Artificial neural networks and the study of evolution of prey coloration

Sami Merilaita*

Department of Zoology, Stockholm University, 10691 Stockholm, Sweden

In this paper, I investigate the use of artificial neural networks in the study of prey coloration. I briefly review the anti-predator functions of prey coloration and describe both in general terms and with help of two studies as specific examples the use of neural network models in the research on prey coloration. The first example investigates the effect of visual complexity of background on evolution of camouflage. The second example deals with the evolutionary choice of defence strategy, crypsis or aposematism. I conclude that visual information processing by predators is central in evolution of prey coloration. Therefore, the capability to process patterns as well as to imitate aspects of predator's information processing and responses to visual information makes neural networks a well-suited modelling approach for the study of prey coloration. In addition, their suitability for evolutionary simulations is an advantage when complex or dynamic interactions are modelled. Since not all behaviours of neural network models are necessarily biologically relevant, it is important to validate a neural network model with empirical data. Bringing together knowledge about neural networks with knowledge about topics of prey coloration would provide a potential way to deepen our understanding of the specific appearances of prey coloration.

Keywords: aposematism; camouflage; evolutionary simulation; mimicry; perception; predation

1. INTRODUCTION

In this paper, I will examine the use of artificial neural networks in the study of prey coloration as an adaptation against predation. Prey coloration provides numerous spectacular examples of adaptation (e.g. Cott 1940; Edmunds 1974; Ruxton *et al.* 2004). These include prey colour patterns used to disguise and make their bearers difficult to detect as well as brilliant colorations and patterns that prey may use to deter predators. As a consequence, prey coloration has been a source of inspiration for biologists since the earliest days of evolutionary biology.

The anti-predation function of prey coloration is evidently a consequence of natural selection imposed by predation. More specifically, it is the predators' way of processing visual information that determines the best possible appearance of the coloration of a prey for a given anti-predation function and under given conditions. Since predators' ability to process visual information has such a central role in the study of prey coloration, it follows that we need models that enable us to capture the essential features of such information processing.

An artificial neural network can be described as a data processing system consisting of a large number of simple, highly interconnected processing elements (artificial neurons) in an architecture inspired by biological nerve systems (Tsoukalas & Uhrig 1997). They provide a technique that has been applied in various disciplines of science and engineering for tasks such as pattern recognition, categorization and

decision making, as well as modelling tools in neural biology (e.g. Bishop 1995; Haykin 1999). The structural and functional similarities between artificial neural networks and biological neural systems are an often mentioned fact that has also drawn the attention of behavioural and evolutionary ecologists to artificial neural networks. These similarities include the structure that constitutes a network of simple processing units (neurons), processing of data in a parallel mode and behaviours that correspond to memory, learning and generalization (e.g. Enquist & Arak 1998; Ghirlanda & Enquist 1998). The features of artificial neural networks make them an appealing tool for modelling biological information processing.

Here, I will scrutinize the use of neural network models for the study of prey coloration. First, I will briefly review the function of prey coloration as an adaptation against predation. Then, I will in general terms describe simple artificial neural network models and how they can be applied for study of prey coloration, followed by a more detailed description of two examples of studies applying neural networks. Finally, I will discuss the suitability of such models for the study of prey coloration. Hopefully, this will help to introduce researchers of prey coloration to neural networks as well as researchers using neural networks to the study of prey coloration.

2. PREY COLORATION AS AN ADAPTATION AGAINST PREDATION

Animals employ colours and patterns in manifold ways to decrease the risk of predation. The two main functions of anti-predator coloration are concealment and signalling to predators. The most obvious way for a

*sami.merilaita@zoologi.su.se

One contribution of 15 to a Theme Issue 'The use of artificial neural networks to study perception in animals'.

prey to achieve concealment (crypsis or camouflage) is through a visual match between its body coloration and the background (Cott 1940; Endler 1978; Ruxton *et al.* 2004; Merilaita & Lind 2005). Such match in the colours and pattern geometry between a prey and its background in the eyes of the predator will minimize the amount of information that the predators can use for detection. This may appear simple to achieve, but there are some complications to it (Merilaita *et al.* 1999). First, practically all habitats vary spatially, and therefore a colour pattern that matches one site or microhabitat does not necessarily provide an equally good match at another part of the habitat. Background matching will in such a case be maximized, either by a coloration that maximizes background matching in that microhabitat only in which the risk of becoming detected by a predator is highest, or by a coloration that compromises the requirements of all the microhabitats, providing the best overall match.

In addition to habitat heterogeneity, also the outline of prey body shape may limit the benefit from background matching. This is because the discontinuity between body surface and background may make the prey outline conspicuous. Thus a characteristic shape of body or some part of it may give away a prey even if it were to closely match its background. However, there is some evidence supporting the idea that colour pattern may be employed to hamper detection based on outline and body shape (Merilaita 1998; Cuthill *et al.* 2005). Such disruptive coloration may aim to attract predators' attention away from the outline to distractive markings or to false outlines (Cott 1940; Merilaita & Lind 2005).

Another feature that may provide predators with helpful information is pattern symmetry. It has been suggested that for prey that are typically seen by predators from an angle that reveals their symmetry, pattern symmetry may aid detection. A recent study showed that pattern symmetry may indeed decrease crypsis, but such cost due to symmetry was not equally high for all patterns (Merilaita & Lind 2006). This means that we should expect natural selection to favour symmetric, cryptic patterns with minimized cost due to symmetry.

To summarize, in order to reduce the risk of becoming detected by predators, a prey may use its colour pattern to decrease the amount of information necessary for detection through resemblance to background. However, because it is seldom possible to completely eliminate the availability of such information, it may be beneficial for the prey also to use colour pattern to provide predators with misleading information about its appearance. Our surprisingly scarce knowledge about these and other questions concerning cryptic coloration suggests that there are plenty of opportunities for future research in this area.

Coloration may also be employed in post-detection defences. Some prey use coloration to, honestly or deceitfully, signal predators that they possess a defence which makes them an unprofitable or dangerous catch. Prey coloration that is associated with such unprofitability or risk and that decreases the risk of attack, because predators tend to avoid prey with such coloration (due to a learned or an innate response), is

called an aposematic warning signal (Cott 1940; Edmunds 1974; Ruxton *et al.* 2004). Such colorations are often striking. One important factor affecting natural selection on aposematic prey is the strengthening of a predators' avoidance response towards aposematic prey with increasing abundance of the aposematic prey type (e.g. Lindström *et al.* 2001).

An aposematic prey may offer another prey an opportunity to benefit from its defence, if the latter can mimic its appearance (Cott 1940; Edmunds 1974; Ruxton *et al.* 2004). Such mimicry has traditionally been divided into Batesian and Müllerian mimicry. In Batesian mimicry, an undefended prey has evolved a similar appearance to an aposematic prey, resulting in incorrect categorization and reduced predation risk for the mimetic species. On the other hand, the model (aposematic) species may suffer from dilution of its defence, especially if the avoidance is based on a learned response. Thus, while the mimetic species is under selection for increased resemblance, the model species is expected to be under selection for distinctiveness from the mimetic and other less defended prey species.

In Müllerian mimicry, there are two or more prey species with similar appearances and the predator experiences them all as unprofitable (Cott 1940; Edmunds 1974; Ruxton *et al.* 2004). In contrast to Batesian mimicry, in Müllerian mimicry all the involved prey species benefit from the common resemblance. This is because a large number of prey with similar appearance can be expected to increase the efficacy of avoidance learning by predators and to decrease the predation risk by naive predators per individual prey. Note, however, that two prey species which have similar appearances and are both unprofitable, need not obviously be equally harmful to a predator, and it has been suggested that it may be more realistic to consider Batesian and Müllerian mimicry as the opposite ends of a continuum instead of considering them as two discrete strategies (Speed 1993; Ruxton *et al.* 2004). Interestingly, a recent experiment suggests that qualitative variation in defence between Müllerian mimics is important: different defences may strengthen the avoidance response and hence increase the mutualistic benefit from Müllerian mimicry (Skelhorn & Rowe 2005). These factors are likely to affect the evolutionary dynamics of the appearances of the mimetic species.

In addition to honest or deceitful warning of predators, anti-predator coloration may have other post-detection functions. These include patterns that are used for intimidation of predators, such as patterns imitating a pair of relatively large eyes that are usually suddenly revealed only when a predator is at very close distance (e.g. Vallin *et al.* 2005). In addition, it has been suggested that some markings, such as lateral wing markings of some butterflies, may function to direct an attack to a less vulnerable area of the body and increase the chance of the prey to escape and survive an attack (Wourms & Wasserman 1985; Ruxton *et al.* 2004).

As in the case of camouflage, also in the case of anti-predator signals, there are plenty of unanswered questions about the optimization of the appearance of the signal (or prey coloration) for the cognitive system

of predators to maximize the desired response, and how this optimization is affected by factors such as presence and level of a secondary defence, the appearance of the background, the appearance of other prey types and their levels of secondary defences, or the presence of multiple predator species. For example, although aposematic coloration often appears conspicuous, it is unclear whether this is to make the signal stand out against the background, to increase the strength of the signal, or to make the prey distinct from other prey with weaker secondary defence (e.g. Gittleman & Harvey 1980; Sherratt & Beatty 2003). In addition, little is known about how a combination of different anti-predator functions affects the optimization of prey coloration appearance. For example, it has been suggested that it may be possible to combine the apparently conflicting functions of aposematic signalling and crypsis in a distance-dependent fashion (e.g. Tullberg *et al.* 2005), but how common this strategy is and which circumstances promote it have yet to be explored.

Generally, most questions about the evolution of anti-predation coloration deal with optimization of the appearance of coloration of a given prey in relation to its environment. Prey coloration is optimized to the predators, particularly to their perception, with respect to the predators' information about that and other potential prey items, and their ability to detect, distinguish and remember that prey. Basically, natural selection on anti-predation coloration can be thought of as adjustment of visual information (honest or deceitful) emitted by a prey and received by its predators in an attempt to control the response of the predators.

When modelling approaches for the study of prey coloration are considered, three points that are essential can be concluded from the review above. First, the coloration of a prey is a pattern, and important information may be lost if it is reduced to a simple variable such as 'conspicuousness'. Second, natural selection on anti-predator coloration is caused by processing of visual information by the predators. This also means that the abilities as well as the biases and limitation of the cognitive systems (vision and information processing) of predators have a central role in the natural selection on prey coloration (cf. Endler 1992; Guilford 1992; Arak & Enquist 1993; Dukas 1998; Enquist & Arak 1998). Third, owing to the multiple factors (various prey types, their appearances, qualities as prey and abundances) at play, evolution of prey coloration may often involve complex interactions.

3. NEURAL NETWORKS AND SIMULATION OF EVOLUTION OF ANTI-PREDATOR COLORATION

In behavioural and evolutionary ecology, artificial neural networks have been used to simulate a predator in studies about anti-predation coloration and a receiver in studies about signalling between and within species. The most often used network design in these studies has been a feed-forward network with three layers (e.g. Arak & Enquist 1993; Enquist & Arak 1993, 1994, 1998; Holmgren & Enquist 1999; Merilaita 2003; Kenward *et al.* 2004). In a feed-forward network,

the signals traverse in one direction only, from the input layer towards the output layer. Each layer consists of a set of neurons which are connected to neurons on the adjacent layers. A connection-specific 'weight' value is associated to each connection. Each neuron consists of a transfer (activation) function and a neuron-specific 'bias' value. When input data (usually values between -1 and 1) are projected on the input layer, signals will pass through the connections and neurons and reach the output layer. Since an artificial neural network consists of a network of connections, a single neuron typically receives signals from multiple input cells or neurons and may also forward its signal to more than one neuron. Signals coming to a neuron are first multiplied by the weights of the connections and the products are then summed together with the bias value. The resulting sum is then fed in the transfer function and the output of the function is used as an input to the connections leading to the neurons on the next layer. Threshold functions are usually used as transfer functions.

In order to achieve the desired output values for given input values, a neural network has to be trained. Thus, training in this context means adjustment of weight and bias values of a network to produce mapping between input data and output values. There are various training algorithms based on the back-propagation algorithm. Back propagation is based on series of iterative calculations to decrease the errors in network weights. During training the errors are first propagated backwards through the network in order to evaluate the derivatives of error function with respect to the weights, and these derivatives are then used to compute the adjustment to be made to the weights (Bishop 1995; Haykin 1999). This also means that the back-propagation algorithm requires the transfer functions to be differentiable.

Usually, the training dataset only includes a subset of all possible values of input data. Yet, an appropriately trained network can be expected to produce correct outputs even for input data not included in the training dataset, a feature called generalization. This ability to generalize is interesting because it is one of the features that artificial neural networks share with biological neural networks (Enquist & Arak 1998; Ghirlanda & Enquist 1998).

Another way to train a network is through evolution (e.g. Enquist & Arak 1993, 1994; Holmgren & Enquist 1999; Kenward *et al.* 2004). This means that in a population of networks, random variation in weight values are produced, and the best performing networks (judged by a fitness function) will be used to form the next generation of networks. Thus, it is comparable to an evolutionary process with genes coding for the weight values. In general, especially when used in disciplines other than evolutionary biology, such a method of searching for optimal solutions in a way that resembles biological evolution is often referred to as the genetic algorithm (e.g. Mitchell 1996). Neural network training through evolution may locate the neighbourhood of an optimal solution quicker than back-propagation methods due to its global search strategy, but once in the neighbourhood of the optimal solution it tends to converge to the optimal solution slower than

back-propagation methods because its convergence is controlled by mutation operations (Tsoukalas & Uhrig 1997). Typically, models with evolving prey that have used training through evolution have produced slower responses in relation to prey evolution than do models that have used a back-propagation algorithm. But evidently, control over the rate of training in relation to prey evolution is fully in the hands of the researcher. Anyway, the choice of training method may depend on whether the training simulates processes corresponding to evolutionary or learned change in response, or how desirable stochastic events are in the training process. Training that is based on a back-propagation algorithm and is repeated each prey generation with an updated training dataset corresponds better to a change in a predator's response through learning than does training through evolution with the same generation time as the prey.

Input data for an artificial neural network can be described as points in a multidimensional hyperspace, the points representing colour patterns of one or more species of prey or samples of the visual background. The number of cells in the input layer corresponds to the number of the dimensions. Most studies relevant in the context of prey coloration have used a feed-forward neural network with threshold functions as transfer functions and one output neuron. Such networks can categorize the input data into two classes, such as 'detected' and 'not detected' in the case of cryptic coloration or 'attacked' and 'not attacked' in the case of warning coloration. Such classification can be compared with drawing a 'decision boundary' that separates input data points to the different categories. Thus, training a neural network adjusts the decision boundary. Insufficient training will result in inaccuracy, whereas too much training will result in 'over-fitting' of the decision boundary and decreased generalization (i.e. the ability to correctly classify input data not included in the training dataset). Importantly, the optimal network design (i.e. the number of layers and the number of neurons in the layers as well as how the neurons are connected to each other) depends on the complexity of the decision boundary and the dimensionality of the input vector.

Conversely, neural network performance depends on both the network design and training. This is in many ways analogous to fitting a polynomial curve to a dataset (e.g. Bishop 1995). Such a nonlinear curve not only describes the data, but it can also be used to predict output values for input values not included in the dataset. However, to obtain a curve that satisfactorily fits the data, one has to find the right number of terms in the polynomial as well as the right parameter values for the terms. Similarly, an artificial neural network satisfactorily performs a given mapping task (such as a categorization task) if it fits the training dataset and, moreover, can be used to predict output values for input values not included in the training data. To obtain such network performance, one has to choose an appropriate number of layers and neurons (network design) as well as find appropriate weight values for the connections between the neurons (training).

The design and training of a network have to be optimized for the modelling task in question. A

network that does not have enough neurons and layers, or has not been trained well enough, will not produce correct mapping between the input and output values in the training data. On the other hand, if the network design is too complex or the network has been trained too much, there is a risk that the network will over-fit the training data: the network will represent well the specific aspects of the training data (including any noise) at expense of representing poorly the more general or systematic pattern (e.g. Bishop 1995). In other words, the generalization ability of the network will be poor. Consequently, the aim usually is to use the simplest design and shortest training that are sufficient for the task at hand. For small networks used for relatively simple tasks, it is possible to gain an idea of the optimal architecture by deduction. Generally, information about appropriateness of network design and training can be gained empirically by investigating network performance (which can be measured as the mean of squared errors between observed and desired outputs) and how it varies when network design or number of iterations of the training algorithm used is changed. Note that if another dataset (called the validation or test dataset) is used instead of the training dataset for measuring performance, then the measures of performance are less likely to be confounded by over-fitting, a method called cross-validation (Bishop 1995; Haykin 1999). More detailed and deeper discussion about various methods for optimization of network design and training can be found elsewhere (e.g. Bishop 1995; Haykin 1999). However, when it comes to optimization of design and training of neural networks in evolutionary models, one may have to consider some additional points that are typically not relevant for static non-biological applications. In practice, the evolutionary process may affect training. For example, input data for the network are likely to change with proceeding evolution, and therefore the training conditions may also change. Furthermore, it is important to bear in mind that the choice of design and training constitute part of the assumptions of the model. Maximized or unconstrained predator performance may not always be the most realistic assumption, and constraints in information processing may have evolutionary importance (see §1).

Evolutionary optimization of prey coloration can be studied by letting prey coloration evolve through selection imposed by the neural network. This can be done by simulating natural selection on a population of prey with genetically coded coloration (cf. genetic algorithm; e.g. Mitchell 1996). The aim of this approach is to find optimal or good solutions as well as to produce adaptive responses (i.e. find new solutions) if the conditions are changing. Here, the prey population corresponds to a set of candidate solutions. Prey coloration is described by chromosomes consisting of sets of alleles. It has been shown that a relatively small population size of about 30 individuals (20–100) usually works well, but this also depends on the rate of mutation and recombination events (Mitchell 1996). The prey population is subjected to selection by the neural network. Thus, the output of the neural network for a given coloration phenotype is used to determine the predation risk, which in turn can be

used to determine the fitness of an individual with that phenotype. A fraction of the population with lowest fitness is removed and replaced by the offspring of the remaining individuals. To produce genetic variation that is necessary for evolution, some mutations and recombination events take place when the prey reproduces. Usually, the prey species can be assumed to be a haploid hermaphrodite if the study focuses on optimization of coloration and not on the effects of specific genetic mechanisms. In a system that combines a continually trained neural network and a virtual prey species, all the parties adapt to the responses of the other parties, resulting in evolution through counter-adaptive steps (and possibly co-adaptive steps among different prey species). Obviously, simulations like this have to be replicated owing to the stochasticity involved in the system.

So far, there are only few studies that have employed neural networks to explicitly study prey coloration. In contrast, there are several neural network studies concerning the evolution of signalling. Some of these have studied questions that are relevant in the context of anti-predator signals. For example, Enquist & Arak (1993, 1994, 1998) studied the effect of biases in receivers' recognition mechanism on selection for different aspects of signal form, such as signal symmetry or amplification of a signal. In a recent study, Kenward *et al.* (2004) investigated the evolution of repetitive patterns in visual signals.

Holmgren & Enquist (1999) used a neural network model to explicitly study prey coloration. They simulated a Batesian mimicry system using neural networks as predators and studied the evolutionary dynamics between the model and the mimic. Their study suggests that the appearance of the model and the mimic are constantly changing, and monotonically increasing response gradients cause the appearance of the model and the mimic to change in the same direction. Merilaita (2003) studied the effect of visual complexity of habitat on natural selection on cryptic coloration, and Merilaita & Tullberg (2005) studied the evolutionary choice of defence strategy between crypsis and aposematism. I will describe the two latter studies in more detail in §4.

4. EXAMPLES OF STUDIES

(a) *Prey camouflage and visual background complexity*

Merilaita (2003) used a three-layer feed-forward network to study the effect of visual background complexity on the evolution of cryptic coloration. Somewhat surprisingly, this appears to be the only neural network study which focuses on the evolution of cryptic coloration so far. In this study the predator's processing of visual information was assumed to be based on visual samples (vectors with eight cells) of its environment, which it classified either as background or prey. In the evolving prey, this resulted in selection for coloration that the predator would incorrectly classify as background.

In the model, two factors were varied. The first factor was the level of visual complexity of the background, meaning the number of different visual

elements that occurred in the habitat. Here, visual elements were defined as the components of colour patterns and can be thought of as colours, or features such as stripes or spots. In this study, visual complexity refers to variation at a scale smaller than the size of the visual samples and it should not be confused with large scale heterogeneity, such as differences between different microhabitats. The complex habitat consisted of four different visual elements and the simple habitat consisted of three visual elements. The eight-cell samples of the backgrounds were created using simple rules. First, one of the visual elements was chosen as the 'basal element'. A sample of the complex habitat always contained two cells of the basal element and two cells of each of three remaining visual elements. In the simple habitat, the samples always consisted of two cells of the basal element and three cells of each of the other two elements. Since the order of the cells was not determined, there were 2520 different possible samples of the complex habitat, but only 560 different possible samples of the simple habitat.

The second factor that varied was the presence of a constraint in the colour pattern evolution of the prey. The prey was able to produce either all the visual elements found in the habitat (unconstrained evolution) or to produce all except one of the elements found in the habitat (constrained evolution). The initial prey coloration in each simulation run was constituted by eight cells of the basal element.

The feed-forward network had eight input cells, five neurons in the hidden layer and one output neuron. Logistic sigmoid (i.e. smooth threshold) functions were used as transfer functions. The task of the neural network was to categorize a visual sample either as prey or as background. The network outputs varied from zero to unity, and the correct output for prey was unity and for background it was zero. A back-propagation algorithm was used in the training. Before prey evolution, the network was first trained to categorize between the initial prey coloration and the background until the mean square error was decreased to 2×10^{-6} , a small value that was achieved in every run of the simulation. Then, while the prey was evolving, increment training was used, such that the network was presented with a training dataset once every prey generation, and the weights and biases of the network were adjusted after each presentation of a vector in the training set. The training set consisted of a sub-sample of the prey population and an equally large sample of randomly chosen background samples. This training procedure was used because it was important to ensure that neural network performance was adjusted at the same rate in all the four combinations of habitat complexity (simple or complex) and type of evolution (unconstrained or constrained).

Fitness of an individual prey was determined by the output of the neural network for that colour pattern phenotype, such that the prey for which the output deviated most from unity had the highest fitness. This caused prey coloration to evolve towards increased crypsis. As a consequence, the prey became less susceptible to predation (i.e. network outputs for prey decreased) and its resemblance to background increased during the course of evolution in all four

Figure 1. Predation susceptibility of the prey per generation, measured as the mean output of the neural network for the most cryptic 25% of the prey population. (a) Unconstrained evolution in the complex habitat; (b) constrained evolution in the simple habitat; (c) unconstrained evolution in the complex habitat and (d) unconstrained evolution in the simple habitat. A simulation run ($N=500$ in each category) was stopped when the predation susceptibility threshold value of 10^{-3} was reached. In (d) 34% of the runs of the simulation did not reach the threshold in 200 generations. One dot may represent multiple points. From Merilaita (2003) with permission.

combinations of habitat complexity (simple or complex) and type of evolution (unconstrained or constrained). The prey was allowed to evolve until either the prey susceptibility to predation was below a threshold value (i.e. the mean output for the most cryptic 25% of the prey population was lower than 0.001) or 200 prey generations had passed.

The main result from this study suggests that it was more problematic for the prey to evolve efficient crypsis in the simple habitat than in the complex habitat (figure 1). Accordingly, in the complex habitat, the detectability of the prey decreased below a threshold value in every run of the simulation both under unconstrained and constrained evolution of coloration. In contrast, in the simple habitat, this was the case only under unconstrained evolution, whereas under constrained evolution the prey was unable to reach the threshold value in 34% of the simulation runs. In other words, the evolutionary constraint had a much more severe impact on predation susceptibility in the simple than in the complex habitat. This suggests that for a prey colour pattern to achieve a given probability of escaping detection (i.e. a given level of crypsis), more will be required from it, for example in terms of background matching, in a visually simple than in a visually complex habitat.

This result has several implications for the study of prey coloration. For one thing, it suggests that estimates of degree of crypsis based on similarity between colour pattern and habitat of prey (e.g. Endler 1984) are not comparable between habitats that differ in complexity. Moreover, the result suggests that prey should make use of constraints in predators' processing of visual information. This is interesting, because

the importance of predators' sensitivity to different wave-lengths of light (colour vision) has received more attention, whereas the role of processing of visual information after it has passed the retina has not received much attention. The reason for this bias may be that more is known about colour vision than about processing of visual information after it has passed the retina. Hence, this bias should encourage the application of neural network models in behavioural and evolutionary ecology, as they can be used to study the evolutionary consequences of information processing, such as processing of visual information by predators.

The rationale behind the study is based on the simple fact that the brain has a limited capacity to process visual information (e.g. Dukas 1998), a limitation that is demonstrated by the trade-off between search rate and detection rate in predators (Gendron & Staddon 1983; Gendron 1986). Therefore, if the processing capacity (determined by neural network design and training) is constant, an increase in visual complexity (diversity of visual information) makes the detection task more difficult and a prey, even with constrained crypsis, will be less likely to become detected. In spite of its simple rationale, this model allowed conclusions that were novel in the study of cryptic coloration. There were at least two benefits from the use of an artificial neural network in this study. First, it enabled a feasible way to study selection on a 'colour' pattern and to draw conclusions about crypsis. With another modelling technique, it would be more problematic to study these two features separately without confusing them. Second, the similarity in information processing between predators and artificial neural networks, in this case the constrained processing capacity and its consequences, was valuable.

(b) *Troubles of crypsis and aposematism and the choice of defence strategy*

Several previous studies on aposematism have emphasized the difficulty aposematism has to evolve. An aposematic prey type has to be common enough for the predator efficiently to learn to associate its warning coloration with the secondary defence that the aposematic prey carries. Therefore, it has been argued that the evolution of aposematism is paradoxical, as a rare aposematic mutant with a conspicuous aposematic coloration is unlikely to invade a cryptic prey population. For some reason, this argument has been based on a more or less explicit assumption that the prey population is initially cryptically coloured and, furthermore, that this crypsis is highly efficient. However, there is no reason to expect crypsis to be an unproblematic adaptation either. As already touched upon, a general problem for cryptic prey is that visual heterogeneity of environment often makes it difficult to achieve, through an invariable coloration, a low risk of detection in every part of the habitat the prey uses (Edmunds 1974; Merilaita *et al.* 1999, 2001; Ruxton *et al.* 2004). Similarly, prey mobility may constraint crypsis because mobility as such facilitates detection (Cott 1940; Edmunds 1974; Ruxton *et al.* 2004). Thus, it is reasonable to assume that constraints on crypsis are rather common.

Merilaita & Tullberg (2005) addressed these questions by studying the evolutionary choice of an optimal defence strategy, particularly between crypsis and anti-predator signalling. The first part of the study consisted of an evolutionary simulation model in which crypsis was either constrained or not. More specifically, the prey either lived in a visually homogeneous habitat that allowed the evolution of a high degree of crypsis, or in a heterogeneous habitat consisting of two equally common but visually very different microhabitats which imposed a constraint on the evolution of crypsis. There were two species of prey, prey 1 and prey 2, which differed only in one respect. Prey 1 was always edible, whereas prey 2 might become inedible. Thus, initially prey 2 was edible, but due to mutation the secondary defence was likely to arise and spread in the population sooner or later. Both the species had evolving colour pattern. In the beginning of each simulation run, both the prey species had the same randomly chosen colour pattern. Thus, the study concentrated on optimization of prey colour pattern and choice of defence strategy when a constraint on crypsis caused by habitat heterogeneity was either present or absent. The prey was considered to have reached a cryptic optimum if the prey colour pattern matched the homogeneous habitat or matched either of the microhabitats in the heterogeneous habitat. The prey was considered to have reached an aposematic optimum if the prey was inedible, it had evolved a colour pattern that deviated from the habitat (or the microhabitats), and this colour pattern was not invaded by another during 30 successive prey generations.

Prey colour patterns and samples of the habitats were described by four-cell vectors. Each cell of a prey colour pattern vector was occupied by a coloration element denoted by 1, 2 or 3. The homogeneous habitat was uniform, consisting of one type of element only. This element was randomly chosen at the beginning of each simulation run to be either 1 or 3. In the heterogeneous habitat, one of the microhabitats consisted uniformly of the element 1 and the other microhabitat uniformly of the element 3. The prey used both the microhabitats with equal probabilities in the heterogeneous habitat. Since habitat heterogeneity was assumed to constrain crypsis, the model parameters were chosen so that coloration adapted to one of the microhabitats would yield higher crypsis than any coloration that compromised the requirements of both the microhabitats.

In this model, prey fitness was assumed to depend on the following factors. Prey crypsis increased, and thus the probability of becoming detected by predators decreased with increasing resemblance between prey colour pattern and the background. In the heterogeneous habitat, the two microhabitats were equally common and thus prey crypsis there was given by the average crypsis in the two microhabitats. Prey 2 could also benefit from aposematism. Accordingly, an increase in the proportion of inedible individuals among individuals with a given colour pattern decreased the risk of becoming attacked by a predator for all individuals with that colour pattern. Thus, for prey 1, fitness was given by the probability of avoiding detection and for prey 2 fitness was determined by that

of the two strategies (crypsis or aposematism) that yielded the better protection against predation.

To simulate predation according to the assumptions of the model, two different neural networks were used. The first network was a radial basis function network. The task of the radial basis function network was to give the probability of detection for each prey colour pattern, based on the assumption that probability of detection is determined solely by the resemblance between the background and prey colour pattern. Radial basis function networks are used in pattern recognition and they consist of an input layer, a hidden layer and an output layer. The output of the transfer function of a radial basis function, which typically is a Gaussian function, is determined by the Euclidian distance between the input vector and a template vector (e.g., Bishop 1995; Tsoukalas & Uhrig 1997; Haykin 1999; Theodoridis & Koutroumbas 1999). As the name implies, such a radial basis function produces radially symmetric activations that decrease from unity to zero with increasing distance from the function's centre (the template vector). Merilaita & Tullberg (2005) applied a very simplistic variant of such a network. It had one radial basis unit only and a vector representing the background was used as the template vector. Consequently, the network transferred the amount of difference between prey colour pattern and background to hypothetical detection probability values.

A number of experimental studies on aposematism demonstrate a tendency of predators to learn to avoid prey that they experience as unpleasant or dangerous by its appearance (Ruxton *et al.* 2004). The second neural network in Merilaita & Tullberg (2005) was used to simulate such an avoidance response. This was a feed-forward network with four input cells, eight neurons in a first hidden layer, four cells in a second hidden layer and one output neuron. The assumption that an increase in the proportion of inedible individuals with a given coloration increased the avoidance response and decreased the probability of attack towards that colour pattern requires that the output of the second neural network had to vary continuously from zero to unity, instead of representing a binomial decision. Therefore, a typical feed-forward network with one hidden layer and one output neuron with a threshold function as a transfer function was not suitable for the task. Instead, the output neuron of such a network was replaced by the second hidden layer and an output layer with one neuron that exceptionally had a linear transfer function (all the other transfer functions were smooth threshold functions). Thus, instead of a single output neuron assigning the summed signals from the first hidden layer in either of two classes, the multiple neurons on the second hidden layer enabled several different categorical outputs. These outputs were then fed into the output neuron which, with the help of the linear transfer function, transformed the signals to continuous values of attack probability. Consequently, training of the network created a relationship between colour pattern phenotypes and attack probabilities. The training dataset included the colour pattern of each prey individual (prey 1 and 2) and the proportions of inedible individuals within each colour pattern phenotype

were used as training target values. The training was based on a back-propagation algorithm with an adaptive learning rate (e.g. Demuth & Beale 2000) and the network was trained during each prey 2 generation until the mean square error of the output was less than 0.001 or until the training dataset had been presented 150 times.

As might be expected, prey 1, which could not become inedible, always evolved to a cryptic optimum. When evolution of camouflage was not constrained by habitat heterogeneity, prey 1 evolved a colour pattern that matched the background in each of the 1000 replicates of the simulation. In addition, in all the 1000 replicates of the simulation in the heterogeneous habitat, prey 1 evolved coloration matching either of the microhabitats (in 22 of the runs, the population was polymorphic and consisted of individuals with either of the background matching colorations).

Owing to the favourable conditions (e.g. high mutation rate for the secondary defence), prey 2 evolved aposematism a number of times in both the habitats. However, the main result from the study was that in the homogeneous habitat prey 2 evolved aposematism only 543 times out of 1000 compared with 915 times in the heterogeneous habitat. This suggests that constraints on camouflage (in this case due to habitat heterogeneity) favour the evolution of aposematism. In the second part of the study, Merilaita & Tullberg (2005) backed up this conclusion with empirical evidence. The evidence comes from a comparison of the commonness of crypsis and anti-predator signalling among day-active lepidopteran taxa in which camouflage is constrained by mobility, and in night-active taxa which rest during the day and therefore are not mobile when susceptible to visual predation. The comparison showed a significant association between day activity and anti-predator signalling. Merilaita & Tullberg (2005) concluded that while focusing on the costs of aposematic or mimetic coloration, previous studies on the evolution of anti-predator signals have neglected the possible costs of constraints for the alternative strategy, crypsis, and that this may provide an explanation for why the evolution of anti-predator signals has been argued to be paradoxical.

5. CONCLUSIONS

As the brief review in §2 on prey coloration as an anti-predator adaptation indicated, there are some features which are essential when considering a modelling approach for the study of prey coloration. First, in some cases, it is possible to use models in which colour patterns are represented by simplistic variables such as 'conspicuousness'. A simple model is preferable if it meets the requirements set by the question posed. However, natural selection acts on phenotypic traits such as colour pattern, instead of (directly) acting on conceptual traits such as predation risk or conspicuousness. Thus, although simplistic variables are suitable for modelling some questions, in other cases, additional or more correct insights can be obtained by using a more realistic representation of colour patterns. Therefore, it may often be important to capture the

multi-dimensional nature of colour patterns as variables. It is evident that artificial neural networks are well-suited for processing such variables, as for example their use in pattern recognition indicates (Bishop 1995; Theodoridis & Koutroumbas 1999). Moreover, neural networks provide a relatively simple method for dealing with patterns.

The second essential feature is the central role of visual information processing by predators in natural selection for prey coloration. Therefore, the capacity to reproduce a desired aspect of information processing by a predator or to imitate a desired response of a predator to visual information with the required level of biological reality is important when one considers modelling approaches for the study of prey coloration. Neural networks are not just capable of information processing and responding to patterns; the structural and functional similarities between artificial neural networks and biological neural systems may be helpful when the processing of visual information by predators is modelled. This is because these similarities may enable artificial neural networks to be used to model information processing and decision making in a biologically plausible fashion (e.g. Enquist & Arak 1998; Ghirlanda & Enquist 1998; Enquist *et al.* 2002).

The neural networks used in the studies of prey coloration and signalling have been rather simple. However, complexity as such is not a goal in modelling. On the contrary, simple models are generally easier to interpret and less likely to suffer from confounding, irrelevant factors or effects. Accordingly, when it comes to neural network models, comprehending the behaviour of a model, which is important, is easier if the model is simple than if it is complex. Thus, simplicity ought to be preferred as far as it does not constrain the function of the model. In addition, it is important to realize here that a neural network model is used for reproducing a given aspect of a predator's information processing or response to visual information. This means that it is wrong to interpret the aim to be something more, such as representing the whole visual system or the brain. For example, in the first neural network model (Merilaita 2003) studying crypsis and background complexity, described above, the interesting aspect of predator behaviour was the limited capacity of information processing and its effect on a predator's ability to detect prey, and eventually on natural selection for crypsis. In the study about evolutionary choice of defence strategy between aposematism and crypsis (Merilaita & Tullberg 2005), the behavioural outcome (i.e. response to prey coloration) was more important than the process by which it was achieved.

The third essential feature in modelling prey coloration is that it may involve dynamic and complex interactions. Neural network studies of prey coloration have used simulation of evolution to study prey coloration (Holmgren & Enquist 1999; Merilaita 2003; Merilaita & Tullberg 2005). Evolutionary simulations are well suited for studying adaptation and optimal phenotypes under dynamic and complex interactions, such as successive counter-adaptive or co-adaptive responses over multiple evolutionary steps, or cases where fitness of a phenotype varies due to

multiple dynamic factors involved (see also Peck 2004). For example, in the study about evolutionary choice of defence strategy between aposematism and crypsis (Merilaita & Tullberg 2005), predation risk of a prey was affected by the appearance of the prey in relation to the background, the characteristics of the background, the presence or absence of the secondary defence, its appearance in relation to other prey, and the abundance of prey with similar coloration phenotype. In addition, in the study by Holmgren & Enquist (1999) the focus was on the dynamics of selection for appearance in Batesian mimicry systems. Generally, artificial neural networks are well suited as components of evolutionary simulation models because they can produce adaptive responses through training and they can respond to novel stimuli by generalizing information from familiar stimuli.

An artificial neural network may exhibit various interesting behaviours. However, it is of paramount importance to bear in mind that, although there are similarities between artificial neural networks and biological neural systems, not all neural network behaviours are biologically relevant. In other words, these general similarities seldom suffice to validate a specific model. As for modelling in general, also in the case of neural network models, empirical data are needed to confirm the validity of the assumptions or the results of a model. In the study about crypsis and background complexity (Merilaita 2003), the model was based on a known fact that the brain has a limited capacity to process visual information simultaneously (Dukas 1998). Further, psychological experiments using humans as subjects show that the difficulty of a visual detection task increases with increasing complexity of background patterns and colours, lending empirical support for the results of that study (Gordon 1968; Farmer & Taylor 1980). In the study about evolutionary choice of defence strategy between aposematism and crypsis (Merilaita & Tullberg 2005), neural networks were used to imitate the decreasing probability of detection by predators with increasing resemblance between prey and its background and the ability of predators to learn to avoid defended prey by their appearance (Ruxton *et al.* 2004). Further, empirical support for the result of the model was provided by the phylogenetic comparison in the second part of the study (Merilaita & Tullberg 2005).

To summarize, a capacity to deal with patterns as well as to reproduce aspects of biological information processing or behavioural responses to visual information are primary qualities of a modelling approach well suited for the study of prey coloration. In addition, capability to deal with dynamic interactions is an advantage. Neural network models satisfy these requirements and may therefore provide an appropriate tool for many questions in the study of prey coloration. Although there are so far only few studies that have applied neural networks to study prey coloration, the studies described above show that neural network models can help us to gain novel insights and promote the understanding of the specific appearances of prey coloration. However, when a neural network model is applied, it is important to biologically validate the behaviour of the model.

I thank Magnus Enquist, Charlotta Kvarnemo and Ian Henshaw for their valuable comments on the manuscript. This study was supported by the Swedish Research Council.

REFERENCES

- Arak, A. & Enquist, M. 1993 Hidden preferences and the evolution of signals. *Phil. Trans. R. Soc. B* **265**, 1059–1064.
- Bishop, C. M. 1995 *Neural networks for pattern recognition*. Oxford, UK: Oxford University Press.
- Cott, H. B. 1940 *Adaptive coloration in animals*. London, UK: Methuen.
- Cuthill, I. C., Stevens, M., Sheppard, J., Maddocks, T., Párraga, C. A. & Troscianko, T. S. 2005 Disruptive coloration and background pattern matching. *Nature* **434**, 72–74. (doi:10.1038/nature03312)
- Demuth, H. & Beale, M. 2000 *Neural network toolbox for use with MATLAB, version 4*. Natick, MA: The MathWorks, Inc.
- Dukas, R. 1998 Constraints on information processing and their effects on behavior. In *Cognitive ecology: the evolutionary ecology of information processing and decision making* (ed. R. Dukas), pp. 89–128. Chicago, IL: University of Chicago Press.
- Edmunds, M. 1974 *Defence in animals*. Harlow, UK: Longman.
- Endler, J. A. 1978 A predator's view of animal color patterns. *Evol. Biol.* **11**, 319–364.
- Endler, J. A. 1984 Progressive background matching in moths, and a quantitative measure of crypsis. *Biol. J. Linn. Soc.* **22**, 187–231.
- Endler, J. A. 1992 Signals, signal conditions and the direction of evolution. *Am. Nat.* **139**, S125–S153. (doi:10.1086/285308)
- Enquist, M. & Arak, A. 1993 Selection of exaggerated male traits by female aesthetic senses. *Nature* **361**, 446–448. (doi:10.1038/361446a0)
- Enquist, M. & Arak, A. 1994 Symmetry, beauty and evolution. *Nature* **372**, 169–172. (doi:10.1038/372169a0)
- Enquist, M. & Arak, A. 1998 Neural representation and the evolution of signal form. In *Cognitive ecology: the evolutionary ecology of information processing and decision making* (ed. R. Dukas), pp. 21–87. Chicago, IL: University of Chicago Press.
- Enquist, M., Arak, A., Ghirlanda, S. & Wachtmeister, C.-A. 2002 Spectacular phenomena and limits to rationality in genetic and cultural evolution. *Phil. Trans. R. Soc. B* **357**, 1585–1594. (doi:10.1098/rstb.2002.1067)
- Farmer, E. W. & Taylor, R. M. 1980 Visual search through color displays: effects of target-background similarity and background uniformity. *Percept. Psychophys.* **27**, 267–272.
- Gendron, R. P. 1986 Searching for cryptic prey: evidence for optimal search rates and the formation of search images in quail. *Anim. Behav.* **34**, 898–912. (doi:10.1016/S0003-3472(86)80076-8)
- Gendron, R. P. & Staddon, J. E. R. 1983 Searching for cryptic prey: the effect of search rate. *Am. Nat.* **121**, 172–186. (doi:10.1086/284049)
- Ghirlanda, S. & Enquist, M. 1998 Artificial neural networks as models of stimulus control. *Anim. Behav.* **56**, 1383–1389. (doi:10.1006/anbe.1998.0903)
- Gittleman, J. L. & Harvey, P. H. 1980 Why are distasteful prey not cryptic? *Nature* **286**, 149–150. (doi:10.1038/286149a0)
- Gordon, I. E. 1968 Interactions between items in visual search. *J. Exp. Psychol.* **76**, 248–355. (doi:10.1037/h0025482)
- Guilford, T. 1992 Predator psychology and the evolution of prey coloration. In *Natural enemies: the population biology of predators, parasites and diseases* (ed. M. J. Crawley), pp. 377–394. Oxford, UK: Blackwell.
- Haykin, S. 1999 *Neural networks: a comprehensive foundation*, 2nd edn. Upper Saddle River, NJ: Prentice-Hall.

- Holmgren, N. M. A. & Enquist, M. 1999 Dynamics of mimicry evolution. *Biol. J. Linn. Soc.* **66**, 145–158. (doi:10.1006/bijl.1998.0269)
- Kenward, B., Wachtmeister, C.-A., Ghirlanda, S. & Enquist, M. 2004 Spots and stripes: the evolution of repetition in visual signal form. *J. Theor. Biol.* **230**, 407–419. (doi:10.1016/j.jtbi.2004.06.008)
- Lindström, L., Alatalo, R. V., Lyytinen, A. & Mappes, J. 2001 Strong antiapostatic selection against novel rare aposematic prey. *Proc. Natl Acad. Sci. USA* **98**, 9181–9184. (doi:10.1073/pnas.161071598)
- Merilaita, S. 1998 Crypsis through disruptive coloration in an isopod. *Proc. R. Soc. B* **265**, 1059–1064. (doi:10.1098/rspb.1998.0399)
- Merilaita, S. 2003 Visual background complexity facilitates the evolution of camouflage. *Evolution* **57**, 1248–1254. (doi:10.1554/03-011)
- Merilaita, S. & Lind, J. 2005 Background-matching and disruptive coloration, and the evolution of cryptic coloration. *Proc. R. Soc. B* **272**, 665–670. (doi:10.1098/rspb.2004.3000)
- Merilaita, S. & Lind, J. 2006 Great tits (*Parus major*) searching for artificial prey: implications for cryptic coloration and symmetry. *Behav. Ecol.* **17**, 84–87. (doi:10.1093/beheco/arj007)
- Merilaita, S. & Tullberg, B. S. 2005 Constrained camouflage facilitates the evolution of conspicuous warning coloration. *Evolution* **59**, 38–45. (doi:10.1554/04-410)
- Merilaita, S., Tuomi, J. & Jormalainen, V. 1999 Optimisation of cryptic coloration in heterogeneous habitats. *Biol. J. Linn. Soc.* **67**, 151–161. (doi:10.1006/bijl.1998.0298)
- Merilaita, S., Lyytinen, A. & Mappes, J. 2001 Selection for cryptic coloration in a visually heterogeneous habitat. *Proc. R. Soc. B* **268**, 1925–1929. (doi:10.1098/rspb.2001.1747)
- Mitchell, M. 1996 *An introduction to genetic algorithms*. Cambridge, MA: MIT Press.
- Peck, S. L. 2004 Simulation as experiment: a philosophical reassessment for biological modeling. *Trends Ecol. Evol.* **19**, 530–534. (doi:10.1016/j.tree.2004.07.019)
- Ruxton, G. D., Sherratt, T. M. & Speed, M. P. 2004 *Avoiding attack: the evolutionary ecology of crypsis, warning signals and mimicry*. Oxford, UK: Oxford University Press.
- Sherratt, T. N. & Beatty, C. D. 2003 The evolution of warning signals as reliable indicators of prey defense. *Am. Nat.* **162**, 377–389. (doi:10.1086/378047)
- Skelhorn, J. & Rowe, C. 2005 Tasting the difference: do multiple defence chemicals interact in Müllerian mimicry? *Proc. R. Soc. B* **272**, 339–345. (doi:10.1098/rspb.2004.2953)
- Speed, M. P. 1993 Müllerian mimicry and the psychology of predation. *Anim. Behav.* **45**, 571–580. (doi:10.1006/anbe.1993.1067)
- Theodoridis, S. & Koutroubas, K. 1999 *Pattern recognition*. San Diego, CA: Academic Press.
- Tsoukalas, L. H. & Uhrig, R. E. 1997 *Fuzzy and neural approaches in engineering*. New York, NY: Wiley.
- Tullberg, B. S., Merilaita, S. & Wiklund, C. 2005 Aposematic and crypsis combined as a result of distance dependence: functional versatility of the colour pattern in the swallowtail butterfly larva. *Proc. R. Soc. B* **272**, 1315–1321. (doi:10.1098/rspb.2005.3079)
- Vallin, A., Jakobsson, S., Lind, J. & Wiklund, C. 2005 Prey survival by predator intimidation: an experimental study of peacock butterfly defence against blue tits. *Proc. R. Soc. B* **272**, 1203–1207. (doi:10.1098/rspb.2004.3034)
- Wourms, M. K. & Wasserman, F. E. 1985 Butterfly wing markings are more advantageous during handling than during initial strike of an avian predator. *Evolution* **39**, 845–851. (doi:10.2307/2408684)