
A plasticity scale factor (evolved): $\mathbf{S} = \{s(t); t = 1, ..., N\}$
Then during each simulated year:
An output scale factor matures between birth and age 10
The plasticities mature: $\mathbf{P}(t) = \{PI(t), PP(t), PT(t), PB(t)\} = s(t)\mathbf{P}(0)$
Good weights are learned: $\mathbf{W}(t) = \{WI(t), WP(t), WT(t), WB(t)\}$
The least fit individuals tend to die (with tournament selection)
The oldest individuals tend to die (at a constant rate after age 30)
The most fit individuals tend to procreate (with tournament selection)
The procreation process has:
Children's innate parameters set randomly in range spanned by two parents
Random variations/mutations may take parameters outside parental range

tive to this *generational* approach is a *steady state* strategy in which only a sub-set of the population is replaced at each time step (Whitley, 1989; Syswerda, 1991). Here, as in most biological systems, 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 their 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. Naturally, the innate characteristics that enable individuals to acquire good performance most quickly and reliably will tend to proliferate in the population.

Note that even when the best possible set of innate weights has evolved, the control system will still generally benefit from being plastic since that will allow each individual to fine tune its performance after their statistically noisy procreation process and/or being born into an unpredictable environment. Many biological systems also need some degree of plasticity to compensate for the changes (e.g. growing sizes) that naturally take place during their own maturation periods, and robots will need to compensate for related factors such as plant drift (e.g. gear wear). For the current study, such changes were simulated by introducing 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 to the general pattern of results that emerge.) In humans this maturation might correspond to changes in inter-pupillary 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 final adult

values. However, the pattern of plasticities that evolve will allow the system to learn most effectively how to optimize its weights throughout its life. The complete evolutionary process is summarised in table 1.

Ultimately, the simulations might benefit from more realistic encodings of the parameters, concepts such as recessive and dominant genes, gender differences, 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 proves adequate. An important consideration, however, is that limited computational resources will generally only permit rather coarse approximations of biological evolutionary processes, and it is important to fix the various simulation parameters appropriately. For example, if all the individuals were able to learn how to perform the given task perfectly by the end of their first year, and we only tested their performance once per year, then the advantage of those that learn in two months over those that take ten is lost and the simulated evolution would not be very effective. Since the individuals were allowed to evolve their own learning rates, this had to be controlled by limiting the amount of training data each individual experienced in each year. Choosing a fixed population size of only 100 was a trade-off between maintaining genetic diversity and running the simulations reasonably quickly. The death rates were set in order to produce reasonable age distributions, and to prevent the population from becoming dominated by skilled adults who killed off most of the children before they had the chance to learn how to perform well. This meant around 10 deaths per year due to competition, and another 4 individuals over the age of 30 dying each year due to old age. The procreation and mutation parameters were chosen to speed the evolution as much as possible without introducing too much noise into the process. Coding such a system in C typically resulted in around 1,000 simulated years per CPU hour on an average UNIX workstation. These details were kept constant across all the simulations I shall now present.

6. Simulation results

Inevitably, starting with random initial populations, and evolving them in randomly generated environments, will lead to some variability between the results from different runs of the same system. However, the general patterns of results were found to be quite robust with respect to these randomizing effects, and so I shall present results from typical runs, rather than averages over many runs which tend to mask many of the crucial details.

The main simulation results for a typical run of the basic system described above are shown in figure 3. The initial population's initial weights $\mathbf{W}(0)$ and learning rates $\mathbf{P}(0)$ were assigned random values from the range $[0, 20]$, but only a relatively small sub-space of those parameters actually correspond to stable control systems. Any individuals that proved unstable on a small input test sequence were immediately replaced by new random individuals until the whole initial population was stable with respect to that test sequence. This ensured diversity in the early generations by preventing the whole population from being derived from the few stable individuals in the initial random set. As can be seen from graph (a) in figure 3, this resulted in the initial population averages of the initial weights starting off quite

near their final evolved values, though the initial populations did actually have quite a spread around those mean values. It is a major advantage of the evolutionary approach that the evolving population naturally tends to keep itself away from the unstable regions of parameter space, and thus to a large extent obviates the need for a stability analysis phase of the design process.

The mean learning rates/plasticities of the initial generation were not as constrained as the corresponding initial weights, but still they quickly evolved to take on appropriate values, as shown in graph (b) of figure 3. An important point to notice is the large variation between the learning rates that emerge for the different components – for example, there is nearly two orders of magnitude difference between PT and PI. Using a single learning rate for the whole system, as is common in neural network modelling, would clearly not be a good strategy in this type of system. Moreover, determining appropriate different values without employing an evolutionary strategy would be a formidable task for a human designer.

The evolved parameters have emerged because they result in good values for the weights throughout the individuals' lives. All the weights will need an initial fine tuning to remove the noise introduced by the procreation process, then some weights (WI and WP) need to adjust during the maturation period, while others (WT and WB) need little further change. In figure 3, the plots (c) and (d) of $WP(t)$ and $WT(t)$ against age t for our typical evolved population show this quite clearly. We can also see that, even for individuals of the same age, there is still a fair degree of variability in the parameter values. Graph (e) shows this variability more clearly by plotting the standard deviations across the evolved population of the parameter values at each age, normalised by the corresponding average values. These ranges of *individual differences* vary with age and ultimately depend upon the sensitivity of the fitness function with respect to the corresponding parameters. For the biological case of oculomotor control discussed earlier, this is particularly important since, although the individual differences emerge because of their irrelevance to fitness under normal conditions, they may be crucial to the responses under unnatural viewing conditions such as the wearing of virtual reality head-mounted displays. Subjects from some sub-ranges of individual differences may be unaffected, while others may have serious problems (Morse & Jiang, 2000). For robotic systems, we may wish to employ the entire range of individuals so that there is a good chance that at least one of them will be able to deal with any unforeseen circumstances, or we may wish to constrain the range further by imposing additional performance or reliability requirements. In either case, a good understanding of the ranges and causes of individual differences will be of great benefit.

In the final graph (f) of figure 3 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 an evolutionary process. There are two competing factors that determine what is an appropriate plasticity for each age. In order to survive in competition with fitter adults and/or a hostile environment, a newborn needs to be able 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 or extreme experiences can potentially result in a large shift of the systems' parameters with a serious reduction in overall fitness. Lower
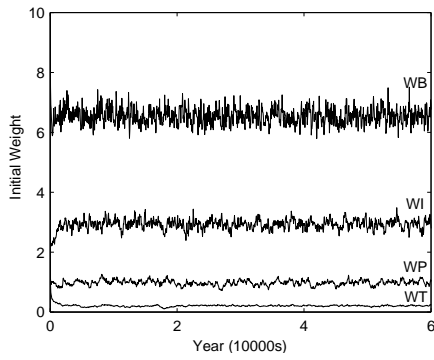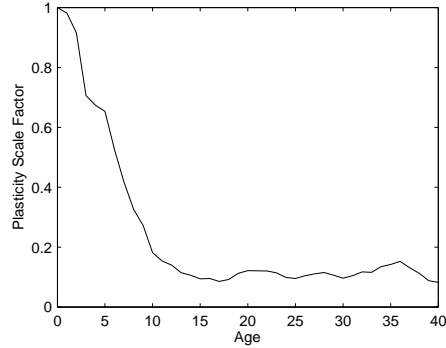
(a) Evolution of the innate weights.
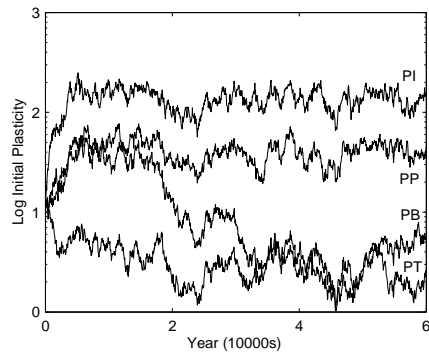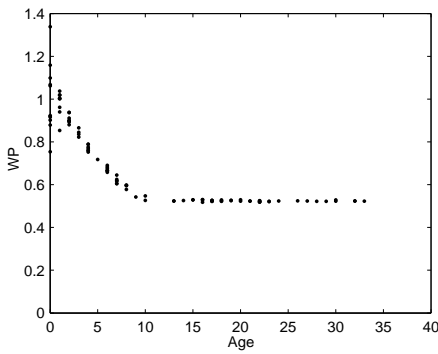
(b) Evolution of innate initial plasticities.

(c) Adaptation of the weight WP.
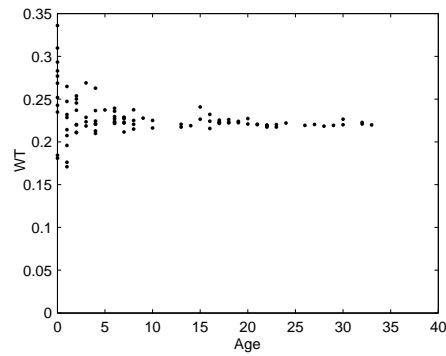
(d) Adaptation of the weight WT.
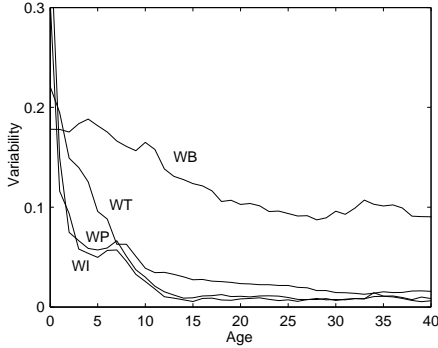
(e) Age dependence of individual differences.

(f) Age dependence of the plasticity

Figure 3. Evolution and learning in a typical simulation of the basic system. Appropriate initial weights and learning rates evolve quite quickly (a, b), and these result in good weights at each age (c, d). The remaining ranges of individual differences (e) vary between the four weights depending on how crucial each is to the fitness, and on how quickly they are optimized by the learning. Individuals in the evolved population have plasticities that fall rapidly between birth and the end of their maturation period (f).

learning rates in this situation will allow smoother optimal parameter estimation and more consistently good responses in a varied environment.

The results from the basic system naturally lead to the question of what happens if an individual does need to be able to learn or adapt later in life, after the standard learning period is over. For example, a human might need to adapt to new prescription spectacles, or a robot may need to adjust because a particular component invariably wears out and gets replaced after a certain number of years. There is a traditional saying that 'old dogs cannot learn new tricks', but it seems unlikely that evolution would allow the plasticities to decay away to very small values in situations where late life adaptation is regularly required. To investigate this, three representative forms of late life change were inflicted on the model's output scale factor. There is no need to specify whether this variation corresponds to an internal factor (such as the need to compensate for system deterioration or damage) or an external factor (such as the need to adapt to changes in the operating environment), since they will have the same effect.

The simplest case to analyse is when there is always the same adaptation needed at a particular age. 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. (The precise details of this variation are not crucial to the general pattern of results obtained.) Figure 4 shows how this changes the simulation results from those of the basic model in figure 3. The most direct consequence is in the graph (c) of $WP(t)$ where we see that the required step change at age 20 is learned successfully and quickly. We also see a corresponding local peak in the individual differences graph (e) bought about by variations in the speed of the adaptation process. The plot (f) of the plasticity scale factor $s(t)$ shows the same initial fall as before, but then a peak to give the increased plasticity required 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.

Obviously, the need for real late-life adaptation will rarely be so predictable, and so a second set of simulations was carried out in which the sudden change in output scale factor occurred at random ages after the end of the maturation period. Figure 5(a) shows that in this case the previous peak in increased plasticity becomes spread out over all ages so that each individual is able to adapt whenever it is required. The plot of WP against age clearly includes some individuals that have not yet experienced the need to adapt, and some who have successfully adapted. The level of late life plasticity here is sufficient, but still somewhat lower than the peak in figure 4(f), suggesting that suffering a slower adaptation when needed is being balanced against having too much plasticity when it might not be.

The third form of late-life adaptation considered involved the need to respond to a gradual change in the output scale factor from maturation till death. This might correspond to gradual plant deterioration in biological or robotic systems. The simulation results presented in figure 5(b) show that a reduced level of late-life plasticity is again sufficient to cope with the level of adaptation required here. It seems safe to conclude that, despite the tendency to evolve plasticities that fall rapidly between birth and the end of the maturation period, our models can deal appropriately with the need for late life adaptation.
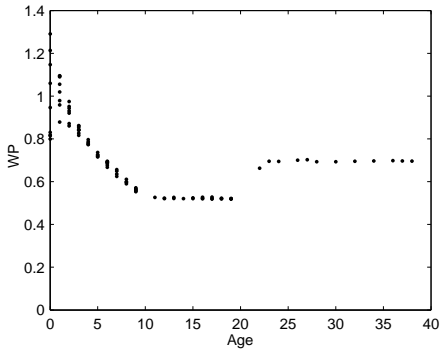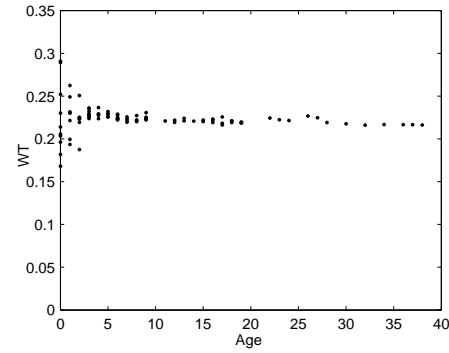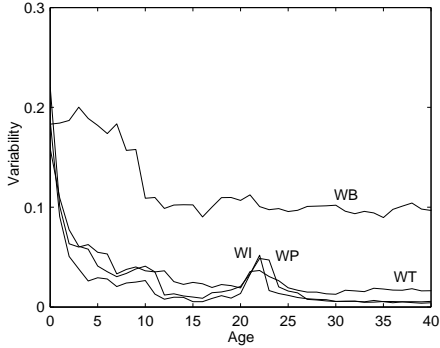
The third and final situation we shall consider here, that often arises in biological development, is when one level of processing is highly dependent on signals coming
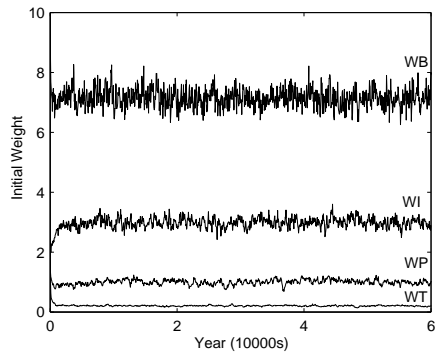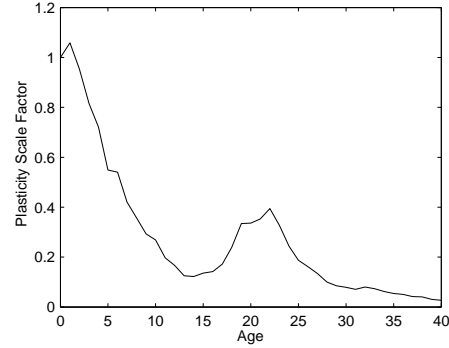
(a) Evolution of the innate weights.
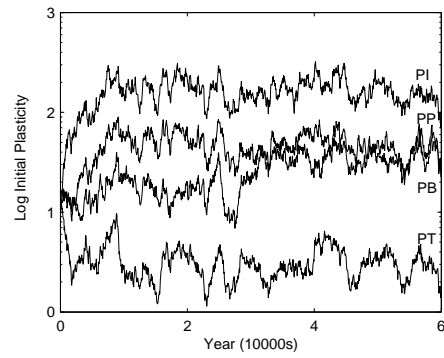
(b) Evolution of innate initial plasticities.

(c) Adaptation of the weight WP.

(d) Adaptation of the weight WT.

(e) Age dependence of individual differences.

(f) Age dependence of the plasticity

Figure 4. Evolution and learning in a typical simulation when late-life adaptation is required suddenly and predictably at age 20. Again the initial weights and learning rates evolve quickly and appropriately (a, b) to result in good weights at each age (c, d). Evolution here produces individuals in the final population with plasticities appropriate for the learning or adaptation that is forced upon them, with a peak around the age of 20 that enables them to cope efficiently with the changes required at that age (f).
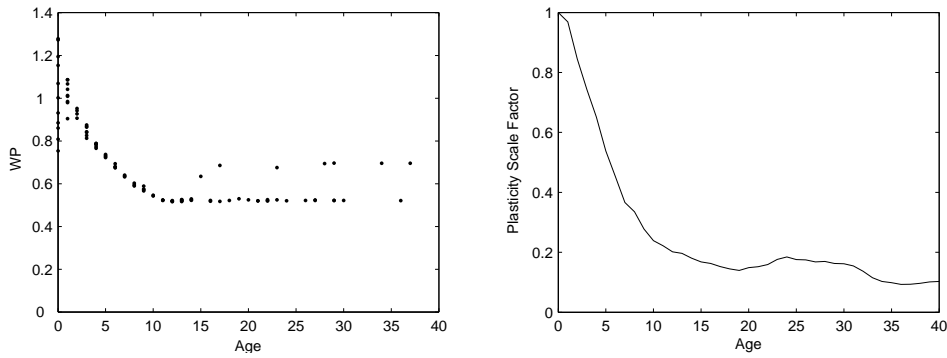
(a) Sudden adaptation required at random ages.
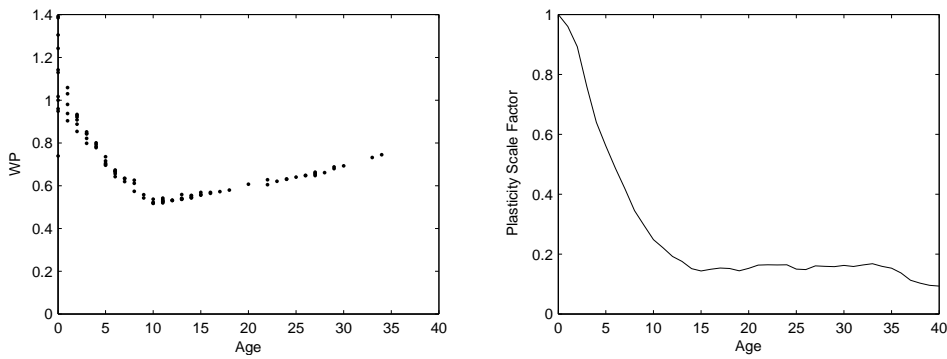
(b) Gradual adaptation required from age 10 onwards.

Figure 5. Age dependencies of parameter WP and plasticity when different forms of late life adaptation are required.

from another sub-system. If the sub-system supplying those signals is not fully functional, it might be sensible to wait until it is before beginning to learn how to use its signals. For example, the adult human vergence (eye rotation) system uses an image disparity signal, and humans have to wait until 12-16 weeks of age before that signal relatively suddenly becomes available. A robotic system may similarly have to wait until its sensors have been properly calibrated for its environment. To simulate such effects in our basic model, the error signal was replaced by low-level noise for each individual until they reached three years of age.

Figure 6 shows how this affects the standard results of figure 3. The changes here are rather clear. First, we see in graph (a) that the initial/innate weights WI, 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. Interestingly, this kind of constant response is actually observed in the accommodation (eye focusing) of new-born humans (Hainline, Riddell, Grose-Fifer & Abramov, 1992). Naturally, the initial learning rates shown in graph (b) are also all very low, because learning from noise is obviously not a good strategy, but the plasticity scale factor $s(t)$ seen in graph (f) evolves to ensure that the plasticities

(a) Evolution of the innate weights.

(b) Evolution of innate initial plasticities.

(c) Adaptation of the weight WP.

(d) Adaptation of the weight WT.

(e) Age dependence of individual differences.
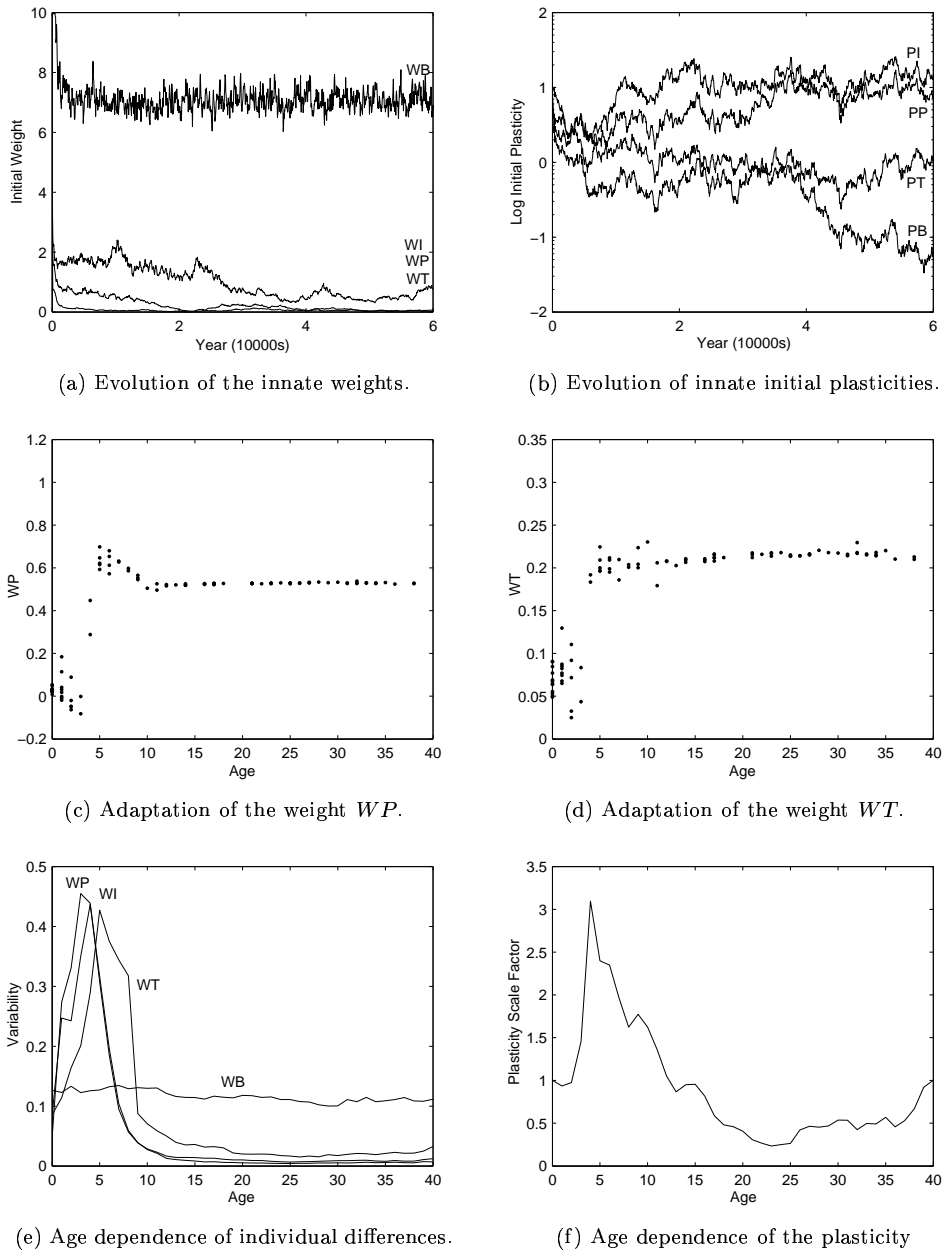
(f) Age dependence of the plasticity

Figure 6. Evolution and learning in a typical simulation when there is a dependency on the development of lower level sub-systems. Here, relatively low values emerge for the initial weights (apart from the bias WB) and initial learning rates (a, b), so that the system avoids processing or learning from signals that are just noise (c, d). This results in a population of individuals with critical periods for learning starting around age three (f), and a correspondingly different profile of individual differences (e).

rise quickly to coincide with the onset of the useful input signal at the age of three. By the age of seven, the system has caught up with the performance levels of figure 3 as seen by the parameter values in graphs (c) and (d). In graph (e) we see that there are increased ranges of individual differences during the learning period, but once again our biological-style evolutionary approach leads to sensible patterns of initial parameters and plasticity variations.

7. Robustness of the results

As with all modelling endeavours, it is important to test the robustness of the results with respect to the implementational details. The choice of representation for the parameters is often a crucial factor. For the current study, the weights and learning rates are reasonably straightforward, but the encoding of the plasticity scale factor $s(t)$ is something 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 initial weights and learning rates, we would actually end up with the rather noisy results shown in figure 7. We get the same general pattern of critical periods for learning, but there is a noticeable lack of smoothness.

Biological systems will generally have their plasticity controlled by physical or chemical processes which are likely to contribute to smoothing of the plasticity time variations, and they also have overheads that will discourage plasticity when it is not necessary. In our simplified models, or artificial systems, the plasticity is more likely to be just another software parameter with no additional constraints, unless we build some in explicitly. As we have already noted, there are individual performance advantages that should keep the curves smooth, and reduce any unnecessary plasticity, but these are rather weak and tend to get lost in the noise of our rather coarse simulations. We see from figure 7 that this is particularly apparent in individuals over the age of about ten. The weakness is partly due to the error signals being relatively low anyway 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 introducing some simple variations into the plasticity scale factor mutations. First, we can prevent unnecessary plasticity, which will quite likely have an intrinsic cost in biological systems, by allowing mutations which set random points $s(t)$ to zero. Then, since it is unlikely to be efficient for any system to have $s(t)$ varying wildly with age, it is reasonable to encourage smoothness of $s(t)$ by allowing mutations which swap the values of random adjacent points $s(t)$ and $s(t+1)$. The standard procreation processes will then tend to convert them to local average values after several generations. It was these natural mutations which turned noisy and relatively inefficient results like those of figure 7 into the smooth and efficient results seen in figures 3, 4, 5 and 6. They clearly do not prevent high plasticity or sharp variations in plasticity when they are needed, but they do discourage them when they are not.

Another detail that one should always check for is any significant dependence on the evolutionary initial conditions. Such dependencies on initial conditions are well known to be widespread in systems, such as neural networks, which learn how

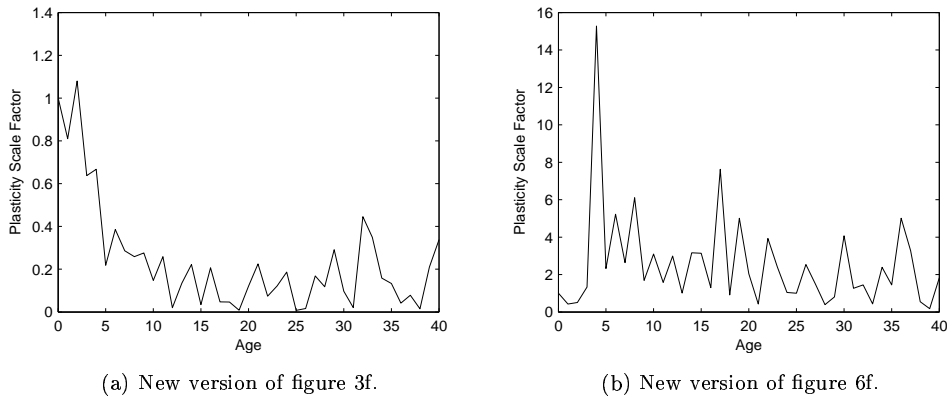(a) New version of figure 3f.   (b) New version of figure 6f.

Figure 7. Plasticity scale factors arising in typical runs that employed a slightly different implementation to that used for figures 3 and 6. Although the general patterns of age dependence are the same, they are considerably noisier.

to perform (e.g. Kolen & Pollack, 1991). Here we started with a diverse population of initial weights and learning rates because an earlier systematic study of a similar control system (Bullinaria, 2001a) indicated that this approach tends to lead to the fastest and most reliable evolution. Figure 8 demonstrates how this factor affects the results by plotting the initial (newborn) and final (adult) distributions of a typical parameter as the system evolves. Graph (a) shows the situation for the approach described above – the initial weights quickly settle around an appropriate value but have quite a wide distribution as a result of all the noise in the procreation process. Learning, which is available from the outset, allows individuals to adjust their weights during their maturation period so they end up with the somewhat narrower distribution of appropriate adult values.

In graph (b) we see that starting with very low (near zero) initial weights and learning rates tends to result in good general purpose weights evolving very quickly, and then the ability to adapt the weights to coincide appropriately with maturation emerges somewhat later (after about 30000 years in this particular simulation). Graph (c) shows that starting the evolution with a wide distribution of learning rates, but low initial weights, results in the initial weights being much slower to evolve to sensible values because the individuals are able to cope reasonably well by learning from any initial weights. In this case it takes some time for the cost of learning to cause the initial weights to drift into their optimal values (about 10000 years in this particular simulation). This last case provides a clear and explicit demonstration of how the genetic assimilation of learned behaviour (i.e. learned parameter values) can occur automatically, without Lamarckian inheritance, to reduce the inherent costs of learning (i.e. periods of poor performance).

One might wonder if the ability to learn or adapt will still evolve in the absence of the strong driving force that results from the need to compensate for maturation. Graph (d) in figure 8 arises in the case where there is no maturation, and the initial population has very small values for both the initial weights and learning rates. Not surprisingly, the initial and adult weight values are now both distributed around
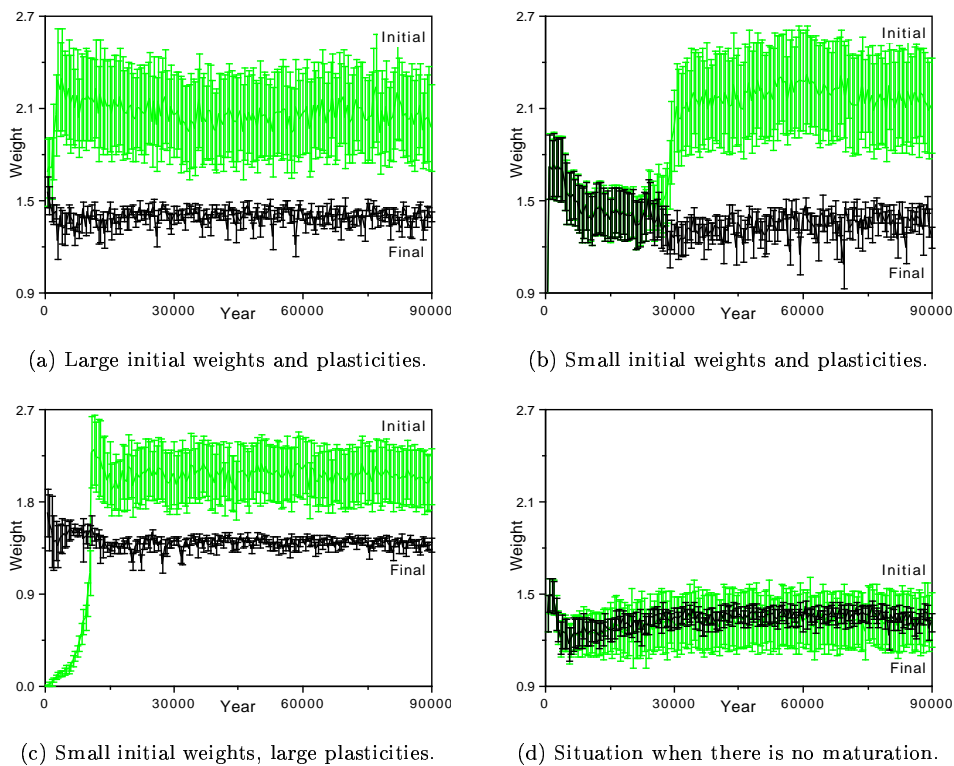
(a) Large initial weights and plasticities.

(b) Small initial weights and plasticities.

(c) Small initial weights, large plasticities.

(d) Situation when there is no maturation.

Figure 8. The variability across the population of the newborn and adult values of the weights WI throughout the evolutionary process. Graph (a) corresponds to the basic model described above. Then graph (b) shows what happens if the initial population all have very small initial weights and learning rates, rather than a wide range of values. If the initial population has very small values for the initial weights, but a large distribution of learning rates, we end up with the situation in graph (c). Finally graph (d) shows what happens when the system has no maturation to compensate for.

the same appropriate values, but still the ability to learn evolves to enable the wide initial weight distribution to be narrowed quickly so that all individuals are close to the most appropriate weight values for good performance.

In the above we have seen how the evolutionary initial conditions affect the speed of evolution, but they have little effect on the final evolved population. In more complex systems, however, there may be many local maxima in the fitness with respect to the system parameters, and the evolutionary initial conditions may also affect where the population ends up. In particular, we may not necessarily arrive at the most fit system possible. A preliminary investigation into this problem has already been carried out (Bullinaria, 2001*a*). The simple control system of figure 1 was augmented by an additional open loop input that injects an approximate burst signal into the system whenever there is a step change in required response. Such an approach is often used by biological systems to get them into the range where a feedback signal is available. In human oculomotor control, for example, there is a

point at which the disparity between the images from the two eyes becomes too large to allow the computation of an error signal, but some estimation of object distance generates an open loop signal that brings the system into range. For robots, an object might be too far away for a particular sensor to provide a reliable feedback signal, but some other information may be available to provide an appropriate open loop signal. In simulations of this situation, particular properties of the environment allowed the parameter determining the strength of the open loop signal, and the bias parameter, to be able to compensate partially for each other. Their evolved values tended to depend on the starting points of the evolution, and have much larger ranges of individual differences than usual. Such compensatory effects are likely to become increasingly common as we scale up to systems of ever increasing complexity and redundancy.

Making sure that the population starts and remains diverse enough to explore all the potential local fitness maxima will be an additional challenge for this approach if we are aiming to create the best possible artificial systems. With biological systems modelling, on the other hand, we may have to constrain carefully the diversity in order to take account of known evolutionary history. For example, the accommodation and vergence systems in human oculomotor control are both feedback control systems like figure 1, one primarily driven by blur, and the other primarily driven by disparity. However, blur and disparity both depend on the distance of the visual target and are consequently highly correlated. This makes it potentially useful for cross-links to develop (or evolve) between the two systems, especially if one happens to be faster or more reliable than the other under certain conditions. It is quite likely, though, that accommodation evolved much earlier than vergence (Horridge, 1987; Fernald, 2000), and so there may be more of a natural tendency for accommodation to drive vergence than vice versa. We know empirically that there are actually wide ranges of individual differences in these cross links in humans, and that they are likely to be crucial when it comes to problems arising from using non-standard viewing devices. Clearly we must exercise some care to deal appropriately with these issues.

The fact that we need not, and probably should not, constrain our evolved robotic systems to match known biological evolutionary history, is one area where our robotic systems have the potential for performing even better than the corresponding human systems.

8. Conclusions

In this paper I have demonstrated how allowing populations of simple adaptable control systems to evolve by natural selection can produce individual systems that are able to perform and adapt efficiently under a range of conditions. A Baldwin (1896) type interaction between learning and evolution results in the emergence of appropriate innate values for each parameter in the system, along with appropriate (different) learning rates for each parameter. We also observe a natural propensity for the evolution of leaning rates (plasticities) that vary sensibly with age, quite independently of any physical overheads of the plasticity. These are able to cope well with normal maturation, when late-life adaptation is required, and when there are dependencies on the prior development of other sub-systems. Ranges of individual

differences arise naturally, and we can see how these relate to the redundancy in the corresponding parameters.

This whole approach was originally motivated by the need to build realistic models of human/biological systems, consistent with the well known 'critical periods' of human brain development, nature versus nurture distinctions, and patterns of individual differences. However, all the indications are that such an evolutionary approach will also prove a profitable strategy for obtaining improved performance in systems for larger scale real world applications requiring adaptable controllers – such as robotics. By appropriate weakening of the constraints of biological evolution, or enforcement of additional performance or reliability constraints, we may even be able to use this approach to produce systems superior to those found in nature.

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. 2001a Exploring the Baldwin effect in evolving adaptable control systems. In *Connectionist models of learning, development and evolution* (ed. R. F. French & J. P. Sougne), pp. 231-242. London: Springer.

Bullinaria, J. A. 2001b Evolving variable plasticity in neural systems. In *Advances in fuzzy systems and evolutionary computation* (ed. N. Mastorakis), pp. 236-241. World Scientific and Engineering Society Press.

Bullinaria, J. A. & Riddell, P. M. 2000 Learning and evolution of control systems. *Neural Network World* **10**, 535-544.

Bullinaria, J. A. & Riddell, P. M. 2001 Neural network control systems that learn to perform appropriately. *International Journal of Neural Systems* **11**, 79-88.

Bullinaria, J. A., Riddell, P. M. & Rushton, S. K. 1999 Neural network modelling of the linkages between accommodation and vergence. *Investigative Ophthalmology and Visual Science* **40**, S362.

Eadie, A. S. & Carlin, P. J. 1995 Evolution of control system models of ocular accommodation, vergence and their interaction. *Medical & Biological Engineering & Computing* **33**, 517-524.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. 1996 *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Fernald, R. D. 2000 Evolution of the eyes. *Current Opinion in Neurobiology* **10**, 444-450.

Hainline, L., Riddell, P., Grose-Fifer, J. & Abramov, I. 1992 Development of accommodation and vergence in infancy. *Behavioural Brain Research* **49**, 33-50.

Hinton, G. E. & Nowlan, S. J. 1987 How learning can guide evolution. *Complex Systems* **1**, 495-502.

Horridge, G. A. 1987 The evolution of visual processing and the construction of seeing systems. *Proceedings of the Royal Society London* **B230**, 279-292.

Greenough, W. T., Black, J. E. & Wallace, C. S. 1987 Experience and brain development. *Child Development* **58**, 539-559.

Jacobs, R. A. 1988 Increased rates of convergence through learning rate adaptation. *Neural Networks* **1**, 295-307.

Julesz, B. & Kovacs, I. 1995 *Maturational windows and adult cortical plasticity.* Reading, MA: Addison-Wesley.

Kolen, J. F. & Pollack, J. B. 1991 Back propagation is sensitive to initial conditions. *Complex Systems* **4**, 269-280.

Kotulak, J. C. & Morse, S. E. 1995 Oculomotor responses with aviator helmet-mounted displays and their relation to in-flight symptoms. *Human Factors* **37**, 699-710.

Levine, W. S. 1996 *The control handbook.* London: CRC Press.

Mon-Williams M., Wann, J. & Rushton, S. 1993 Binocular vision in a virtual world: Visual deficits following the wearing of a head mounted display. *Ophthalmic Physiological Optics* **13**, 387-391.

Morse, S. E. & Jiang, B. C. 2000 Oculomotor function after virtual reality use differentiates symptomatic from asymptomatic individuals. *Optometry and Vision Science* **76**, 637-642.

Nolfi, S. & Floreano, D. 2000 *Evolutionary robotics.* Cambridge, MA: MIT Press.

Rushton, S. K. & Riddell, P. M. 1999 Developing visual systems and exposure to virtual reality and stereo displays: Some concerns and speculations about the demands on accommodation and vergence. *Applied Ergonomics* **30**, 69-78.

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.

Sheehy, J. B. & Wilkinson, M. 1989 Depth perception after prolonged usage of night vision goggles. *Aviation, Space, and Environmental Medicine* **60**, 573-579.

Syswerda, G. 1991 A study of reproduction in generational and steady-state genetic algorithms. In *Foundations of Genetic Algorithms* (ed. G. Rawlins), pp. 94-101. San Mateo, CA: Morgan Kaufmann.

Whitley, D. 1989 The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. In *Proceedings of the Third International Conference on Genetic Algorithms* (ed. J. D. Schaffer), pp. 116-123. San Mateo, CA: Morgan Kaufmann.

Yao, X. 1999 Evolving artificial neural networks. *Proceedings of the IEEE* **87**, 1423-1447.