The Next Generation of Immunity-Based Systems: From Specific Recognition to Computational Intelligence

Yoshiteru Ishida

Department of Knowledge-Based Information Engineering, Intelligent Sensing System Research Center, Research Center for Future Vehicle, Toyohashi University of Technology, Tempaku, Toyohashi 441-8580, Japan, ishida@tutkie.tut.ac.jp

1 Introduction

The post-genome era proved that DNA sequence data [11, 26] with structural and functional analysis on genes archived in many data bases can support in developing new bio-engineering technologies and can drive systemic views for biological systems. However, the post-genome era also proved that sequence data alone is not sufficient, but revealed that higher knowledge of the function of proteins is indispensable. Personalized medicine required not only sequence data, but further knowledge such as SNP (single nucleotide polymorphism) and of functioning of proteins and its deployment to interacting systems such as gene networks, giving birth of a new territory called proteome.

Considering such trends in the post-genome era, we propose possible directions for immunity-based systems (IMBS). One such approach is a constructive systems approach, taking fundamental properties of the component and trying to construct a fundamental function. The synthetic approach has been extensively studied [3, 5, 18, 24], to mention but a few). A constructive approach that assumes an intrinsic character of the components (such as antibodies), and constructs the fundamental function of the immune system from the component. Although it should not be limited to two, another possible direction of next generation immunity-based systems is to extend and enhance models and simulations to be operational: that is, involving medications as a control to the systems with the immune system and pathogen interactions. This would be made possible by using post-genome genetic data. The operational models and simulations allow, for example, involving the immune system in personalized medicine. Information of disease agents, medicine, and host agents are required for personalized medicine.

When functions are focused and more pathways are revealed, biological systems will be studied as a system of interacting components and processes.

Y. Ishida: *The Next Generation of Immunity-Based Systems: From Specific Recognition to Computational Intelligence*, Studies in Computational Intelligence (SCI) **115**, 1091–1121 (2008)
www.springerlink.com

Restricting discussions on immunity-based systems, the post-genome era naturally proceeds to study immune systems focusing not only on discovery of genes related to the immune system such as MHC (major histocompatibility complex), but also its systemic organization and the knowledge of the organism [9]. While on the other hand, studies on the immune system are indispensable for personalized medicine, since the immune systems is one important component for personalization of medicine, and such personal differences are integrated in the immune system. The other two components for personalized medicine are: pathogens and medications.

Too close mimicking and superficial analogy could be misleading, not only for biologically inspired systems but also for computing (circuits) that use biological system components. Biological systems can be neither simple nor optimal. One reason for apparently complex and roundabout implementations is that biological systems have large-scale interactions in a spatio-temporal sense. In space, they interact with an environment that includes not only nonself but self. In time, they undergo an adaptation within an individual time scale, as well as evolution in a species time scale. Thus, it is suggested that a superficial analogy could be misleading in mimicking biological systems; biological mimicking should not be done at a phenomenological level, but instead on a principle level.

Another reason for the complexity and intangibility of biological systems seems largely due to the feature of the material they comprise – namely, proteins. This would suggest that a constructive systems approach to biological mimicking systems can be not only an alternative to modeling and simulations but also a complementary tool supporting and guiding the modeling and simulation. The huge information available in post genome era allows a systems approach to biology, and this trend is accelerated for immunology as well. Next generation immunity-based systems may depend not only on a modeling/simulation approach, but also on a constructive approach that might bridge between the material and experimentally-based immunology and model/simulation-based informatics on bio-systems.

In summary, next generation immunity-based systems should focus on the following:

- A constructive systems approach to computational intelligence and artificial systems by assuming material similar to the real biological systems
- Extension and enhancement of models and simulations so that several operations are possible, involving genome data in the post-genome era, and targeting bioinformatics incorporating the immune system (such as personalized medicine involving the immune system).

This Chapter explores the first issue – that is, we consider next generation immunity-based systems by first revisiting conventional immunity-based systems (focusing on recognition capability), and next by extending them by

restricting antibodies (or peptides in general) as a base material for a constructive systems approach to immunity-based systems. Another Chapter in this volume explores the models without recognition – that is, all the agents mounting only effectors but without sensors (see Chap. 4). In this Chapter, each agent is assumed to be capable of recognizing the state of other agents.

This Chapter is organized as follows: Sect. 2 focuses on the preliminary problem of whether recognition is indeed needed, focusing on the specific task of abnormal state eradication on a simple network. Section 3 addresses the problem of networked recognition that involve action counterpart, hence agents can not only recognize but also be recognized. Section 4 further introduces adaptation by assuming agents can not only reproduce but also mutate in the receptor counterpart. Section 5 considers arrayed recognition, which is the very first step, even before networked recognition; however, it assumes specific recognition capability of antibody-antigen recognition.

2 Impact of Recognition

It is still controversial whether the immune system actually needs to discriminate 'self' and 'nonself' in order to eliminate *nonself* [20], however, elimination is actually done, and hence the double-sided property that elimination could be directed not only towards nonself, but also to self. Thus, the immune system is a double-edged sword.

This Section considers the impact of recognition in a simple model. To observe the impact, a simplified problem of network cleaning is considered. In information systems, the repairing units can repair others simply by copying their content, but could have spread contamination when the repairing units themselves are contaminated. We consider the possibility of cleaning up the network by mutual copying. However repair by copying in information systems is also a 'double-edged sword', and it needs to be identified when the network can really eradicate abnormal elements from the system.

The self-repairing network consists of units capable of repairing other connected units. We call the connected units as neighbor units based on the terminology of cellular automata (CA). Although mutual repair and other interactions involved may be done in an asynchronous manner, our model considers synchronous interactions for simplicity. Each unit tries to repair the units in its neighborhood, however whether it can really repair or not depends on several factors: the state of the repairing unit and the success rate of the repair.

In a mathematical formulation, the model consists of three elements $(\mathbf{U}, \mathbf{T}, \mathbf{R})$ where $\mathbf{U}$ is a set of units, $\mathbf{T}$ is a topology connecting the units, and $\mathbf{R}$ is a set of rules of the interaction among units. In the simulations to come, a set of units is a finite set with N units, and the topology is restricted

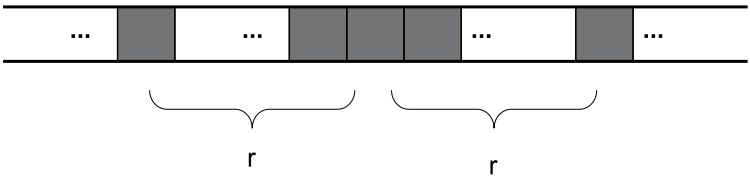

Fig. 1. One-dimensional lattice with the neighborhood radius r; the next state of the cell will be determined by $2r + 1$ nodes in the neighborhood

to the one-dimensional lattice as shown in Fig. 1. The network structure could be an n-dimensional array, complete graph, random graph, or even a scale-free network. In our one- or two-dimensional lattice, each unit has S neighbors and the lattice with a boundary condition – in other words, the structure of the lattice is a ring with unit 1 adjacent to the unit N in the case of a one-dimensional lattice. Also, we restrict our discussion to cases where each unit has a binary state: normal (0), and abnormal (1).

2.1 An Impact of Recognition is a Double-Edged Sword

Our model also involves recognition of the states (normal or abnormal) of a target node before trying to repair it. For simplicity, frequency of recognition is controlled by a recognition rate γ. When recognition is undertaken (with a probability γ), successful recognition occurs with a recognition success rate γ_0 when performed by normal nodes, and γ_1 by abnormal nodes. If the target node is identified as 'abnormal', repair action take place. When recognition does not occur (with a probability $1 - \gamma$), the repair action takes place with the probability μ. Thus, if recognition is completely suppressed ($\gamma = 0$), this new model reverts to the original model. Figure 2 shows the procedure of recognition and repair.

Computer simulations are conducted in a one-dimensional array with a ring structure (periodic boundary condition). The parameters listed in Table 1 are fixed throughout the simulations. Other parameters: $\gamma, \gamma_1, \mu, and \alpha_1$ are varied to observe the impact of recognition.

We are concerned with the problem: "Is recognition really necessary?" Moreover, if 'yes', then when and how should the recognition should be incorporated? In the following simulations, we pursue the problem of identifying an appropriate level of recognition (namely, γ) when the adverse effect of abnormal units (that is, γ_1 and α_1) is given.

When the rate of successful repair by abnormal nodes (that is, α_1) is given, what is the minimum level of recognition (namely, γ) required for abnormal node eradication? Figure 3 plots the minimum level of γ. As observed and already reported, we do not care about the level of repair and/or recognition when α_1 exceeds a threshold (0.4 in this simulation). However, when α_1 is less

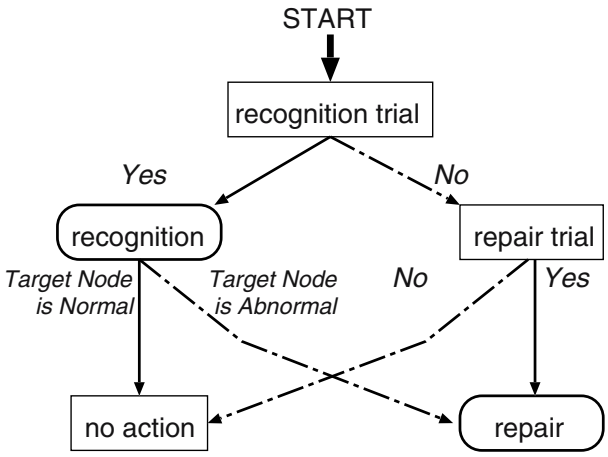

Fig. 2. Recognition carried out prior to repair

Table 1. Parameter list for the simulations

	Description	Value
N	number of nodes	500
$N_f(0)$	initial number of failure nodes	250
r	neighborhood radius	1
T	number of time steps for each trial	5000
N_T	average number of trials	10
α_0	repair success rate by normal nodes	1
γ_0	recognition success rate by normal nodes	1

than the threshold, recognition is needed (γ is positive) to eradicate abnormal nodes. Furthermore, the smaller the level of repair (μ), the smaller the level of recognition (γ) can be.

In this simulation (and with the specific model parameters as indicated), only repair by copying suffices for abnormal node eradication when the rate of successful repair by abnormal nodes exceeds some level. However, recognition before repair is required when the rate does not exceed this level.

3 Immunity-Based Systems: Evolved Recognitions

3.1 Definition of Immunity-Based Systems

Although recognition may not be needed under an optimistic situation in a simple network model, as in the previous Section, immunity-based systems (IMBS) [13] assume each agent mounts receptor counterpart. IBMS as a design

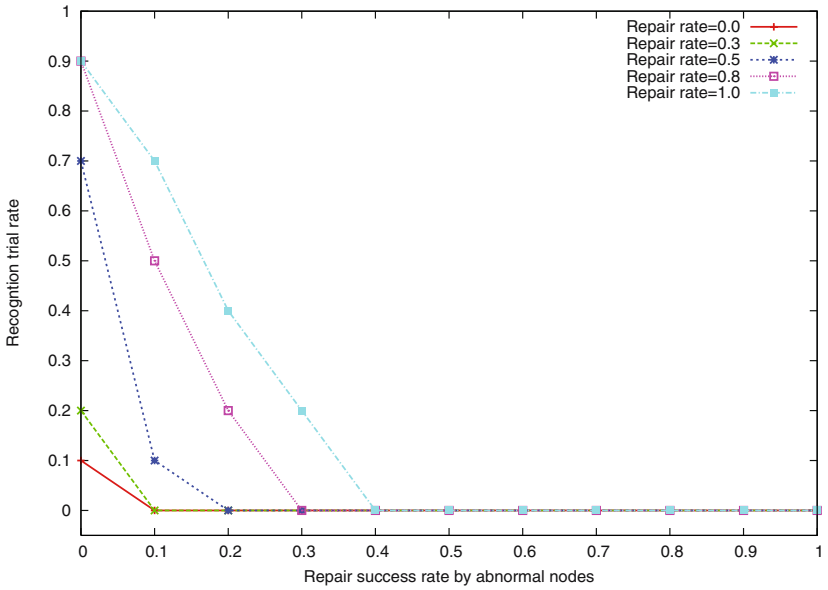

Fig. 3. Minimum level of γ to eradicate abnormal nodes when α_1 is given

paradigm has the following three properties:

1. a *self-maintenance system* with monitoring not only of the nonself but also of the self
2. a *distributed system* with autonomous components capable of mutual evaluation
3. an *adaptive system* with diversity and selection

In the following Sections, networked recognition focuses on the first two, while adaptive recognition involves the third one of these.

3.2 Networked Recognition

[17] proposed the immune network. In network theory, the immune system is not merely a 'firewall' but a network of antigen-antibody reactions. That is, when an antigen is administered, it stimulates the related cells and causes them to generate antibodies. However, the generated antibodies themselves are antigens to other cells, and consequently result in another antibody generation. The antigen-antibody reaction percolates like a chain reaction and hence requires a regulation mechanism. An analogy of this problem in engineering design is the 'alarm problem' of placing mutually activating and non-activating alarms whose sensitivity must be appropriately set to avoid false negative and false positive.

There is a variety among immune system models, even if we restrict ourselves to those by differential equations. If they were to be described by a single

equation with x_i: the number of recognizing (or recognized) sets (T-cells, B-cells, antibodies, and antigens) and a_{ij}: interactions between type i and type j (positive for stimulation and negative for suppression), the equation would be:

$$\frac{dx_i(t)}{dt} = F(\{x_i(t)\}, \{a_{ij}(s_i(t), s_j(t), aff_{ij}(t))\}) \tag{1}$$

where s_i denotes the state of the type i entity (for example, activated/inactivated, virgin/immune, and so on); and aff_{ij} the affinity between these two types. The dimension of x_i (the number of types) can vary, since a new type can be born, mutated from other types, or just injected in the case of antigens.

So far, this is not much different from the population dynamics of general ecological systems described by the Lotka-Volterra equation, for example. What makes this equation peculiar to the immune system is that interactions a_{ij} vary depending on the states of type i and type j entities, as well as the affinity between them. It is this affinity that models of the immune system devised by several techniques, such as the 'shape-space' model [23], where antigens and antibodies are expressed as points in the space, which allows the affinity between them to be measured as a distance between the points. Several spaces such as continuous and discrete ones are considered, hence several distances too (such as Euclidean and Hamming distance).

In such dynamical models, immunological concepts such as immune memory and tolerance are mapped to attractors of the dynamical systems. Within the context of problem solving, attractors of the system are mapped to solutions, thus the perturbed state (nonself) will be attracted to the solution (self), and hence nonself will be eliminated and self will be preserved. Positive and negative regulation will be interpreted as reinforcement and elimination.

Let us consider a credit assignment problem where high credit should be assigned to the self and low credit to nonself. Weighting the vote and propagating the information correctly identifies the abnormal agents. A continuous dynamic network is constructed by associating the time derivative of the state variable with the state variables of other agents connected by the evaluation chain. Further, considering not only the effect from evaluating agents, but also that from evaluated agents leads to the following dynamic network:

$$\frac{dr_i(t)}{dt} = \sum_j T_{ji} R_j + \sum_j T_{ij} R_j - 1/2 \sum_{j \in \{k: T_{ik} \neq 0\}} (T_{ij} + 1) \tag{2}$$

where $R_i(t) = \frac{1}{1+\exp(-r_j(t))}$ and

$$T_{ij} = \begin{cases} -1 & \textbf{if } \textit{evaluating agent} \text{ i } \textit{is normal and evaluated agent} \text{ j } \textit{is faulty} \\ 1 & \textbf{if } \textit{both agents} \text{ i } \textit{and} \text{ j } \textit{are normal} \\ \pm 1 & \textbf{if } \textit{evaluating agent} \text{ i } \textit{itself is faulty} \\ 0 & \textbf{if } \textit{there is no evaluation from agent} \text{ i } \textit{to agent} \text{ j} \end{cases} \tag{3}$$

In evaluating agents, agent j will stimulate (inhibit) agent i when $T_{ji} = 1(-1)$. We call this model the *black-and-white model*, meaning that the network tries to separate an abnormal agent clearly from a normal agent; namely, the *credibility* (which differs from the probabilistic concept of *reliability*) of an agent tends to be 1 (fully credible) or 0 (not credible), not an intermediate value. Moreover, we have proposed several variants of this dynamic network, such as the *skeptical model* and the *gray model* for different engineering needs. The results presented in this Chapter are generated only from the *black-and-white model*.

Figure 3 shows an example of the evaluation chain of mutual voting. The pattern associated with the evaluation arc shows a case when agents 4 and 5 are faulty. A positive arc from agent i to agent j indicates that agent i voted positively for agent j (in other words, considered 'normal'), and a negative arc negatively (that is, considered 'abnormal'). Formally, evaluation results are assumed to give the following pattern:

$$T_{ij} = \begin{cases} -1 & \textbf{if } \textit{evaluating agent} \text{ i } \textit{is normal and evaluated agent} \text{ j } \textit{is faulty} \\ 1 & \textbf{if } \textit{both agents} \text{ i } \textit{and} \text{ j } \textit{are normal} \\ \mp 1 & \textbf{if } \textit{evaluating agent} \text{ i } \textit{itself is faulty} \\ 0 & \textbf{if } \textit{there is no evaluation from agent} \text{ i } \textit{to agent} \text{ j} \end{cases} \tag{4}$$

Simple voting at each agent does not work, since three agents (2, 3, and 5) are all evaluated as 'faulty' by two other agents, and hence cannot be ranked in terms of credibility. Since an abnormal agent may give faulty results, these votes should be weighted. Next, let us introduce a binary weight for each agent: 0 (inactive or abnormal) when the sum of votes for the agent is negative, and 1 (active or normal) when the sum of votes for the agent is zero or positive. Starting with all agents active, evaluating the weight would synchronously result in the sequence of credibility vector $(R_1 R_2 R_3 R_4 R_5)$, as shown on the right of Fig. 4.

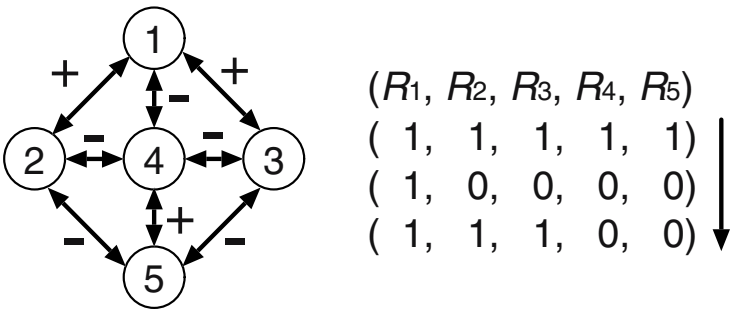

Fig. 4. An example evaluation chain of mutual voting (*left*), and the credibility vector sequence (*right*)

Example: Application to Automobile Engine Sensor Diagnosis [12]

A dynamic relational network can be built in roughly two steps:

1. *Line up candidates of relational arcs*: find causally related sensors by investigating correlation by checking indices such as coefficient of correlation.
2. *Narrow down the above candidates*: remove those arcs from sensor A to B if the test from sensor A to B generates false positives or false negatives.

A time series analysis is carried out for step 1 (using mutual correlation matrix), and/or for step 2 (prediction by the models of time series analysis). As reported below in the case of both the combustion control system of an automobile engine and for a particular fault in an air-flow sensor, a statistical analysis of up to step 1 for building the network suffices. However, time series analysis (with the VAR model) is used to determine the sign of an arc (evaluation from node i to node j) in online diagnosis.

In this Section, a case study with statistical analysis for building the relational network is reported. S_a indicates the data from sensor A. In step 1, arcs between A and B are added if | coefficient of correlation between S_a and S_b $|\geq 0$. Figure 5 shows a network built when $\theta = 0.4$ and only step 1 in the algorithm is used. The network turned out to be complete. Signs are a snapshot of evaluation based on the sensor data. Gray level in the nodes indicates credibility. Dark nodes correspond to high credibility, while light nodes to low credibility (that is, evaluated as 'faulty') [14]. The signs of arcs in the network

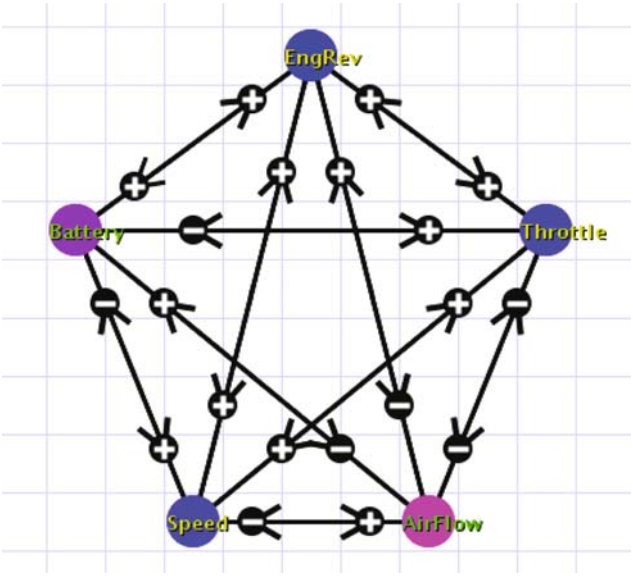

Fig. 5. A network with arcs added when | coefficient of correlation $|\geq 0.4$ in cruise phase (EngRev: engine revolutions; Battery: battery voltage) [14]

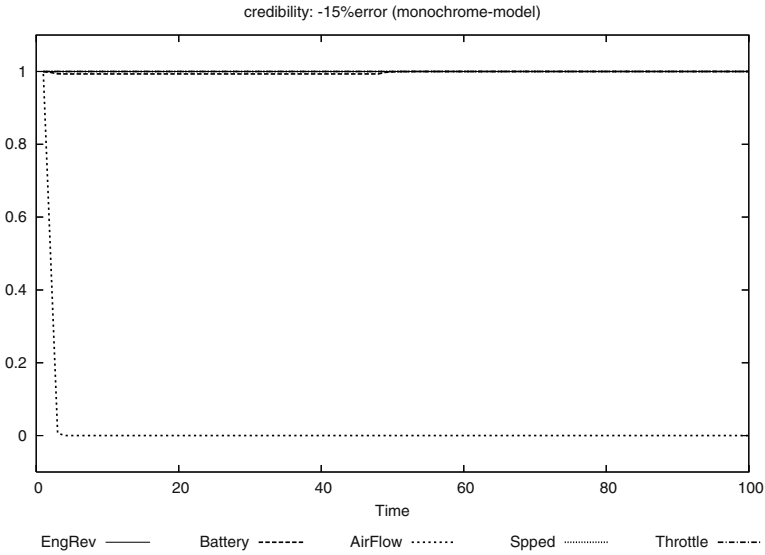

Fig. 6. Diagnosis by the evaluations calculated from the VAR model when the air flow sensor is faulty

change dynamically in online diagnosis; Fig. 5 shows only a snapshot of signs. The network structure does not change during the diagnosis.

It should be noted that the above calculation is done using only normal sensor data. Figure 6 shows the time evolution of credibility calculated by the time series analysis stated above. The dotted line shows the time evolution of sensor credibility. Only the credibility of the faulty sensor (Air Flow) becomes 0, hence the diagnosis is successful.

As heuristics for solving problems by the networked recognition, the following remarks apply:

- Signals from different agents should be related by signal processing models and statistical analysis [14] to map from the signals to evaluations (positive/negative sign of the network as in Fig. 5).
- Interactions among agents should be designed so that attractors of the entire network correspond to solutions to be obtained

Compared with the Bayesian Network [22], the above networked recognition is not able to obtain probabilities of events, however the problem solving mechanism can be directly embedded in the system where many agents are able to relate with each other. When applied to the signal processing domain, as in the above example, networked recognition is able to utilize the information embedded in the relations between the signals, as well as the information embedded in each signal itself – that is, both absolute and relative information in multiple signals can be involved. Networked recognition can be applied

not only to signal processing domain but also to other domains, such as data mining, and search engines, if distributed agents are involved and mutually related.

3.3 Adaptive Recognition

[4] speculated on clonal selection theory based on antibody production. An immune algorithm for a population of agents is proposed based on the clonal selection concept [16]. The most naive immune algorithm has the following three steps carried out in parallel by agents distributed over the system. In the algorithm, agents (corresponding to the immune cells) have not only recognition and communication capabilities, but also reproduction capability with possible mutation.

1. *Generation of diversity*: diverse agents with distinct specificity of the receptor and the effector are generated;
2. *Establishment of self-tolerance*: agents are adjusted to be insensitive to 'known patterns' (self) during the developmental phase;
3. *Memory of nonself*: agents are adjusted to be more sensitive to 'unknown patterns' (nonself) during the working phase.

Figure 7 shows a process which basically mimics the affinity maturation; affinity will increase by exploring diverse agents with slightly varied receptors. Diversity is generated by recombination of genetic counterpart, which is due to

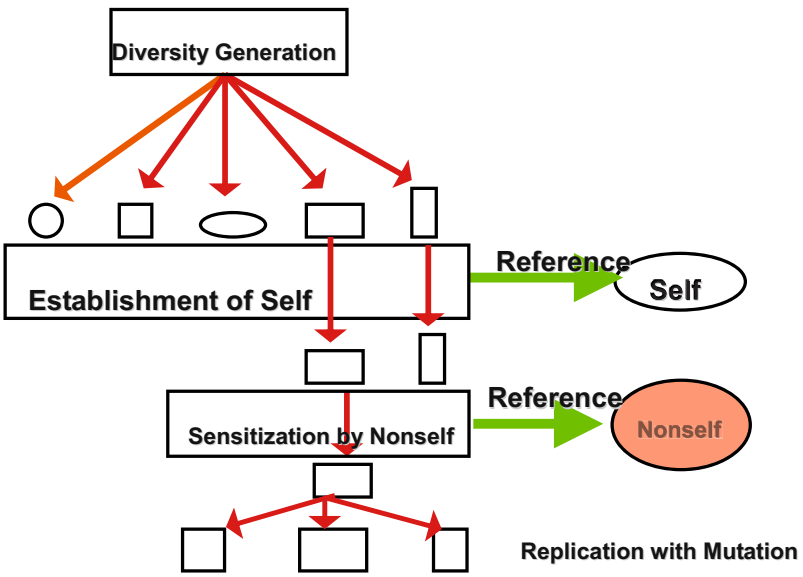

Fig. 7. Utilizing diversity for affinity maturation by agent filtering and agent sensitization [13]

the finding by [25]. In using diversity for exploring further possibility of affinity increase, slight variations of not only structure but also function (affinity) can contribute. For the immune system, the environment with which it must interact is not only the nonself from the outer world, but also the self from the internal world.

Immune algorithms are meant for specific problems where self-nonself discrimination and openness to the environment are critical. Further, the immune algorithm assumes 'agents' as a primitive to build immunity-based systems. In summary, the significance of the immune system used by the immune algorithm is:

- indirect information transfer from the environment by 'selection', as opposed to 'instruction';
- adaptive character driven by continuous diversity generation;
- involvement of self-reference as well as nonself-reference.

An outline of the immune algorithm is depicted in Fig. 7. The algorithm is described in a general context – it is for any adaptive system for self-nonself.

This action part formalized as an immune algorithm has been used for noise cancelation, where noise corresponds to the nonself and the control signal to the self. Since the signal is not labeled beforehand, agents must discriminate the self signal from the nonself one by the specific features of these signals. Further, the cancelation signal from agents must be discriminated for other agents. Although the noise cancelation can apply even to the unknown noise, it must deal with self-reactive agents (that try to cancel the control signal) as if auto-immune disease could happen to the immune system.

Example: Noise Neutralization by Agents

Agents with diverse receptors are first needed. As a set of gene data for initial agents, primitive ones such as shown in Fig. 8 can be used. Diversity may be provided by genetic operations such as recombination. In the simulation, however, ten different gene data with different base lengths but identical heights are used. Since genes will change by adaptation to the noise in the immune algorithm, the initial set of genes may be arbitrary as long as they have variations. However, primitive genes are required so they can approximate many shapes of disturbance signal. During adaptation in the memory of nonself step, genes mutate and higher affinity is attained.

To observe immunologic memory, a noise is first imposed, then the different noise imposed at 15,000 step. Finally, the first noise is again imposed at 30,000 step. Figure 9 shows the response (output from the system) to this noise imposition. It is known that the neutralizer more efficiently neutralizes the noise in the second encounter, if we compare the responses at the initial and second (after 30,000 step imposition of other disturbances) encounter. This

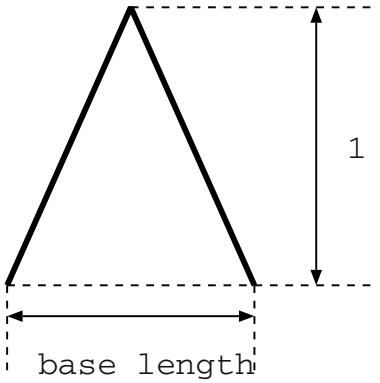

Fig. 8. Initial gene data; ten different base lengths are prepared initially [15]

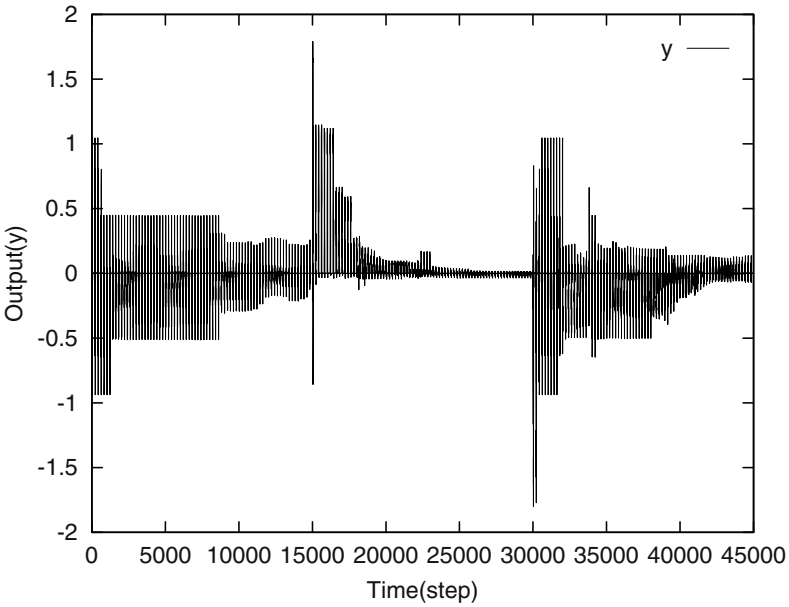

Fig. 9. Time evolution of error when first encounter with the disturbance of a type at step 0 and again at step 30,000, after imposition of a different disturbance from step 15,000 to 30,000 [9]

comes from the adaptation of the agent: memory attained by elongation of the lifespan in this case.

For observing the step of 'Establishment of Self-Tolerance' in the immune algorithm, a sine wave is imposed to the reference input, then agents cannot discriminate whether the signal is disturbance or the reference input.

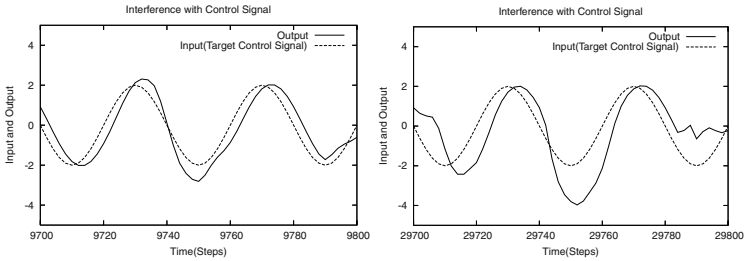

Fig. 10. Response from the system during 100 steps in the early phase (*left*) and the final phase (*right*) in a 30,000-step simulation when the self-reactive agents are not filtered

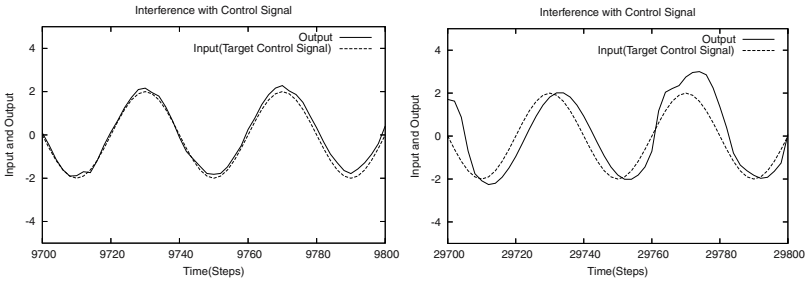

Fig. 11. Response from the system during 100 steps in the early phase (*left*) and the final phase (*right*) in a 30,000-step simulation when the self-reactive agents are filtered

In this and the next simulation, a training phase is added before the noise neutralization.

During the training phase, only the reference input is imposed without noise. In this simulation, agents are not filtered in the training phase. Figure 10 shows the response during 100 steps in the early phase (left: from 9,700 to 9,800 steps) and that in the final phase (right: from 29,700 to 29,800 steps) in a 30,000-step simulation in the noise neutralization phase after training. Noise is not well neutralized due to the self-reactive agents.

In another simulation, the self-reactive agents are removed at the training phase. After this training phase, the neutralizer is placed at the same environment as the previous simulation. Figure 11 shows responses both in the early and final 100, as in Fig. 10. In the final phase (righthand plot of Fig. 11), it is observed that the noise is well canceled while preserving the self (that is, the target signal). We also observe, however, that the self-reactive agents will appear after long time steps, due to the affinity increasing by mutation of existing agents. This would suggest that the self-reactive agents should not only be removed during the training phase, but also memorized at this

phase so that agents similar to the memorized one (hence self-reactive) will be removed whenever they appear by mutation.

In this simulation, one extreme of the IMBS is used – that is an adaptive system (open to the environment) where agents will evolve by adapting to the exogenous nonself. However, agents could form a network by communication and cooperation both in elimination of the disturbance and in memorizing the disturbance pattern.

For example, if the neutralizing signal can affect the other agents, then it would be close to the network model (networked recognition). Further, if the neutralizing signal from agents can be an error signal to other agents, then agents may be connected by signal similarly to Jerne's network [17].

As heuristics for solving problems by adaptive recognition, the following remarks apply:

- Signals (phenotype) should be mapped to gene data (genotype) by decomposing and expressing signals as a primitive signal such as a triangular (Fig. 8) or any other form (such as those found in wavelets), allowing coding of signals where genetic operations are possible;
- Reference to the self as well as nonself should be carefully designed, allowing for the possibility of either or both changing.

Compared with other population-based methods, adaptive recognition can handle not only changing nonself but changing self. This feature also leads to the risk of 'auto-immune disease'. As an application to intrusion detection, adaptive recognition can handle intrusion from both inside as well as outside, by preparing and diversifying profiles for legitimate users as well as illegitimate ones. That is, profiles of legitimate users within the firewall can be taken and processed not only to identify legitimate users (the self) but also to identify illegitimate masqueraders, by diversifying the profiles and even synthesizing the profiles of non-legitimate users (by mutating and recombining available profiles). Here, profiles are any signature that can be obtained by monitoring activities during login. This would provide the possibility of trade-off between internal masqueraders and external masqueraders, other than that between false positive and false negative. Intrusion detection (or equivalently legitimate user identification) becomes more important in the era of ubiquitous computing.

4 Antibody-Based Computing: Arrayed Recognition

[1] demonstrated that Hamiltonian circuits can be achieved by DNA-based computing. Many researchers established that not only DNA but also other macro molecules could have computational capability comparable to DNA. For example, protein-based computing had been proposed by [10] and extended by [2].

Antibody-based computing has a possibility of extension to an immunity-based problem solver that incorporates not only specific recognition of antibodies but adaptive nature supported by diversity generation, selection, and reinforcement of the selected antibodies.

4.1 Definition of Antibody-Based Computing

The immune system is capable of recognizing even artificially synthesized substances. Also, it can further classify substances into the self (those derived from the individual) and oneself. Among those bearing recognition capabilities, antibody is no doubt bearing important component and has been studied in great detail.

Similarly to the DNA-based computing, antibody-based computing utilizes affinity between macro molecules: antibodies. Since the computational capabilities that DNA-based computing could be inherited to antibody-based computing, we rather focused on the difference between them.

Affinity between antigens and antibodies can be measured and their intensities can be ordered (as formatted in an affinity matrix). That is, in contrast to **Matching**$(DNA_i, DNA_j) = 1$ (matched) 0 (not matched), **Affinity**$(Antigen_i, Antibody_j)$ can vary from 0 (no agglutination) to 1 (highest agglutination). This difference would suggest that antibody-based computing could be more general in expressing and solving problems. Also, error tolerance that could be implemented more directly than the DNA-based computing.

4.2 Solving a Combinatorial Problem: The Stable Marriage Problem

The stable marriage problem (SMP) [8] assumes n men and n women, with each member having preference lists of members of the opposite sex. A pair of a man M_i and a woman W_j is called a *blocking pair* if they are not pair in the current solution, but M_i prefers W_j to their current partner, and W_j prefers M_i to their current partner as well. A matching between men and women with no such blocking pair is called *stable*.

Let us consider the stable marriage problem by antibody-based computing. The stable marriage problem can be mapped to antigen-antibody reaction so that preference order of each person in SMP will be reflected in the affinity level between an antibody and an antigen. It should be remarked that agglutination process could be any agglutination (not necessarily between antibodies and antigens) if their affinity levels are measurable and ordered. After agglutinogen and agglutinin are adequately arranged, the solution of SMP will emerge by observing concentration of the agglutination.

Although obtaining a stable matching shows some computational power, it can be solved in $O(N^2)$ time, where N is the size of men (and women). A well-known algorithm exists for giving stable matching for man-oriented matching or woman-oriented one [7, 12]. By further assuming that the concentration observed at a cross-point can reflect the amount of antibodies imposed, the array is capable of obtaining any stable matching in the array from the man-oriented (man optimal and woman pessimal) matching to the woman-oriented (woman optimal and man pessimal) one. By regulating the quantities of all the antibodies $AbM_i(i = 1...n)$ (or equivalently antigens $AgW_j(j = 1...n)$, from a unit to α, the matching would become close to the man-oriented one. Similarly, increase of $AbW_i(i = 1...n)$ will bias the matching towards the one of woman-oriented one.

4.3 Mapping the Stable Marriage Problem to Antibody-Based Computing

Mapping a combinatorial problem to antibody-based computing can be done by composing antigen-antibody compounds corresponding to a problem entity. As for the stable marriage problem, the entity is an individual corresponding to a man or a woman. Antibodies and antigens for a compound corresponding to a particular individual will be determined by considering her(his) preference list over men(women).

Let us consider a scheme for synthesizing antigen-antibody compounds that realize mapping from given preference lists to the compounds. If the woman W_i prefers the man M_j to other men, the compound corresponding to W_i must contain antibody AbW_i and the compound corresponding to M_j contains antigen AgM_j that satisfies $\mathbf{Aff}(\mathbf{AbW_i}, \mathbf{AgM_j})$ being highest among other $AgM_j(j = 1...n)$. If M_j is second in the preference list of W_i, then $\mathbf{Aff}(\mathbf{AbW_i}, \mathbf{AgM_j})$ must be second highest, and so on. AgM_j must realize the order from women W_k other than W_i, hence the affinity $\mathbf{Aff}(\mathbf{AbW_k}, \mathbf{AgM_j})$ must realize the order accordingly (if AgM_j alone cannot realize the order, then new antigen realizing the order must be added to the corresponding compound). Constraints for selecting antibodies and antigens for a compound corresponding to a person can be summed up as follows:

- $\mathbf{Aff}(\mathbf{AbW_i}, \mathbf{AgM_j}) > \mathbf{Aff}(\mathbf{AbW_i}, \mathbf{AgM_k})$ if the woman W_i prefers M_j to M_k in her preference list; and
- $\mathbf{Aff}(\mathbf{AbM_i}, \mathbf{AgW_j}) > \mathbf{Aff}(\mathbf{AbM_i}, \mathbf{AgW_k})$ if the man M_i prefers W_j to W_k in his preference list.

Let us next consider how to solve SMP with an array format. In the array shown in Table 2, row i and column j correspond to the compound for man i (namely, AbM_i and AgM_i), and that for woman j (that is, AbW_j and AgW_j). In other words, at the cross-point ij, two antigen-antibody reactions between AbM_i and AgW_j (reflecting man i's preference), and between AbW_j and AgM_i (reflecting woman j's preference) will take place.

Table 2. Arrayed compounds to solve the stable marriage problem: $M_i(W_i)$ stands for the compound for a man i (woman j), the symbol * at the ij cross-point indicates that M_i and W_i is selected as a stable pair due to a high affinity (each row and each column has only one pair [15])

Compounds	M_1	M_2	$\cdots$	M_i	$\cdots$	M_n
W_1						
W_2						
$\vdots$						
W_j				*		
$\vdots$						
W_n						

Table 3. Landsteiner's ABO blood group system [15]

Blood Type	A	B	AB	O
Antigen (agglutinogen)	A	B	A,B	none
Antibody (agglutinin)	β	α	none	α, β

Under the assumption that the concentration observed at each cross-point is proportional to both $\mathbf{Aff}(\mathbf{AbM_i}, \mathbf{AgW_j})$ and $\mathbf{Aff}(\mathbf{AbW_j}, \mathbf{AgM_i})$, the array can find a stable matching by selecting one cross-point with highest concentration from each row and column. This matching is certainly stable one, for suppose otherwise there must be a blocking pair M_k and W_l such that $\mathbf{Aff}(\mathbf{AbM_k}, \mathbf{AgW_l}) > \mathbf{Aff}(\mathbf{AbM_k}, \mathbf{AgW_p}(\mathbf{M_k}))$ and $\mathbf{Aff}(\mathbf{AbW_l}, \mathbf{AgM_k}) > \mathbf{Aff}(\mathbf{AbW_l}, \mathbf{AgM_p}(\mathbf{W_l}))$, where $p(M_k)$ denotes a partner of M_k in the current matching. Then both concentration at the cross-point k_l is higher than those of $k_p(M_k)$, and those of $p(W_l)l$ reflecting the affinity level.

Example: A Trivial two-by-two Stable Marriage Problem

Landshteiner's ABO blood group system [19] may be used as an example of antibody-based computing. His blood type system is based on antigens (as agglutinogen) on red blood cells and antibodies (as agglutinin) in the blood serum. Table 3 shows agglutinogen and agglutinin of each blood type. Affinity between antibody and antigen is shown in Table 4. Table 5 indicates the well-known incompatible transfusion among the blood type A, B, AB, and O.

Table 4. Affinity matrix: a circle indicates that the antibody-antigen reaction would occur if the antibodies in the column meet with antigens in the row

Antigen; Antibody	A	B
α(anti-A)	O	
β(anti-B)		O

Table 5. Agglutination when the blood type in the column is transfused with the blood type to the row: a circle indicates that the blood type of the column when transfused to that of the column would agglutinate (a double circle indicates an agglutination higher than circles)

Blood Type	A	B	AB	O
A		O		O
B	O			O
AB	O	O		◎
O				

Table 6. A trivial preference list for the two-by-two stable marriage problem

	M_1	M_2		W_1	W_2
W_1	1	2	M_1	1	2
W_2	2	1	M_2	2	1

In this example, we map the relation the woman W_i (the man M_i) prefers the man M_j (the woman W_j) to other to the relation that if the blood of W_i (M_i) would be agglutinate when the blood of M_j (W_j) were transfused. That is, if the woman W_i prefers the man M_j most, the blood type should be so assigned that the type for W_i comprises of antibody AbW_i and antigen AgW_i; and that for M_j of antibody AbM_j and antigen AgM_j and the affinity $\mathbf{Aff}(\mathbf{AbW_i}, \mathbf{AgM_j})$ are highest.

For the trivial case when the preference lists of men and women are as per Table 6, simple assignment would suffice: a man to type A and another man to type B; for the woman who prefers a man with type A to type B, and for another woman type A (Fig. 11). It should be noted that assignment to A for two men and to B for two women would not work, since the assignment does not reflect the preference of men and women.

In the nontrivial preference list shown in Table 7, one assignment would be type O to both M_1 and W_1, type A to M_2, and type B to W_2 (Fig. 13).

Table 7. A non-trivial preference list for the two-by-two stable marriage problem [23]

	M_1	M_2		W_1	W_2
W_1	2	1	M_1	2	1
W_2	2	1	M_2	2	1

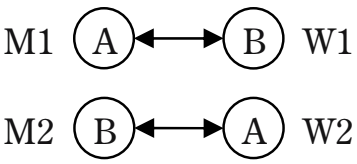

Fig. 12. A blood type assignment reflecting the preference [15]

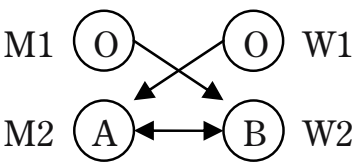

Fig. 13. A blood type assignment reflecting the preference of Table 7 [23]

For other two preference lists (with a different graph topology than that of Figs. 12 and 13), it is not possible to map the blood type with the above correspondence, and other compounds should be synthesized for realizing the preference lists.

We have shown that antibodies, a macro molecule of the immune system, with specific recognition capability could be used for computation as DNA are used for DNA computing. We suggest a possibility of extending antibody-based computation to the solver, rather than showing its computational capability or universality. The application aims not at replacing current electronic computers, but rather at being a supportive tool for bioinformatics.

5 Toward a General Problem Solver: Immunity-Based Problem Solver

Problem solving by Means-Ends Analysis (MEA) [21] organizes a search in a dynamically constructed search space of problem-subproblem decomposition. It embodies and simulates human problem solving by the recognition-action cycle shown in Fig. 14, where solid arcs are recognitions and white arcs are actions. An important feature of MEA is that application of operators is not

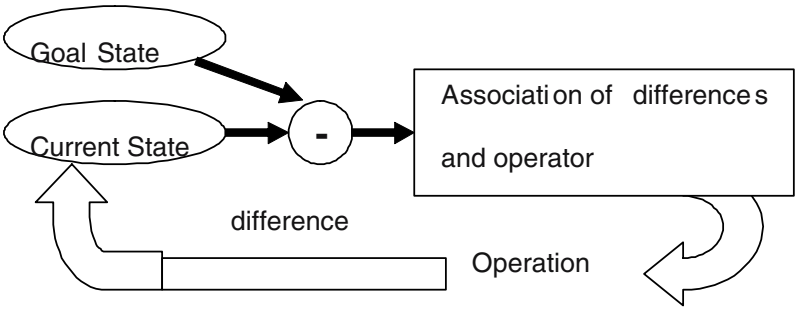

Fig. 14. Problem solving by Means-Ends Analysis (MEA) [15]

very rigid: if an operator selected in the heuristics part is not directly applicable to the current problem, the problem will be divided into subproblems. This flexibility allows a certain degree of freedom in identifying the heuristics, and further contributes to generality in the problems that MEA can handle. In fact, MEA had been implemented as General Problem Solver (GPS) that can deal with many well-known puzzles [6].

As an intermediate stepping stone from general problem solving by MEA to immunity-based problem solving, let us briefly investigate more general biological problem solving. It should be first emphasized that the following discussions and Fig. 15 is for bridging purposes between MEA and immunity-based problem solving, and hence may be rough and approximate. Here solid arcs are recognitions, and white arcs actions; recognition action cycles are iterated until there is no difference between the goal state and the current state. One difficulty is that a unit of biological system such as DNA, cells, individuals, and species do not constitute a usual hierarchical system understood in component-system relation found in most artificial systems. Another difficulty, hence making the biological problem solving remarkable, is that biological systems use concepts distinct from those used in artificial systems to realize robustness (robustness is a solution implemented and embedded in the system by a biological problem solving to challenges to system survival). Biological problem solving utilizes the variations for implementing robustness, although artificial problem solving considers the variations as disturbance and trying to prevent them from occurring and minimizing the effect. Thus, the 'difference' in Fig. 15 of biological problem solving can be quite different from that shown in Fig. 14.

The third difficulty that makes biological problem solving more complicated and entangled is that the given problem is a challenge to the survival of the problem solver (the biological system) itself. This self-referential aspect must be paid attention in capturing the immunity-based problem solving shown in Fig. 16.

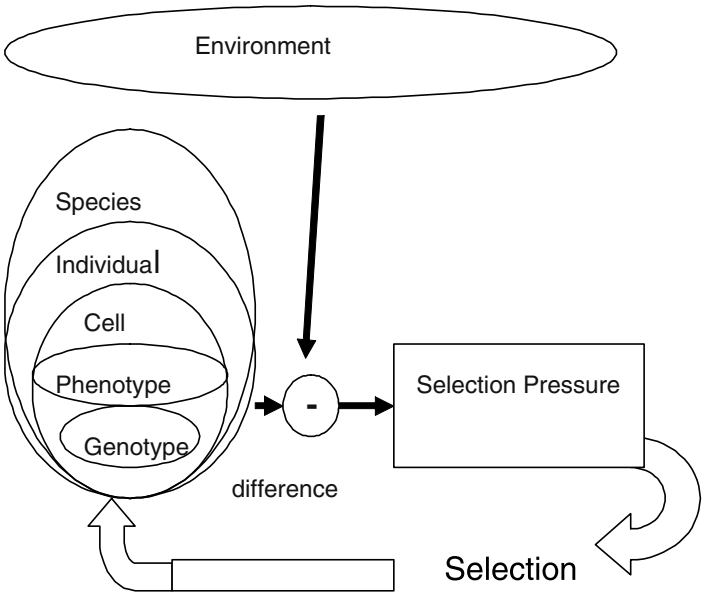

Fig. 15. Problem solving by Means-Ends Analysis (MEA)

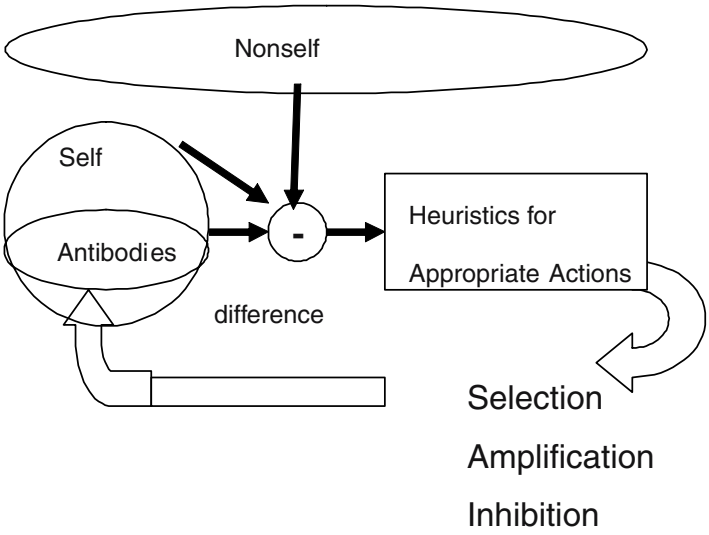

Fig. 16. The immune system as a problem solver: only antibodies are focused

Throughout Fig. 14 (MEA), Fig. 15 (biological problem solving) and Fig. 16, the framework for problem solving is to recognize differences and deploy actions based on these differences. However, actions are oriented toward the system itself for both biological and current immunity-based problem solv-

ing, hence the process is intrinsically adaptation (or in Fig. 15, evolution). Figure 16 and the following discussion focuses only on the immune system, involving antibodies, hence that of adaptive immunity.

In means-ends analysis on the one hand, the problem solving process constitutes an intrinsic part of the solution. That is, the *order* of operator applied is a critical part of the solution of a given puzzle. Thus, the problem is fixed throughout the problem solving; hence, the solver deals with a static problem.

On the other hand, in immunity-based problem solving, the problem itself undergoes changes, because the environment, including the nonself, is changing and the solver involving the self must change accordingly. Therefore, there is no complete solution and there will always be a gap between the current solution and the current problem. However, the current solution can be used for the next problem when the next problem (the change) also evolved from the current problem. Problem solving does not have a beginning and an end. The current solution is not good for the current environment because the latter is ever changing; therefore the gap between these two must be compensated for the next solution. However, the next solution is not built from scratch but rather from the current solution. The solution must always chase the environment, which is an online and dynamical adaptation to the dynamical environment. In immunity-based solving, the typical environmental change is either a challenge from the outside (for example, bacteria and viruses) or from inside (say, cancer). To deal with these challenges, the solver (a collection of agents) must prepare a diverse set for being selected by these problems (challenges) and the selected agents must be further increased. Since there is always a difference from the current solution and the current environment, there must be a diversification of agents.

6 Conclusion

This Chapter has explored the possibility of antibody-based computing that use antibodies or peptides in general. We first investigate several types of recognitions by revisiting immunity-based systems. The very primitive form of arrayed recognition is shown to have a computational capability comparable of DNA based computation by taking an example of a combinatorial problem: the Stable Marriage Problem. This would suggest that more sophisticated forms of recognition such as networked or selected recognition will have computational capabilities not only in a static context but in a more dynamic context as seen in the environments which the immune systems face. Thus, main feature of the antibody-based computing is that allows extension to a problem solver, since the immune system is a problem solver embedded in individuals, taking care of challenges to the individuals.

Acknowledgements

I am grateful to Prof. Fulcher, not only for giving me the opportunity for presenting this work, but also for his great assistance in editing and clarifying the chapter. I am also grateful to the anonymous reviewers, whose comments were quite helpful in improving the chapter. Mr. Sasajima, Mr. Mori, Mr. Oohashi, Mr. Sugawara, and Mr. Hiraizumi assisted with the numerical simulations. This work was supported in part by a Grant-in-Aid for Scientific Research (B) 16300067, 2004. This work was also partly supported by the 21st Century COE Program 'Intelligent Human Sensing' of the Ministry of Education, Culture, Sports, Science and Technology of Japan. I am also grateful to Mr. Abe, Mr. Hattori, and Mr. Toyofuku of Design Department No. 21, Electronics Engineering Division II, Vehicle Engineering Group, Toyota Motor Corporation, for incisive discussions on this project. Lastly, I am indebted to Dr. Watanabe who supported the project launch.

References

1. Adleman LM (2004) Molecular computation of solutions to combinatorial problems. *Science,* 266(11): 1021–1024.
2. Balan MS, Krithivasan K (2004) Parallel computation of simple arithmetic using peptide-antibody interactions. *Biosystems,* 76(1–3): 30–307.
3. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R (2005) A synthetic multicellular system for programmed pattern formation. *Nature,* 434: 1130–1134.
4. Burnet FM (1957) A modification of Jerne's theory of antibody production using the concept of clonal selection. *Australian J. Science,* 20: 67–69.
5. Elowitz MB, Leibler SA (2000) Synthetic oscillatory network of transcriptional regulators. *Nature,* 403: 335–338.
6. Ernst G, Newell A (1969) *GPS: A Case Study in Generality and Problem Solving.* Academic Press, New York, NY.
7. Gale D, Shapley L (1962) College admissions and the stability of marriage. *American Mathematical Monthly,* 69: 9–15.
8. Gusfield D, Irving RW (1989) *The Stable Marriage Problem: Structure and Algorithms.* MIT Press, Cambridge, MA.
9. Hood L, Galas D (2003) The digital code of DNA. *Nature,* 421: 444–448.
10. Hug H, Schuler R (2001) Strategies for the development of a peptide computer. *Bioinformatics,* 17(4): 364–368.
11. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature,* 409: 860–921.
12. Irving RW, Leather P, Gusfield D (1987) An efficient algorithm for the optimal stable marriage. *J. ACM,* 34(3): 532–543.
13. Ishida Y (2004) *Immunity-Based Systems: a Design Perspective* Springer-Verlag, Berlin.
14. Ishida Y (2006) Designing an immunity-based sensor network for sensor-based diagnosis of automobile engines. In: Gabrys B, Howlett RJ, Jain LC (eds.) *Proc. Knowledge-Based Intelligent Informaiton and Engineering Systems Conf.*

(KES'2006). Lecture Notes in Computer Science LNCS 4252. Springer-Verlag, Berlin: 146–153.
15. Ishida Y (2007) A constructive systems approach to understanding the immune system as a biological problem solver. In: *Proc. 1st IEEE Symp. Foundations of Computational Intelligence (FOCI'07)*, 1–5 April, Honolulu, Hawaii. IEEE/Omnipress, Madison, WI: 646–650 (CD-ROM).
16. Ishida Y, Adachi N (1996) Active noise control by an immune algorithm: adaptation in immune system as an evolution. In: *Proc. 3rd IEEE Intl. Conf. Evolutionary Computation (ICEC'96)*, 20–22 May, Nagoya, Japan. IEEE Press, Piscataway, NJ: 150–153.
17. Jerne NK (1973) The immune system. *Scientific American,* 229(1): 52–60.
18. Kollmann M, Lovdok L, Bartholome K, Timmer J, Sourjik V (2005) Design principles of a bacterial signalling network. *Nature,* 438(24): 504–507.
19. Landsteiner K (1900) Zur kenntnis der antifermentativen, lytischen und agglutinierenden wirkungen des blutserums und der lymphe. *Zentralblatt Bakteriologie,* 27: 357–362.
20. Langman RE, Cohn M (eds.) *Seminars in Immunology* (available online at: http://www.ingentaconnect.com/content/ap/si/2000/00000012/00000003 – last accessed: April 2007).
21. Newell A, Simon H (1972) *Human Problem Solving.* Prentice Hall, Englewood Cliffs, NJ.
22. Pearl JJ (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference (Revised 2nd Printing).* Morgan Kaufmann, San Mateo, CA.
23. Perelson AS, Oster GF (1979) Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *J. Theoretical Biology,* 81: 645–670.
24. Sprinzak D, Elowitz MB (2005) Reconstruction of genetic circuits. *Nature,* 438(7067): 438–424.
25. Tonegawa S (1983) Somatic generation of antibody diversity. *Nature,* 302: 575–581.
26. Venter JC, Adams MD, Myers EW, et al. (2001) The sequence of the human genome. *Science,* 291: 1304–1351.
27. Volterra V (1931) Variations and fluctuations of the number of individuals in animal species living together. In: Chapman RN (ed.) *Animal Ecology.* McGraw-Hill, New York, NY.

Resources

1 Key Books

Burnet FM (1959) *The Clonal Selection Theory of Immunity.* Cambridge University Press, London, UK.

Dawkins R (1976) *The Selfish Gene.* Oxford University Press, Oxford, UK.

Dennett DC (1995) *Darwin's Dangerous Idea: Evolution and the Meaning of Life.* Simon & Schuster, New York, NY.

Edelman GM (1992) *Bright Air Brilliant Fire: On the Matter of the Mind.* Basic Books, New York, NY.

Goldberg DE (1989) *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, Reading, MA.

Holland J (1975) *Adaptation in Natural and Artificial Systems.* University of Michigan Press, Ann Arbor, MI.

Ishida Y (2004) *Immunity-Based Systems: A Design Perspective.* Springer-Verlag, Berlin.

Janeway CA, Travers P, Walport M, Shlomichik M (2002) *Immunobiology: The Immune System in Health and Disease. (5th ed.).* Garland Publishing, New York, NY.

Maturana H, Varela F (1980) *Autopoiesis and Cognition: The Realization of The Living.* D. Reidel, Dordrecht, Germany.

Paul WE, et al. (1999) *Fundamental Immunology (4th ed.).* Lippincott-Raven, Philadelphia, PA.

Perelson AS (ed.) (1988) *Theoretical Immunology, Part I and II.* Addison Wesley, Reading MA.

Segel LA, Cohen IR (eds.) (2001) *Design Principles for the Immune System and Other Distributed Autonomous Systems.* Oxford University Press, New York, NY.

Tauber AI (1997) *The Immune Self.* Cambridge University Press, Cambridge, UK.

Varela FJ (1979) *Principles of Biological Autonomy.* Elsevier/North-Holland, New York, NY.

von Neumann JJ (1966) *Theory of Self-Reproducing Automata.* In: Burks AW (ed.) University of Illinois Press, Urbana, IL.

von Neumann JJ (1956) Probabilistic logics and the synthesis of reliable organisms from unreliable components. In: Shannon, C.E., McCarthy, J. (eds.) *Automata Studies.* Princeton University Press, Princeton, NJ: 43–98.

2 Key Survey/Review Articles

Farmer JD, Packard NH, Perelson AS (1986) The immune systems: adaptation and machine learning, *Physica D,* 22: 187–204.

Forrest S (1993) Genetic algorithm: principles of natural selection applied to computation. *Science,* 261: 872–878.

Jerne NK (1973) The immune system. *Scientific American,* 229(1): 52–60.

Jerne NK (1974) Clonal selection in a lymphocyte network. *Society General Physiologists Services,* 29: 39–48.

Jerne NK (1985) The generative grammar of the immune system. *The EMBO J.,* 4(4): 847–852.

Perelson AS, Weisbuch G (1997) Immunology for Physicists. *Review of Modern Physics,* 69: 1219–1267.

Tonegawa S (1985) The molecules of the immune system. *Scientific American,* 253(4): 122–131.

3 Organisations, Societies, Special Interest Groups

Santa Fe Institute
http://www.santafe.edu/

4 Research Groups

Artificial Immune Systems at University of York
http://www-users.cs.york.ac.uk/jtimmis/

Artificial Immune Systems at the Center for Information Assurance and Intelligent Security Systems Research Lab
http://www.cs.memphis.edu/~dasgupta/

Immunity-Based Systems at TUT Systems Sciences Lab
http://www.sys.tutkie.tut.ac.jp/~ishida/en/index.html

5 Key International Conferences/Workshops

FOCI 2007: The 1st IEEE Symposium on the Foundations of Computational Intelligence
http://events.cs.bham.ac.uk/foci07/index.php

ICARIS 2007: International Conference on Artificial Immune Systems
http://lsin.unisantos.br/icaris2007/

IMBS-KES 2007: Special Session on Immunity-Based Systems
http://www.sys.tutkie.tut.ac.jp/~ishida/IMBS-KES07CFP.php.htm

Knowledge-Based and Intelligent Engineering Systems Conf. (KES2007)/ XVIIth Italian Workshop on Neural Networks (WIRN2007)
http://ra.crema.unimi.it/AIS2007/