Int. Journ. of Unconventional Computing, Vol. 1, pp. 00–00
Reprints available directly from the publisher
Photocopying permitted by license only

Conceptual Frameworks for Artificial Immune Systems

SUSAN STEPNEY[1], ROBERT E. SMITH[2], JONATHAN TIMMIS[3], ANDY M. TYRRELL[4], MARK J. NEAL[5] AND ANDREW N. W. HONE[6]

[1]*Dept. Computer Science, University of York, YO10 5DD, UK*

[2]*The Intelligent Computer Systems Centre, University of the West of England, Bristol, BS16 1QY, UK*

[3]*Computing Laboratory, University of Kent, Canterbury, CT2 7NF, UK*

[4]*Dept. Electronics, University of York, YO10 5DD, UK*

[5]*Dept. Computer Science, University of Wales Aberystwyth, SY23 2AX, UK*

[6]*Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, CT2 7NF, UK*

Received 16 December 2004; Accepted 4 February 2005

We propose that bio-inspired algorithms are best developed and analysed in the context of a multidisciplinary conceptual framework that provides for sophisticated biological models and well-founded analytical principles, and we outline such a framework here, in the context of Artificial Immune System (AIS) network models, and we discuss mathematical techniques for analysing the state dynamics of AIS. We further propose ways to unify several domains into a common meta-framework, in the context of AIS population models. We finally discuss a case study, and hint at the possibility of a novel instantiation of such a meta-framework, thereby allowing the building of a specific computational framework that is inspired by biology, but not restricted to any one particular biological domain.

1. INTRODUCTION

The idea of biological inspiration for computing is as old as computing itself. It is implicit in the writings of von Neumann and Turing, despite

the fact that these two fathers of computing are now more associated with the standard, distinctly non-biological computational models.

Computation is rife with bio-inspired models (neural nets, evolutionary algorithms, artificial immune systems, swarm algorithms, ant colony algorithms, L-systems, . . .). However, many of these models are naive with respect to biology. Even though these models can work extremely well, their naivety often blocks understanding, development, and analysis of the computations, as well as possible feedback into biology.

2. A CONCEPTUAL FRAMEWORK

The next steps in bio-inspired computation should be to develop more sophisticated biological models as sources of computational inspiration, and to use a conceptual framework to develop and analyse the computational metaphors and algorithms.

We propose that bio-inspired algorithms are best developed and analysed in the context of a multidisciplinary conceptual framework that provides for sophisticated biological models and well-founded analytical principles.

Figure 1 illustrates a possible structure for such a conceptual framework. Here *probes* (observations and experiments) are used to provide a (partial and noisy) view of the complex *biological system*. From this limited view, we build and validate simplifying abstract representations, *models*, of the biology. From these biological models we build and validate *analytical computational frameworks*. Validation may use

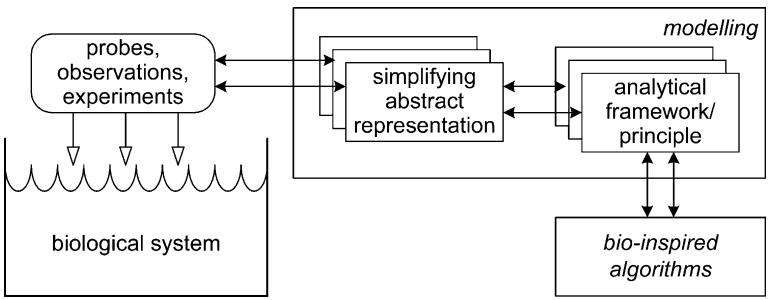

FIGURE 1
An outline conceptual framework for a bio-inspired computational domain.

mathematical analysis, benchmark problems, and engineering demonstrators. These frameworks provide principles for designing and analysing *bio-inspired algorithms* applicable to non-biological problems, possibly tailored to a range of problem domains, and contain as much or as little biological realism as appropriate. The concept flow also supports the design of algorithms specifically tailored to modelling the original biological domain, permits influencing and validating the structure of the biological models, and can help suggest ideas of further experiments to probe the biological system. This is necessarily an interdisciplinary process, requiring collaboration between (at least) biologists, mathematicians, and computer scientists to build a complete framework.

An important observation is that none of the representation and modelling steps outlined above is *unbiased*. There are many possible probes, and many possible representations of the same systems even given the same probes, and they all provide different insights. In particular, models derived specifically for the goals of biological simulation may provide insights that are distinct from those that serve computational goals.

It is very seldom that the modelling steps used in these distinct activities are examined for common properties, and comparative biases. In many instances not all of the representational steps outlined above are taken. In particular, bio-inspired computational algorithms usually proceed directly from a (naive) biological model to an algorithm, with little analytical framing of the representation's properties. Such "reasoning by metaphor" is a troubling aspect of these algorithms. Without the application of suitable analysis techniques to the simplified representations of biological systems, algorithms derived from these representations rely only on the (often weak) analogy to the biological system to support their use. We feel that it is important to recognise the distinct levels of the modelling process outlined above, to avoid naive assumptions.

One example that *can* be described in terms of such a framework, at least partially, is Holland's original adaptive system theories [26,20], founded on a simplified binary-encoded representation of genetics, and analytical principles of building blocks, k-armed bandit theories, the schema theorem, and implicit parallelism. Evolutionary computation theory has developed and deepened in the wake of this work, and it continues to influence the prescription of genetic algorithms. We propose

that other bio-inspired computational domains, including Artificial Immune Systems, should be put on a similarly sound footing.

3. INSTANTIATING THE FRAMEWORK FOR AIS

The natural immune system is a complex biological system essential for survival. It involves a variety of interacting cellular and molecular elements that control either micro- or macro-system dynamics. The effectiveness of the system is due to a set of synergetic, and sometimes competitive, internal strategies to cope with chronic and/or rare pathogenic challenges (antigens). Such strategies remodel over time as the organism develops, matures, and then ages (immuno-senescence). The strategies of the immune system are based on task distribution to obtain distributed solutions to problems (different cells are able to carry out complementary tasks) and solutions to distributed problems (similar cells carrying out the same task in a physically distributed system). Thus, cellular interactions can be envisaged as parallel and distributed processes among cells with different dynamical behaviour, and the resulting immune responses appear to be emergent properties of self-organising processes. Theories abound in immunology pertaining to how the immune system remembers antigenic encounters (maintenance of memory cells, use of immune networks), and how the immune system differentiates between self and non-self molecules (negative selection, self-assertion, danger theory).

We can explicitly exploit the conceptual framework, in order to develop, analyse and validate sophisticated novel bio-inspired computational schemes, including those inspired by complex processes within the natural immune system. This work needs to be done; here we outline a suggested route.

3.1 A first step: Interdisciplinary research

AIS is a relatively new and emerging bio-inspired area and progress has been made from naively exploiting mechanisms of the immune system. Computer security systems have been developed, anti-virus software has been created, optimisation and data mining tools have been created that are performing as well as the current state of the art in those areas.

The original AIS were developed with an interdisciplinary slant. For example, Bersini [2–4] pays clear attention to the development of immune network models, and then applies these models to a control

problem characterised by a discrete state vector in a state space $\mathbb{R}^L$. Bersini's proposal relaxes the conventional control strategies, which attempt to *drive* the process under control to a specific zone of the state space; he instead argues that the metadynamics of the immune network is akin to a meta-control whose aim is to keep the concentration of the antibodies in a certain range of viability so as to continuously preserve the identity of the system.

There are other examples of interdisciplinary work, such as the development of immune gene libraries and ultimately a bone marrow algorithm employed in AIS [25], and the development of the negative selection algorithm and the first application to computer security [16].

However, in more recent years, work on AIS has drifted away from the more biologically-appealing models and attention to biological detail, with a focus on more engineering-oriented approach. This has led to systems that are examples of "reasoning by metaphor". These include simple models of clonal selection and immune networks [12,13,60,42], and negative selection algorithms [6,22,59]. We suggest that even such an engineering-oriented approach may benefit from closer interaction with biologists, and from a more principled mechanism for the extraction, articulation, and application of the underlying computational metaphor.

Freitas & Timmis [17] outline the need to take into account the application domain when developing AIS. The conceptual framework proposal here complements that position: once we have a well-developed conceptual framework, we can specialise it for various application domains in a justifiable way.

3.2 Adopting the conceptual framework for AIS

de Castro & Timmis [10] propose a structure for engineering AIS. The basis is a *representation* to create abstract models of immune organs, cells, and molecules, together with a set of *affinity functions* to quantify the interactions of these "artificial elements", and a set of *general-purpose algorithms* to govern the dynamics of the AIS.

The structure can be modelled as a layered approach (figure 2). To build a system, one typically requires an application domain or target function. From this basis, a *representation* of the system's components is chosen. This representation is domain and problem dependent: the representation of network traffic, say, may well be different from that of a real time embedded system. The representation-specific *affinity measures* quantify the interactions of the elements of the system. There

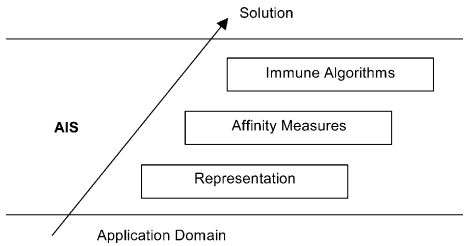

FIGURE 2
A structure for AIS, from [10].

are many possible affinity measures, such as Hamming or Euclidean distances. The final layer involves the use of *algorithms*, such as negative and positive selection, clonal selection, the bone marrow algorithm, and immune network algorithms, which govern the behaviour (dynamics) of the system. Each algorithm has its own particular range of uses.

This layered structure is not complete from the conceptual framework perspective. For example, we propose that AIS algorithms (in some cases) may benefit from asking questions such as: what is "self", or "danger". In addition, AIS algorithms in their current form can be classified as *population based* or *network based* [10]. In the following sections, we adopt this classification, and propose how one might undergo a development of an AIS algorithms adopting the conceptual framework above.

3.3 Population based AIS algorithms

Three common algorithms in AIS, those of positive, negative, and clonal selection, are all based on *populations* of agents trained to recognise certain aspects of interest (see [10] for an overview). There are similarities between the algorithms: positive and negative selection, for example, are merely two sides of the same coin. There are also differences: positive and negative selection involve essentially random generation of candidate recognisers, whilst clonal selection uses a form of reinforcement based on selection and mutation of the best recognisers.

We defer discussion of these models to the meta-frameworks of section 5, and population based models in general.

3.4 Network based AIS algorithms

Jerne's original immune network theory [30] suggests an immune system with a dynamic behaviour even in the absence of non-self

antigens. This differs from the biological basis of the clonal and negative selection algorithms, as it suggests that B-cells are capable of recognising each other.

Theoretical immunologists have been interested in creating models of immune networks in order to introduce new ways of explaining how the immune systems works [52,14]. Researchers have translated some of these ideas into the computing domain, in applications such as optimisation and control [2,4]. This work has also inspired the development of machine learning network models with applications mainly in data analysis [60,13].

However, as we stated earlier, the later work has somewhat deviated from the biological model, being adapted to a particular problem. In addition, Jerne's immune network theory itself is controversial, and not widely accepted by immunologists. This has an impact on the AIS algorithm: if the biology is not correct, then one must re-examine the algorithm to understand what is *really* going on; this would hopefully shed light on the more complex nature of the immune systems, and the networks that are clearly present therein.

The first step to achieving this understanding would be to probe the biological system from the perspective of interpreting the system as a *network* of interaction, cooperation and competition amongst molecules, cells, organs, and tissues. The results could then be used to formulate a suitable mapping between biological properties and framework components. These components could then be used as the basis for the topology and dynamics of new biological models, in addition to re-examining existing models such as [53,55]. The new models would allow a greater understanding of the operation of such systems to be developed in an artificial context.

Within the context of these new immune network models, one could examine, for example, Matzinger's danger theory [37], context of response [31,36], memory mechanisms [58], general alarm response or stress response [50], self/non-self recognition [39], and Varela's self-assertion [64]. Additionally, the constructive role of noise in biological systems, which is an intrinsic feature of such systems, could also be examined [18].

From these biological models, suitable new computational metaphors and analytical frameworks could be created, to include appropriate representations for components, methods of assessing interactions between components, and processes to act on components. The frameworks

should also provide features that allow biological models to be represented and manipulated in a number of ways, and should permit the analysis and identification of generic properties. An instantiation of a framework should permit the capture of properties relevant to the application being developed. In an iterative process, the framework algorithms should be implemented and tested in order to test and develop the biological metaphors prior to their implementation and experimental exercises on the intended platform.

Taking this fuller view of immune networks may yield AIS algorithms that truly mimic the qualities of the diversity of immune network memory mechanisms, and may inform us as to the scalability of immune networks, their ability to cope more effectively with noise, their open nature, and the level of interaction both within the network and external to the network. Biology would benefit from the resulting sophisticated models, too.

3.5 Self, or danger?

Some researchers have begun taking a more interdisciplinary slant again. For example, Aicklen *et al.* [1] describe an ambitious interdisciplinary project investigating novel ideas from immunology such as danger theory [37], with application to computer security. Those authors propose to observe the biological system by undertaking new experiments to identify key signals involved in cell death, and identify the functions of such signals and how these affect immune cells. The aim is to shed light on how the immune system distinguishes self from non-self, in order to build effective immune-inspired computer security systems that no longer rely on the need to define *a priori* the self of the system. Although those authors make no reference to adopting a framework approach such as outlined above, we believe that taking such an approach would to help to ensure not only biologically-plausible algorithms, but effective and general solutions.

4. MATHEMATICAL ANALYSIS OF AIS

The adaptive immune system is able to recognise a vast number of different pathogens. The recognition of potentially harmful foreign invaders is a necessary (but not sufficient [62]) condition for the immune response to be activated. The robustness of the immune response relies on

the diversity of the lymphocyte repertoire. This diversity is achieved in two ways. Firstly, new cells are generated at random (although they may be selectively deleted later). Secondly, the number of lymphocytes is *huge*: the total number of lymphocytes in a mouse is of the order of 10^8, with approximately 10^7 different types of receptors, while in the human body there are of the order of 10^{12} lymphocytes [54]. These two aspects ensure the completeness of the repertoire, in the sense that essentially all possible pathogens that might be encountered will be recognised.

The importance of the size of the lymphocyte repertoire has not yet been addressed for the development of AIS algorithms. Many AIS algorithms, including CLONALG [10], the B-Cell Algorithm (BCA) [29,33,32] and MISA [65], are based purely on mutation and selection mechanisms for randomly generated clones (clonal selection), and these algorithms use very *small* populations of artificial immune cells. The main reason is probably straightforward: very large cell networks are computationally too expensive to implement. However, the emergent properties of complex systems can depend crucially on system size [24].

Lymphocytes interact and respond to one another in a highly complicated and *nonlinear* fashion: a small difference in the input to the system can produce a large (and complicated) change in the output. Nonlinearity is a key feature of complex systems that display emergent phenomena [56,57]. AIS algorithms that are inspired by the network of interactions in the adaptive immune system include the Resource-limited Artificial Immune Network (RAIN) [41], AINE [35] and aINET [11].

In the next subsections, we show how the theory of Markov chains can be used to understand the stochastic nature and convergence properties of AIS, and explain how nonlinear effects in network models may be analysed.

4.1 Optimisation problems and Markov chains

Many AIS algorithms are based purely on clonal selection mechanisms, without any interaction between the different members of the cell populations. Such algorithms are purely stochastic, in the following sense: given the state of the cell population at time t, the subsequent state at time $t + 1$ is a random variable. In many cases, the changing behaviour of the population with the time t (which varies in discrete steps) is naturally described in terms of a Markov chain (see [8,23] for background material).

Convergence is a highly desirable property for optimisation problems (although not necessarily so for other kinds of problem, such as ones to do with openness). The convergence of MISA [65], a multi-objective optimisation algorithm, has been proved via an associated Markov chain, under the assumption of an elitist selection mechanism. Many AIS algorithms have been tested on optimisation problems, because such problems are ubiquitous in applications to optimal control [34] and the calculus of variations [19,28,27]. Furthermore, AIS apparently perform well in tackling extremely hard biological optimisation problems, including protein folding [9,45].

We represent the state of our system by a variable X. The aim of the algorithm is to find the state value X that optimises (for example, minimises) the function $F(X)$. We can represent the state at time t by a random variable X_t, and optimise by making many iterations in time.

To be completely concrete, we illustrate this idea with the example of BCA [33,29,28,32] as an optimisation algorithm, inspired by the notion of contiguous hypermutation in B-cell clones. For BCA, the different cells in the population do not interact, so we can focus on the dynamics of one individual cell, and define the state X_t to be the value of the bit string corresponding to this cell. At each time t, a set of clones is taken from the cell, hypermutation is applied to the clones, and if for some clone C_t it happens that $F(C_t) < F(X_t)$ then the next state value is $X_{t+1} = C_t$, otherwise the original cell is kept and $X_{t+1} = X_t$ (see [33] for details). At each stage there is a definite probability for transition to a new state (bit string) value for X_{t+1}, and BCA is purely elitist in the sense that only mutations that result in improvement are kept (the value of $F(X_t)$ is non-increasing with t).

To describe the evolution of a cell in BCA in terms of a Markov chain, we label the possible states (bit strings) by an index j running from 1 to N, where N is the number of possible states, and we implicitly identify state values with their labels. We model the value of the bit string at time t by the probability distribution vector $\mathbf{v}_t = (v_{t,1}, v_{t,2} \ldots v_{t,N})$, with jth component $v_{t,j} = P(X_t = j)$: the probability of being in state j at time t. The initial value is given by $\mathbf{v}_0$.

The probability of transition between state j and state k is independent of the time t, and so can be represented by the $N \times N$ transition matrix $\mathbf{P}$, with entries $P_{jk} = P(X_{t+1} = k \mid X_t = j)$: the probability of being in state k at time $t + 1$, given being in state j at time t. $\mathbf{P}$ is a stochastic matrix: all of

its entries lie between 0 and 1, and the row sums satisfy $\sum_k P_{jk} = 1$; that is, from state j the cell must make a transition somewhere with probability 1.

To work out the probability distribution at time $t + 1$, we have

$$P(X_{t+1} = k) = \sum_{j=1}^{N} P(X_{t+1} = k \cap X_t = j) \tag{4.1}$$

$$= \sum_{j=1}^{N} P(X_{t+1} = k \mid X_t = j) P(X_t = j) \tag{4.2}$$

and so

$$v_{t+1,k} = \sum_{j=1}^{N} v_{t,j} P_{jk} \quad \text{or} \quad \mathbf{v}_{t+1} = \mathbf{v}_t \mathbf{P} \tag{4.3}$$

Because the transition matrix $\mathbf{P}$ is time-independent, the probability distribution vector $\boldsymbol{vt}$ can be written immediately in terms of the initial distribution, as

$$\mathbf{v}_t = \mathbf{v}_0 \, \mathbf{P}^t \tag{4.4}$$

It is evident from the form of equation (4.4) that if we wish to understand the long-time behaviour of the algorithm, we need to understand what happens to the powers of the transition matrix $\mathbf{P}^t$ as $t \to \infty$. For BCA it is further possible to prove that where there is a unique optimum state, it is reached with probability one in the limit $t \to \infty$;* in the terminology of Markov chains, a unique optimum is an *absorbing state* [23].

There are other instances of biologically-inspired computing being applied to optimisation problems and the calculus of variations, in particular the use of genetic algorithms (see references in [29]) and neural networks [38]. We expect that a similar analysis of stochastic properties will be relevant to some of these other approaches to biologically-inspired computing. Moreover, Markov chain techniques could be used to prove the convergence (or otherwise) of many other existing AIS and genetic algorithms. A thorough analysis of *rates* of convergence for optimisation algorithms would be even more useful, since the limit $t \to \infty$ cannot be reached in practice.

* Edward Clark (2004) private communication.

4.2 Nonlinear dynamics

The dynamics of cell populations in the immune system has been modelled extensively using nonlinear dynamical systems [21,47,54,63]. These models generally involve coupled systems of differential equations, taking the form

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}) \tag{4.5}$$

where the vector $\mathbf{z}(t)$ is the state vector at time t; or coupled discrete (difference) equations

$$\mathbf{z}_{t+1} = \mathbf{g}(\mathbf{z}_t) \tag{4.6}$$

In each case the components of the state vector $\mathbf{z}$ typically correspond to populations of cells or molecules of different species, and in general each of the components of the vectors $\mathbf{f}$ and $\mathbf{g}$ are nonlinear functions of their arguments. (For analytical calculations, it is often simpler to work with systems of ordinary differential equations; but for computer algorithms discrete equations are more appropriate).

The main difference between the dynamics of a Markov chain, defined by equation (4.3), and the dynamical systems here, is that the evolution of the state X_t is *random* and the vector $\mathbf{v}_t$ is a probability distribution, whereas the state vector $\mathbf{z}$ satisfies a purely *deterministic* evolution: the state at time t is determined uniquely by the initial state at $t = 0$.

The original development of AIS algorithms received much inspiration from mathematical models of immune networks (see section 3.4), such as those in [54], which are based on nonlinear interactions between cell populations. Therefore, in order to understand the behaviour of AIS algorithms, it makes sense to apply methods from the theory of nonlinear dynamical systems [48].

In general, given such a system such as that described by equation (4.6), one would like to make sense of the time evolution of the state vector by understanding how it moves in the phase space (the space of all possible states). One can explore this by *iterating* the difference equation: start with an initial value $\mathbf{z}_0$, then calculate $\mathbf{z}_1$ (a nonlinear function $f(\mathbf{z}_0)$), $\mathbf{z}_2$, $\mathbf{z}_3$, . . . and so on.

For a general mapping of the form

$$\mathbf{z}_t \longrightarrow \mathbf{z}_{t+1} = \mathbf{g}(\mathbf{z}_t) \tag{4.7}$$

it is natural to consider the *fixed points*, namely the solutions of the equation $\mathbf{z}_s = \mathbf{g}(\mathbf{z}_s)$ that remain fixed by equation (4.7). The initial value $\mathbf{z}_0$ is said to be in the *basin of attraction* of the point $\mathbf{z}_s$ if $\mathbf{z}_t \to \mathbf{z}_s$ as $t \to \infty$. For each fixed point there is a corresponding basin of attraction, and it is well known [51] that for even simple cases such as Newton-Raphson iteration, these basins are fractal sets. There are also periodic orbits which lie outside the basins of attraction. For a high-dimensional dynamical system the attractors can be more complicated than just points: they may be "strange attractors", which are themselves fractals [48]. The attractors and periodic orbits correspond to the important landmarks in the phase space.

AIS networks are based on nonlinear interactions between artificial cell populations. If an AIS network algorithm is required to perform some recognition task, or solve a multi-objective optimisation problem, then the different objectives might be regarded as attractors in a suitable phase space. Moreover, one would expect that the properties of the nonlinear mathematical models that have inspired the AIS, within the general development framework proposed in this paper, should inform the behaviour of the immune-inspired algorithms. Therefore nonlinear systems analysis will be essential in the development of robust algorithms.

However, this is not the full story, since as we have seen above, even the simplest AIS — those without any nonlinear interactions — are essentially stochastic in nature. (In the Markov chain context, we may also note that absorbing states play the role of attractors.) Hence the full treatment of AIS will require nonlinear *stochastic* dynamical systems. Very little is known about the explicit solutions of nonlinear stochastic differential equations, although there is a considerable interest in the linear case [7]. The application of nonlinear stochastic analysis to AIS is a novel approach which raises many exciting challenges for the future.

5. META-FRAMEWORKS FOR BIO-INSPIRED COMPUTATION

We have so far been speaking particularly of AIS, arguing the case for the framework in figure 1. This shows potentially many representations of the same systems under the same observations, each of which may provide different insights. Such distinct representations, although common, are seldom examined for unifying properties. Once we have a conceptual

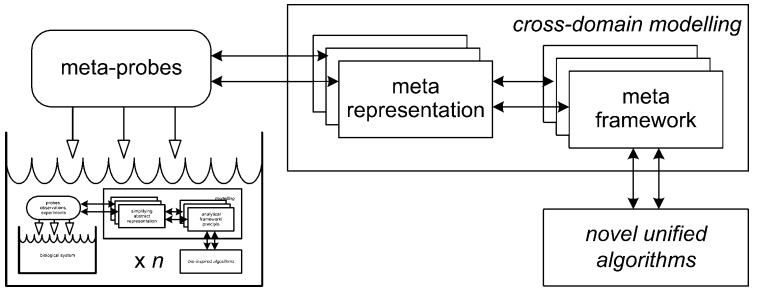

FIGURE 3
An outline conceptual framework for integrating bio-inspired computational domains.

framework, we can not only make such comparisons, we can go a step further: to examine and compare the separate conceptual, mathematical and computational frameworks, to develop more *integrated and generic frameworks*, and to expose essential differences.

To achieve this, we can apply the same conceptual model, at a higher level (figure 3). The key probes here are meta-questions. Just as the questions at the biological level influence the kinds of models developed, so the meta-questions influence the kinds of meta-models developed.

5.1 Meta-probes for complex system frameworks

What kind of meta-questions might we ask? Clearly, the questions asked influence the resulting framework. We have identified some initial areas thought to affect complex behaviour in general; questions that address notions such as openness, diversity, interaction, structure, and scale might lead to models of complex adaptive systems. The idea is to ask each question (suitably refined and quantified) across the range of frameworks being incorporated, and to use the answers as part of the input to build the meta-framework.

Openness: We do not want our computations to halt; we want continual evolution, continual growth, continual addition of resources: that is, *open*, *far-from-equilibrium* systems. How much openness is necessary? How is openness controlled by structure and interaction? How is system unity maintained in the presence of openness?

Diversity (heterogeneity) is present in all complex biological systems, and occurs in structure, behaviour, and interactions. When can we talk of an *average* agent? How much diversity is necessary within a level of

a structure? between levels? What does it cost? How does it combat fragility?

Interaction: Agents interact with their environment and with each other. What are the features of interaction within structural levels? between levels? What is the balance between computation and communication?

Structure: Biological systems have structure on a variety of levels, yet the levels are not crisply delineated. Are the levels we discern artefacts of our modelling framework? How can we recognise levels? When is a hierarchy an adequate structural model? How does structure affect interaction? What are the relationships between physical structures and information structures? What is the relationship with specialisation of function? with localisation of function?

Scale: Biological systems have a vast scale, a vast number of components. When and how does 'more' become 'different'? What are the critical points/phase transitions? How small can a system be, and still be emergent? When is a system too big? How important is multi-scale modelling? What are the relationships between scale and diversity?

Generic questions apply to each meta-probe question area X: What is the role of X within a system? What is the balance between X and not-X at the peak of complexity? How and when does X emerge? How does X evolve? How does physical embodiment affect X? How can we exploit X?

5.2 A meta-framework for population models

Many bio-inspired algorithms are based on *populations* of agents trained to perform some task, or optimise some function. The most obvious one is the area of evolutionary algorithms, based on analogy to populations of organisms breeding and selecting to become "fitter" [40]. In AIS, there are the positive and negative selection, and clonal selection algorithms. Swarm algorithms and social insect algorithms [5] are based on populations of agents whose co-operations result in problem solving. A neural network could be viewed as a population of neurons cooperating to perform a recognition task.

Given the number of underlying commonalities, it seems sensible to abstract a meta-framework from individual population based models. What are the key properties of population models, and how are they realised in the various individual models? Here we outline just a few similarities and differences of these models, which could be used in

constructing a population based meta-framework. (Since these individual frameworks themselves do not yet exist, this section is somewhat meta-speculative!)

All these models contain a population of individual agents. Members of the population usually exhibit a range of *fitnesses*, used when calculating a new population: fitter individuals have a greater effect on the composition of the next generation than do less-fit individuals. The aim is to find a population that is sufficiently fit for the task at hand.

In evolutionary algorithms (EAs), a population of *chromosomes* reproduces in a *fitness landscape*. Fitter individuals are *selected* more frequently, to breed the next generation. When described in these terms, the clonal selection algorithm looks very similar: the population comprises a collection of *antibodies*, which proliferate in an *affinity landscape*. The higher affinity individuals are cloned more, and mutated less, when producing the next generation. Additionally, the lowest affinity cells are replaced by random cells (providing automatic diversity maintenance). In swarm algorithms, a population of *particles* exists and adapts in a *fitness landscape*. Fitter individuals' properties are *copied* more by the next generation. In ant colony algorithms, a population of *paths* exist in a local fitness (path component length) landscape. The use of components from fitter (shorter) paths are *reinforced* by "pheromones" in the next generation, which is then constructed by "ants" following pheromone trails.

In EAs, clonal AIS, and ant algorithms, the fitness of the entire population is evaluated and used for selection and construction of next generation. Swarm algorithm evaluate the fitness of each individual relative to the others in its *local neighbourhood*. (Some EA variants incorporate *niching*, which provides a degree of locality.)

In EAs, swarm and ant algorithms, the result is the fittest member of the final population. In clonal AIS, however, the result is the entire final population of detectors; the individual detectors are each partial and unreliable, yet their combined cooperative effect makes the full robust detector.

Such commonalities and differences as outline above, once exposed and analysed, can be used to suggest more general algorithms. For example, the natural diversity maintenance of clonal AIS suggests ways for similar mechanisms to be added to other population algorithms, in a less *ad hoc* manner than currently. Also, many population algorithms find themselves forced to add some form of *elitism* to preserve the best

solution so far: clonal AIS is naturally elitist. One potentially interesting feature to explore is the relationship between the natural locality of swarm algorithms, and the locality inherent in danger theory.

Such a combination of models permits many of the meta-probe question outlined above to be asked. *Diversity* is a key question: how to maintain diversity within a single population, but additionally, should there be different "species", too? *Interaction* with the environment (laying and sensing pheromones) is crucial in the ant algorithms, and with other agents (at least at the level of copying their behaviour) in swarm algorithms. Co-evolution, with its effect on mutual fitness landscapes, can be regarded as a form of interaction. What *scale*, that is, what population size, is appropriate? The probes also force us to think of new issues: is *openness* a relevant aspect? Should we be concerned with flows of agents into and out of the population (other than by internal mechanisms of generational breeding)? And is there any way to exploit *structure*, given the homogeneity of most population algorithms?

This somewhat simplistic meta-framework sketch is built on the correspondingly simplistic population models. More sophisticated population models developed in terms of full conceptual frameworks would doubtless lead to much richer and more powerful meta-frameworks.

5.3 A meta-framework for network models

AIS networks, metabolic networks, auto-catalytic chemical reaction networks, intra-cellular protein interaction networks, inter-cellular cytokine, hormone and growth factor signalling networks, ecological food webs, are all examples of biological networks. Indeed, most biological processes operate through a complicated network of interactions, with positive and negative feedback control by factors that are themselves subject to similar controls. These networks function in a distributed fashion: most components have a variety of roles, and most functions depend on more than one component. This presumably underpins their robustness, whilst keeping the malleability required for adaptability and evolution. How this is achieved in practice is poorly understood.

Currently mathematical and computational descriptions of the structure of biological networks tend to be static (there is no time component to the architecture), closed (no inputs from the environment), and homogeneous (the types of nodes and connections are uniform, and new instances, and new kinds, of connections and nodes, are not supported). It will be necessary to develop novel mathematical approaches to model

real complex biological networks. Developing these new mathematical models in the context of the proposed conceptual framework will provide mechanisms for evaluating their appropriateness and power.

6. CASE STUDY

On-going work by some of the authors [44] is addressing the development of a more complete biologically inspired model of computation than is currently available, and provides a case study upon which to deploy the meta-probes described in section 5.1. The model is focused on the generic ability of the "higher" organisms to maintain their homeostatic state in a wildly varying environment. Key components of the organism which promote this ability have been identified as:

The immune system, which provides mechanisms for very long term adjustment of an organism's physiological state in a number of ways.

The neural system, which provides mechanisms capable of rapid and widely varying responses from very specific parts of the organism. These may involve interacting with the environment and/or internal organs.

The endocrine system, which provides mechanisms for relatively long term (compared to the neural system) adjustment of behaviour of the organism. The majority of endocrine system function is restricted to communication and control within the organism.

Our model is initially intended for use as a control system for complex electronic and electromechanical systems that would profit from long term autonomous operation. Autonomous robots are the specific exemplar of this type of system which we have chosen to target [43]. Our system contains direct analogues of these three systems:

An artificial immune system, which will consist of two layers of cells: an innate layer to pre-process and filter data from condition sensors distributed throughout the robot, and an adaptive layer capable of monitoring and adapting to problems signalled by the innate system. The performance and adaptation of the adaptive layer will be modulated by the concentrations of hormones generated by the artificial endocrine system.

An artificial neural system, which will consist of relatively conventional artificial neurons. These will be connected as multi-layer perceptrons, but the synapses of the neurons will be sensitive to hormones produced by the artificial endocrine system [43]. Hormone

concentrations will linearly suppress or excite the activity of synapses depending on the sign of the sensitivities associated with each hormone at each synapse. The inputs to the neural system will come from sensors of the environment and the outputs will control the system actuators.

An artificial endocrine system, which will consist of networks of cells connected in a similar manner to artificial neural networks. The terminal cells of each endocrine network will however produce and release "hormone" depending on the inputs that it receives. The hormone concentrations in the system will be global properties and will influence the behaviour of all three of the systems including the endocrine system itself.

The examination of this system in the light of the meta-probes of section 5.1 reveals a number of interesting properties and some key compromises.

The system is really only *open* in the sense that it is intended to operate for very long periods of time without interruption. There is no current mechanism proposed for the addition of elements to the system, although a number of artificial hormone controlled mechanisms are under consideration. Analysis of biological growth control hormones and potential analogues may provide mechanisms which could be integrated into our overall structure. At present however this lack of development is the first major compromise in our structure.

Diversity is present on a number of levels: the division between neurons, gland cells (the artificial endocrine system) and immune cells provides one layer of heterogeneity. The presence of multiple types of immune cell and multiple hormones provides another. All of these components behave in different ways and communicate in some common and some specialised ways. This diversity provides two key advantages: it allows the processing of diverse signals from a number of sources (both within the robot itself and from the environment) and the control of the system on a variety of time-scales in an integrated fashion. The artificial endocrine system provides a common communication channel capable of working across this range of time-scales and mechanisms.

Interaction is varied and locally simple. A large number of relatively low capacity communication channels promotes the emergence of complex behaviour. The transmission of signals across synapses in the artificial neural system and artificial endocrine system provides the shortest time-scale channel and the hormone concentrations provided by the artificial endocrine system the longest time-scale signal. In the biological system the immune system generates endocrine signals over potentially

very long time-scales that are capable of controlling both the neural and endocrine system; in our artificial system we have chosen to ignore these signals. This is a further key compromise that restricts the ability of the artificial immune system to control and avert damage. Further examination of the role of hormones and cytokines and specifically the stress response [15,49] may provide clues about how to successfully exploit such a mechanism.

The *structure* of the system is crudely seen in the description at the beginning of this section: neural system composed of probably several separate neural networks, each composed of hormone sensitive neurons and synapses; an endocrine system broken down in a very similar way to the neural system; an immune system composed of two coupled populations of cells (innate and adaptive). The specialisation of function within each system affects different aspects of the potential for control on different time-scales and of different physical mechanisms within the robot.

The proposed *scale* of the system presented here is currently limited. This is mainly due to the major compromise on the openness of the system. The engineering of a system based on these components will only be possible for relatively small systems. The inherent modularity of structure (especially within the neural and endocrine systems) allows these systems to built out of a number of pre-trained and pre-organised subcomponents. Ultimately the use of a developmental mechanism within the system will allow much larger systems and a more blurred internal structure. For example, cross-linking between the endocrine and neural networks might be permitted to develop and may well allow more flexible, robust and effective control to develop.

Whilst the limitations of the system proposed are self-evident and extensive, we believe that the selection of components and mechanisms provides a sound basis for effective control of complex electro-mechanical systems. We also believe that it fits well with the framework proposed here and that the framework provides a useful tool for the analysis and further development of the system.

7. DISCUSSION AND CONCLUSION

We have argued that bio-inspired algorithms would benefit from exploiting more sophisticated biological models, and from being based

on sound analytical principles; we believe that biology could benefit from the resulting sophisticated models, too. We have outlined what we believe to be a suitable conceptual framework including these various components. We have suggested how AIS network models might fit into this framework.

We have additionally sketched how meta-frameworks, based on the same underlying structure, might be applied at higher levels to unify various kinds of bio-inspired architectures, and we have suggested how population based models, including AIS models, might form one such meta-framework. We do not expect that every individual model will fit perfectly into an integrated model: part of the development process will be to expose essential differences as well as to integrate common abstractions.

One exciting prospect of a unified meta-framework is the possibility of a *novel instantiation*, possibly using concepts from across a range of biological domains, and possibly using concepts from outside biology (since words like "Lamarck" and "teleology" need not be so necessarily dismissed in the artificial domain). This would allowing the building of a chimerical computational framework that is inspired by biology, but not restricted to any one particular biological domain.

8. Acknowledgements

The conceptual framework described in this paper was developed during the EPSRC-funded EIVIS project. We would like to thank the other members of the project for their invaluable contributions to some of the ideas in this paper: Andrew Anderson, Jim Austin, Brian Bell, Peter Bentley, David Broomhead, Robin Callard, Steve Furber, David Halliday, Douglas Kell, Alan Murray, Jaroslav Stark, Stefan Wermter, David Willshaw, Xin Yao, Peter Young.

REFERENCES

[1] Aicklen, U., Bentley, P., Cayzer, P., Kim, J. and McLeod. J. (2003). Danger theory: the link between AIS and IDS? In Timmis *et al. 61*, 147–155.

[2] Bersini, H. (1991). Immune network and adaptive control. In *Proc. First European Conference on Artificial Life*, 217–226. MIT Press.

[3] Bersini, H. (1992). Reinforcement and recruitment learning for adaptive process control. In *Proc. Int. Fuzzy Association Conference (IFAC/IFIP/IMACS) on Artificial Intelligence in Real Time Control*, 331–337.

[4] Bersini, H. and Varela, F.J. (1994). The immune learning mechanisms: reinforcement, recruitment and their applications. In R. Paton, editor, *Computing with Biological Metaphors*, 166–192. Chapman & Hall.

[5] Bonabeau, E.W., Dorigo, M. and Theraulaz, G. (1999). *Swarm Intelligence: from natural to artificial systems*. Addison Wesley.

[6] Bradley, D.W. and Tyrrell, A.M. (2002). Immunotronics: novel finite state machine architectures with built in self test using self-nonself differentiation. *IEEE Trans. Evo. Comp., 6*(3), 227–238. June.

[7] Brzeźniak, Z. and Zastawniak, T. (1999). *Basic Stochastic Processes.* Springer.

[8] Cox, D.R. and Miller, H.D. (1965). *The Theory of Stochastic Processes.* Chapman and Hall, London.

[9] Cutello, V. and Nicosia, G. (2002). An immunological approach to combinatorial optimization problems. In F.J. Garijo, J.C. Riquelme, and M. Toro, editors, *Advances in Artificial Intelligence – IBERAMIA 2002*, volume 2527 of *LNAI*, 361–370. Springer.

[10] de Castro, L.N. and Timmis, J. (2002). *Artificial Immune Systems: A New Computational Intelligence Approach.* Springer.

[11] de Castro, L. N. and Von Zuben, F. N. (2000). An evolutionary immune network for data clustering. In *6th Brazilian Symp. Neural Networks, SBRN* '00, 84–89. IEEE.

[12] de Castro, L.N. and Von Zuben, F.J. (2000). The clonal selection algorithm with engineering applications. In *Workshop on Artificial Immune Systems and Their Applications, GECCO*, 36–37.

[13] de Castro, L.N. and Von Zuben, F.J. (2001). aiNet: an artificial immune network for data analysis. In H.A. Abbass, R.A. Sarker, and C.S. Newton, editors, *Data Mining: a heuristic approach*, chapter XII. Idea Group Publishing.

[14] Farmer, J.D., Packard, N.H. and Perelson, A.S. (1986). The immune system, adaptation, and machine learning. *Physica D, 22*, 187–204.

[15] Feder, M. and Hofmann, G. (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Ann. Rev. Physiol., 61*, 243–282.

[16] Forrest, S., Perelson, A., Allen, L. and Cherukuri, R. (1994). Self-nonself discrimination in a computer. In *Proc. IEEE Symp. on Research in Security and Privacy*, 202–212.

[17] Freitas, A.A. and Timmis, J. (2003). Revisiting the foundations of artificial immune systems. In Timmis *et al. 61*, 229–241.

[18] Gammaitoni, L., Hanggi, P., Jung, P. and Marchesini, F. (1998). Stochastic resonance. *Rev. Mod. Phys., 70*(1), 223–287.

[19] Gelfand, I.M. and Fomin, I.M. (1963). *Calculus of Variations.* Prentice-Hall.

[20] Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison Wesley.

[21] Goldstein, B., Faeder, J.R. and Hlavacek, W. (2004). Mathematical models of immune receptor signalling. *Nature Rev. Immunol., 4*, 445–456.

[22] Gonzalez, F. and Dasgupta, D. (2003). Anomaly detection using real-valued negative selection. *J. Genetic Prog. and Evolvable Machines, 4*, 383–403.

[23] Grimmett, G.R. and Stirzaker, D.R. (1982). *Probability and Random Processes.* Oxford University Press.

[24] Haken, H. and Mikhailov, A. (eds) (1993). *Interdisciplinary Approaches to Nonlinear Complex Systems.* Springer.

[25] Hightower, R.R., Forrest, S.A. and Perelson, A.S. (1995). The evolution of emergent organization in immune system gene libraries. In L.J. Eshelman, editor, *Proc. 6th Int. Conf. Genetic Algorithms*, 344–350. Morgan Kaufmann.

[26] Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems.* University of Michigan Press.

[27] Hone, A. (2004). A piece of the action. *Physics World, 17*(9), 64.

[28] Hone, A. and Kelsey, A. (2004). Optima, extrema and artificial immune systems. In Nicosia *et al. 46*, 80–90.

[29] Hone, A., Kelsey, J. and Timmis, J. (2003). Chasing chaos. In R. Sarker *et al.*, editors, *Proc. Congress on Evolutionary Computation*, 413–419. IEEE.

[30] Jerne, N.K. (1974). Towards a network theory of the immune system. *Annals Immunol., 125C*, 373–389.

[31] Janeway Jr, C.A. and Medzhitov, R. (2002). Innate immune recognition. *Ann. Rev. Immunol., 20*, 197–216.

[32] Kelsey, J. (2004). *An immune system-inspired function optimisation algorithm.* Master's thesis, University of Kent.

[33] Kelsey, J. and Timmis, J. (2003). Immune inspired somatic contiguous hypermutation for function optimisation. In Cantu-Paz *et al.*, editors, *GECCO 2003*, volume 2723 of *LNCS*, 207–218. Springer.

[34] Kirk, D.E. (1997). *Optimal Control Theory: An Introduction.* Prentice Hall.

[35] Knight, T.P. and Timmis, J. (2002). A multi-layered immune inspired approach to data mining. In A. Lotfi, J. Garibaldi, and R. John, editors, *Proc. 4th Intl. Conf. Recent Advances in Soft Computing*, 266–271.

[36] Kourilsky, P. and Truffa-Bachi, P. (2001). Cytokine fields and the polarization of the immune response. *Trends Immunol., 22*, 502–509.

[37] Matzinger, P. (2002). The danger model: a renewed sense of self. *Science, 296*, 301–305.

[38] Meade, A.J. and Sonneborn, H.C. (1996). Numerical solution of a calculus of variations problem using the feedforward neural network architecture. *Advances in Engineering Software, 27*, 213–225.

[39] Medzhitov, R. and Janeway Jr, C.A. (2002). Decoding the patterns of self and nonself by the innate immune system. *Science, 296*, 298–300.

[40] Mitchell, M. (1996). *An Introduction to Genetic Algorithms.* MIT Press.

[41] Neal, M. and Timmis, J. (2001). A resource limited artificial immune system for data analysis. *Knowledge Based Systems, 14*, 121–130.

[42] Neal, M.J. (2003). Meta-stable memory in an artificial immune network. In Timmis *et al.*, Proceedings of the 2nd International Conference of Artificial Immune Systems. LNCS 2787, 168–180.

[43] Neal, M.J. and Timmis, J. (2003). Timidity: a useful mechanism for robot control? *Informatica, 27*(4), 197–204.

[44] Neal, M.J. and Timmis, M.J. (2004). Once more unto the breach . . . towards artificial homeostasis? In L.N. de Castro and F.J. Von Zuben, editors, *Recent Advances in Biologically Inspired Computing*, 340–365. IGP.

[45] Nicosia, G. (2004). Combinatorial landscapes, immune algorithms and protein structure prediction problem. Poster at Mathematical and Statistical Aspects of Molecular Biology (MASAMB XIV), Isaac Newton Institute, Cambridge.

[46] Nicosia, G., Cutello, V., Bentley, P.J. and Timmis, J. (eds) (2004). *ICARIS* 2004, volume 3239 of *LNCS*. Springer.

[47] Nowak, M.A. and May, R.A. (2000). *Virus dynamics.* Oxford University Press.

[48] Ott. E. (1993). *Chaos in dynamical systems.* Cambridge University Press.

[49] Ottaviani, E. and Franceschi, E. (1996). The neuroimmunology of stress from invertebrates to man. *Progress in Neurobiology, 48*, 421–440.

[50] Padgett, D.A. and Glaser, R. (2003). How stress influences the immune response. *Trends Immunol., 24*, 444–448.

[51] Peitgen, H.-O. and Richter, D.H. (1986). *The Beauty of Fractals: Images of Complex Dynamical Systems.* Springer.

[52] Perelson, A.S. (1989). Immune network theory. *Imm. Rev., 110*, 5–36.

[53] Perelson, A.S. (2002). Modelling viral and immune system dynamics. *Nature Rev. Immunol., 2*, 28–36.

[54] Perelson, A.S. and Weisbuch, G. (1997). Immunology for physicists. *Rev. Mod. Phys., 69*, 1219–1267.

[55] Romanyukha, A.A. and Yashin, A.I. (2003). Age related changes in population of peripheral T cells: towards a model of immunosenescence. *Mechanisms of Ageing and Development, 124*, 433–443.

[56] Scott, A. (1999). *Nonlinear Science.* Oxford University Press.

[57] Scott, A. (ed.) (2004). *Encyclopedia of Nonlinear Science.* Routledge.

[58] Sprent, J. and Surh, C.D. (2002). T cell memory. *Ann. Rev. Immunol., 20*, 551–579.

[59] Taylor, D. and Corne, D. (2003). An investigation of the negative selection algorithm for fault detection in refrigeration systems. In Timmis *et al. 61*, 34–45.

[60] Timmis, J. (September 2000). *Artificial Immune Systems: a novel data analysis technique inspired by the Immune Network Theory.* PhD thesis, Department of Computer Science, University of Wales.

[61] Timmis, J., Bentley, P. and Hart, E. (eds) (2003). *ICARIS* 2003, volume 2787 of *LNCS*. Springer.

[62] van den Berg, H.A. (2004). Control of T-cell immunity: design principles without the wetware. Technical Report UKC/IMS/04/36, University of Kent. preprint.

[63] van den Berg, H.A. and Rand, D. A. (2004). Quantitating T cell responsiveness. Technical report, University of Warwick. preprint, submitted to Interface.

[64] Varela, F., Coutinho, A., Dupire, B. and Vaz, N.N. (1988). Cognitive networks: immune, neural and otherwise. In A. S. Perelson, editor, *Theoretical Immunology, part 2*, 359–375. Addison-Wesley.

[65] Villalobos-Arias, M., Coello Coello, C.A. and Hernández-Lerma, O. (2004). Convergence analysis of a multiobjective artificial immune system algorithm. In Nicosia *et al. 46*, 226–235.