Discrimination-based Artificial Immune System:
Modeling the Learning Mechanism of Self and Non-self Discrimination for Classification

Kazushi Igawa and Hirotada Ohashi
Department of Quantum Engineering and System Science
Graduate School of Engineering, The University of Tokyo, Japan

Abstract: This study presents a new artificial immune system for classification. It was named discrimination-based artificial immune system (DAIS) and was based on the principle of self and non-self discrimination by T cells in the human immune system. Ability of a natural immune system to distinguish between self and non-self molecules was applicable for classification in a way that one class was distinguished from others. We model this and the mechanism of the education in a thymus for classification. Especially, we introduce the method to decide the recognition distance threshold of the artificial lymphocyte, as the negative selection algorithm. We apply DAIS to real world datasets and show its performance to be comparable to that of other classifier systems. We conclude that this modeling was appropriate and DAIS was a useful classifier.

INTRODUCTION

Biological mechanisms have attractive applications in computer science. Many algorithms, called biologically inspired algorithms, have been proposed; for example, neural networks, genetic algorithms (GA), genetic programming and artificial immune systems (AIS). AIS is the algorithm mimicking the human immune system for solving various problems[1] and has been studied in many fields, including computer security[2], information filtering[3], data clustering[4], concept learning[5], classification[6], etc. These studies show that AIS is effective for data analysis.

One of the main objectives of AIS is to apply the immune metaphor or mechanism to computations[1]. In a natural immune system, recognition of the T cell is one of the most important mechanisms. It allows the immune system to discriminate self and non-self molecules. This characteristic is conducive to machine learning, particularly, classification and concept learning, in terms that one class (self) is distinguished from other classes (non-self).

In AIS study, however, the main topic has been the recognition mechanism and network theory of B cells[4,5]. Although the mechanism of T cells is essential to immune discrimination, there are a few studies in which this concept is applied to classification problems. Most of these studies either over-simplifies fundamental processes[7] or are too complex by featuring GA with many parameters[8,9]. We need to research the generalized modeling of T cells and then the learning mechanism to apply these to machine learning.

We developed new AIS that is applicable to the classification of data sets by incorporating the training and recognition mechanisms of T cells. It is called a discrimination-based artificial immune system (DAIS). In DAIS, we introduced a variable recognition threshold and a new negative selection algorithm that decides the threshold. DAIS employs only immune inspired algorithms with five parameters, of which only one parameter is important. In this paper, we explain the concept of DAIS and its algorithm. Then, we demonstrate its behavior by applying it to real world data sets. The results show that DAIS has comparable properties to other classifier systems and the mechanism of T cell is useful for classification.

IMMUNE SYSTEM

A biological immune system consists of two different response systems: innate and adaptive immune systems. By cooperation of these, the whole immune system defends the body against harmful foreign antigens. The innate system responds to

Corresponding Author: Kazushi Igawa, Dempartment of Quantum Engineering and System Science, Graduate School of Engineering, The University of Tokyo

foreign antigens. This mechanism has not been used in artificial immune systems. In contrast, the adaptive part functions against specific targets, which has been available for artificial immune systems.

The important functions of the adaptive immune system are to recognize, eliminate and memorize foreign antigens, as well as to discriminate between self and non-self molecules. This means that the immune system is able to react only to non-self molecules. The adaptive immune system consists of two kinds of lymphocytes. One is the B cell, which originates in the bone marrow and has the ability to produce antibodies. B cells can recognize the shape of foreign antigens, eliminate and memorize them. The other is the T cell, which can discriminate between self and non-self molecules in a body. T cells acquire the ability to discriminate by education in an organ named the thymus.

What happens in the thymus? If the T cell recognizes self-molecules, it causes apotosis and dies. This is called negative selection. If it does not react, it is tested with non-self molecules. If the T cell is not able to recognize them, it dies. So, T cells that can react with non-self molecules survive. This is positive selection. T cells that can pass through the thymus are able to react to non-self molecules but unable to react to self molecules (Fig. 1).

After training in a thymus, T cells circulate throughout the body. When B cells react with foreign antigens, they need the help of T cells. B cells do not react with self molecules if T cells do not recognize it. As a result, the adaptive immune system does not attach self-molecules.

DISCRIMINATION-BASED ARTIFICIAL IMMUNE SYSTEM

We prepare a discrimination-based artificial immune system (DAIS), which consists of artificial lymphocytes (ALC), a learning algorithm and a classification method. In this section, first, we describe the idea which is inspired by education in a thymus. Second, we prepare the algorithm to apply this idea to classification, which is the basic algorithm of DAIS. Third, the whole learning algorithm of DAIS is described and finally, we explain the classification method of DAIS.
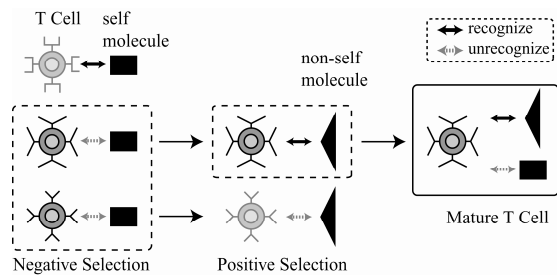

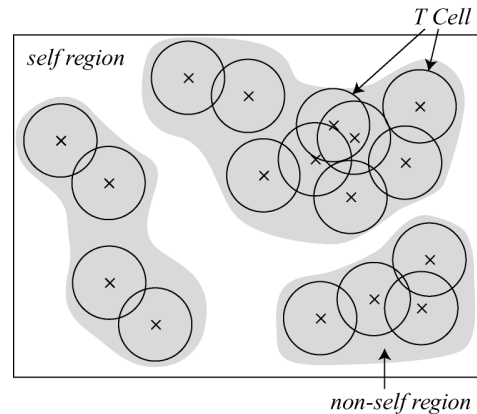Fig. 1: Education of T cells in a thymus
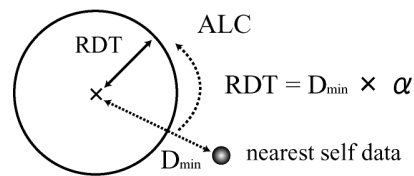

Fig. 2: Self and non-self regions in shape space

$$RDT = D_{min} \times \alpha$$

Fig. 3: Determination of RDT

Self and non-self regions in the shape space: Modeling of the immune receptor is one of the most important items, because the immune reaction is based on recognition via the receptor. In AIS, it is represented in shape space[1]. It is a phase space whose axes represent the properties of the receptor; for example, the length, width, height of any bump or groove in the combining site, electrostatic charge, etc. This means that the position in shape space represents the shape of a receptor. The distance between two positions is the affinity between these shapes. The smaller the distance, the higher the affinity. If the distance is smaller than a threshold, these interact with each other.

To apply education in a thymus to classification,

we need to understand how to represent this mechanism in shape space. There are two kinds of region in shape space. Thymal education makes it possible for survived T cells to be distributed in only one region called the non-self region (Fig. 2). The other region is named self region. Negative selection removes T cells from the self-region and positive selection chooses T cells distributed in the non-self region. This is very useful for classification, because the data set belonging to the same class becomes clusters.

Basic learning algorithm: In this subsection, we describe how to apply the above idea to classification. This algorithm consists of three big parts: self and non-self classes, artificial lymphocyte and learning processes. Also, the data set is normalized.

Self and non-self classes: To apply the self and non-self regions and the mechanism of dividing shape space to classification, it is necessary to divide the data set into two classes: self and non-self classes. Self-class consists of only one class in the problem and non-self class consists of all the remaining classes. The immune system categorizes self-molecules as self-class and all other molecules as non-self class. Data belonging to self-class are named self data and those in the non-self class are named non-self data.

Artificial lymphocyte: In the immune system, each T cell recognizes these molecules. AIS uses artificial lymphocytes (ALC) to recognize data individually in the same way. The ALC consists of a receptor, a recognition distance threshold (RDT) and a self class.

The receptor is represented in shape space. Its definition is such that, if there is a distance (Euclidian distance, hamming distance, etc.) between an ALC and individual data, the ALC can recognize it.

Self class of an ALC is one of the classes. The remaining classes are non-self class for the ALC. Self class is decided in a learning process (next subsection). If class i is a self class in a learning process, self class is i. Then, ALC_i represents the ALC who's self-class is i and the ALCs whose self class is the same are named the same kind of ALC.

Learning processes: Next, we describe the algorithm to acquire the ALC_i set so that these ALC_i cannot recognize the data belonging to their own self class (i). This is inspired by education in a thymus and

consists of iteration of some processes. These processes are proliferation and mutation (clonal expansion), determination of RDT (negative selection) and activation and selection (positive selection). After passing through these processes, the ALC_i is named memory ALC_i.

In the proliferation and mutation process, one ALC_i produces clones. Mutation of a cell is done in the immune system and it is important to use this for computation. Changing the receptor of ALC partially makes a clone. This process is the same as the evolutionary computations and is known to be effective.

In negative selection, we reverse the idea. We don't choose the ALC that cannot recognize the self data set, but decide the RDT of ALC so that it cannot recognize the self data set. The reason is the computational cost. With a real thymus, only about 3% can survive among an enormous number of T cells and this is costly. RDT is determined by multiplying the minimum distance with the self data set by the constant number α equal to or less than 1 (Fig. 3). This ensures that ALCs cannot recognize the self-data that are used as the training data. It introduces the variable recognition threshold of ALC and a new negative selection algorithm. Parameter α is the most important in DAIS.

In positive selection, an ALC_i is selected in order to perform two kinds of activity. To decide the activities, the non-self data set is divided into two kinds. One is the data that are not recognized by the memory ALC_is and the other is the data recognized by them, which are represented as U_i and R_i, respectively. The primary activity of an ALC_i is the number of U_i that it can recognize. The secondary activity of an ALC_i is the number of R_i that it can recognize. An ALC_i is selected in the order of the primary activity. If some ALC_is have the same activity, they are selected in the order of the secondary activity. If some ALC_is have the same secondary activity, they are selected in the order of RDT. The larger RDT an ALC_i has, the wider it can cover. The remaining ALC_is are removed. These processes make an ALC_i more effective.

The above processes are repeated to obtain the more useful ALC_i. It is the only one allowed to pass (Fig. 4). This is repeated a predetermined number of times. The last selected ALC is stored as a memory ALC. The learning process is as follows in detail.
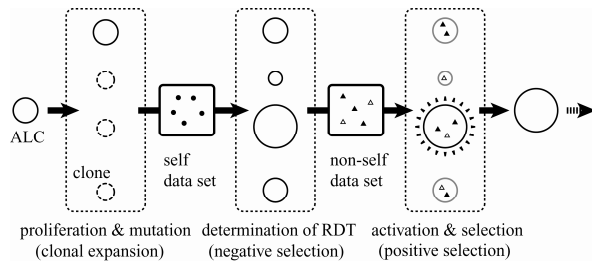
Fig. 4: Schematic of the learning process

Through three big processes with one ALC, we get the more effective ALC. First, in the proliferation and mutation process, the ALC produces its clones. Next, the self data set is presented to these clones and their RDT is determined. In the final process, they are activated by non-self data set and the best ALC is selected.

1. Set class i to self class and the remaining classes to non-self class.
2. Select one among the non-self data set, which the memory ALC_is have not recognized.
3. Make an ALC_i whose receptor is the same as a selected non-self data.
4. Make 100 clones of the ALC_i. (Proliferation)
5. Mutate the receptor of these clones. Each element of the receptor is selected according to predetermined mutation probability (0.2) and given a random value ([-0.1,0.1]). (Mutation)
6. Determine the RDT of these ALC_i by presenting the self data set. (Negative selection)
7. Set two kinds of activity to evaluate the number of U_i and R_i that each ALC can recognize. These numbers are the primary and secondary activity. Remove the ALCs that cannot recognize the data selected in Step 2. (Activation)
8. Select the ALC whose primary activity is the largest. If more than one ALC has the same activity, select the ALC whose secondary activity is the largest among them. If more than one ALC has the same activity, select the ALC which has the biggest RDT. Then, remove the rest. (Selection)
9. If the number of repetitions from Step 4 to Step 8 reaches 100, go to Step 10. If not, return to Step 4.
10. Store the selected ALC in Step 8 as a memory ALC_i.
11. If all non-self data are recognized by the set of the

memory ALC_i, the learning is complete. Otherwise, return to Step 2.

Demonstration: We demonstrate the behavior of the basic learning algorithm and investigate its recognition mechanism visually. This algorithm is applied to two kinds of artificial data set: linear and non-linear separable data sets. Both of them are two-dimensional
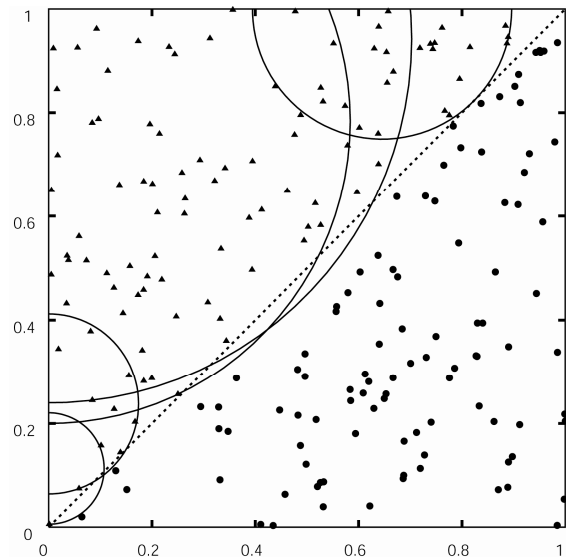

Fig. 5: Result for linear separable data set

The open circles represent memory ALCs. For these memory ALCs, the closed circles are the data which belong to a self class and the closed triangles belong to a non-self class. The memory ALCs are distributed in the non-self region. These figures show that the learning algorithm is useful to obtain one kind of ALC.
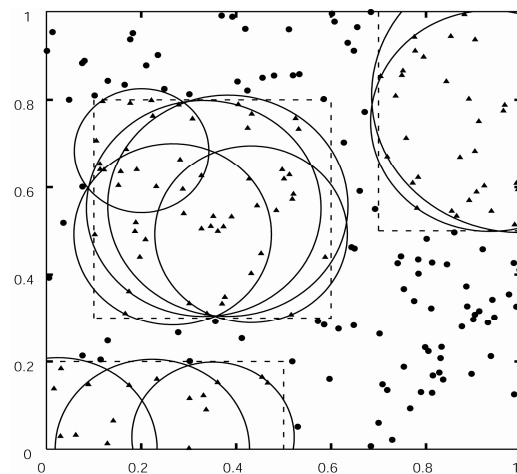

Fig. 6: Result for non-linear separable data set

two-class problems. They are trivial, but our purpose here is to demonstrate their behavior.

Figure 5 is the result for linear separable data set and Fig. 6 is one for non-linear separable data set. Closed circles and triangles represent the data belonging to class 1 and 2, respectively. Class 1 is self class. Open circles are the memory ALC_1s. These show that the memory ALC_1 is distributed in the non-self class region (class 2). Therefore, the prepared learning algorithm allows the distribution of the memory ALC in a non-self region.

Whole learning algorithm: Generally, the number of classes in a classification problem is more than three. Therefore, DAIS needs to acquire all kinds of memory ALC_i ($i \in \{1, 2, \cdots, N_c\}$, N_c: the number of class) and classify the unknown data using these memory ALCs. The method to get all kinds of ALC is to repeat the basic learning algorithm (section 3.2) N_c times. At first, class 1 becomes self class. Next, class 2 becomes self class. DAIS repeats this until N_c becomes self class. This is the whole learning algorithm of DAIS.

After the learning process, there are all kinds of memory ALC_is that can distinguish between one class (their own self class) and the remaining classes.

Classification method: In the classification process, all memory ALCs are used. The classification method is an elimination process. If data are recognized by an ALC_i, the data do not belong to the self class of ALC_i (class i). If only memory ALC_i cannot recognize a data, the class is determined as k. For example, in the case of three classes problem ($i \in \{1, 2, 3\}$), if data are recognized by both ALC_0 and ALC_1, then they belong to class 3 (Fig. 7, left).

This method has two problems. One is that the data are not recognized by more than two kinds of ALC. This case is named hole (Fig. 7, center). The other is that the data are recognized by all kinds of ALC. This case is named overlap (Fig. 7, right).

In the case of hole, DAIS sets a new threshold (*rdt*) to the memory ALCs that cannot recognize data. *rdt* is the value obtained by multiplying RDT with w ($rdt = RDT \times w$, $w > 1$). If there is only one kind of ALC which cannot recognize data, then its own self class is the class of the data. Otherwise, w increases ($w = w + \delta$) and *rdt* is calculated again. This process is repeated until the above-mentioned condition is satisfied. However, if *rdt* is too big and all kinds of memory ALCs can recognize data (overlap), w
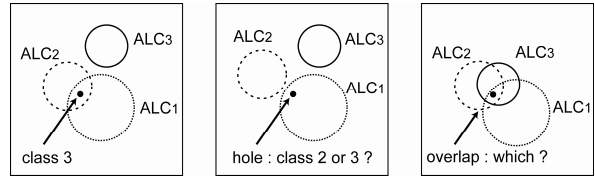

Fig. 7: Three patterns of DAIS's recognition: the left is a normal case, the center is hole and the right is overlap

Table 1:   Data sets

	features	class	data
Iris	4	3	150
Ionosphere	34	2	351
Diabetes	8	2	768
Sonar	60	2	208

decreases ($w = w - \delta$) and δ is made half ($\delta = \delta \times 0.5$). Then, the above process is repeated.

In the case of overlap, DAIS sets a new threshold (*rdt*) to the memory ALCs that can recognize data. *rdt* is the value obtained by multiplying RDT with w ($rdt = RDT \times w$, $w < 1$). If there is only one kind of ALC that cannot recognize data, then its own self-class is the class of the data. Otherwise, w decreases ($w = w - \delta$) and *rdt* is calculated again. This process is repeated until the above-mentioned condition is satisfied. However, if *rdt* is too small and more than two kinds of memory ALCs cannot recognize data (hole), w increases ($w = w + \delta$) and δ is made half ($\delta = \delta \times 0.5$).

RESULTS

Accuracy: We applied DAIS to four real world data sets and compared these results with those of other classifier systems. Data sets chosen are Iris, Ionosphere, Diabetes and Sonar. These sets are all obtained from the UCI machine learning repository[10]. Iris and Diabetes have a few features, but Ionosphere and Sonar have many features. Their details are as follows (Table 1).

Table 2 shows the accuracy and ranking of DAIS and the other classifiers[6,11]. These were obtained by averaging ten runs of DAIS. For Iris, Diabetes and Sonar data sets, a five, ten and thirteen-fold cross validation scheme were employed, respectively. For Ionosphere data set, the first 200 items were chosen for the training set and the remaining 151 items for test set. DAIS compares well with some of the best general purpose classifiers available. Especially, for Ionosphere and Sonar data sets, DAIS showed high

Table 2: Average accuracy and comparison

	Iris		Ionosphere		Diabetes		Sonar	
1	Grobian (rough)	100.0%	3-NN + simplex	98.7%	Logisc	77.7%	TAPMFTBasyesian	92.3%
2	SSV	98.0%	**DAIS**	**96.8%**	IncNet	77.6%	Naïve MFT Bayesian	90.4%
3	C-MLP2LN	98.0%	3-NN	96.7%	DIPOL92	77.6%	SVM	90.4%
4	PVM2 rules	98.0%	IB3	96.7%	Linear Discr. Anal	77.5%	Best 2-Layer MLP + BP	90.4%
5	PVM1 rule	97.3%	MLP + BP	96.0%	SMART	76.8%	**DAIS**	**89.1%**
6	AIRS	96.7%	AIRS	94.9%	GTO DT (5xCV)	76.8%	MLP + BP, 12 hidden	84.7%
7	FuNe-I	96.7%	C4.5	94.9%	ASI	76.6%	MLP + BP, 24 hidden	84.5%
8	NEFCLASS	96.7%	RIAC	94.6%	Fischer discr. Anal	76.5%	1-NN, Manhatten	84.2%
9	CART	96.0%	SVM	93.2%	MLP + BP	76.4%	AIRS	84.0%
10	**DAIS**	**95.8%**	Non-linear perceptron	93.0%	LVQ	75.8%	MLP+BP, 6 hidden	83.5%
11	FUNN	95.7%	FSM + rotation	92.8%	LFC	75.8%	FSM - methodology?	82.6%
12			1-NN	92.1%	RBF	75.7%	1-NN Euclidean	82.2%
13			DB-CART	91.3%	NB	75.5%	DB-CART, 10xCV	81.8%
14			Linear perceptron	90.7%	kNN, k=22, Manh	75.5%	CART, 10xCV	67.9%
15			OC1 DT	89.5%	MML	75.5%		
16			CART	88.9%	SNB	75.4%		
17			GTO DT	86.0%	BP	75.2%		
…					…			
22					AIRS	74.1%		
23					**DAIS**	**73.4%**		
24					C4.5	73.0%		

11 other reported with lower scores

accuracy and its rank was high; proof that DAIS is useful for pattern classification with a lot of features.

To obtain these results, we changed only parameter α. The value of α is 0.85, 0.8, 1.3 and 0.65 for Iris, Ionosphere, Diabetes and Sonar data sets, respectively. The best value is dependent on the problem. However, it is easy to optimize only one parameter.

We also compared DAIS with artificial immune recognition system (AIRS), which is a classification algorithm based on the mechanism of B cell's recognition[6]. For Iris and Diabetes, the accuracy of DAIS is about 1% lower than AISR's. However, for Ionosphere and Sonar data sets, DAIS's accuracy is about 2% and 5% higher than AIRS's. If the problem involves features, DAIS is more useful than AIRS. This means that the mechanism of T cell discrimination is as useful as that of B cell for classification.

Now, we discuss the accuracy of the training set. DAIS is applied to the satellite image data set (STATLOG version) obtained from the UCI machine learning repository. This offers 36 features and 5 classes. In addition, it provides a training set with 4435 items and a test set with 2000 items. The result

Table 3: Average accuracy for training and test data sets of satellite image

|	Satellite image	Accuracy % (training)	Accuracy % % (test)
1	MLP+SCG	96.0	91.0
2	k-NN	-	90.9
3	k-NN	91.1	90.6
4	**DAIS**	**100.0**	**90.5**
5	k-NN	-	90.4
6	k-NN	-	90.0
7	FSM	95.1	89.7
8	LVQ	95.2	89.5
9	k-NN	-	89.4
10	Dipol92	94.9	88.9
11	MLP+SCG	94.4	88.5
12	SVM	91.6	88.4
13	Radial	88.9	87.9
14	Alloc80	96.4	86.8
15	IndCart	97.7	86.2
16	CART	92.1	86.2
17	MLP + BP	88.8	86.1
18	Bayesian Tree	98.0	85.3
19	C4.5	96.0	85.0
20	New ID	93.3	85.0
21	QuaDisc	89.4	84.5
22	SSV	90.9	84.3
23	Cascade	88.8	83.7
24	Log DA, Disc	88.1	83.7
25	LDA, Discrim	85.1	82.9
26	Kohonen	89.9	82.1
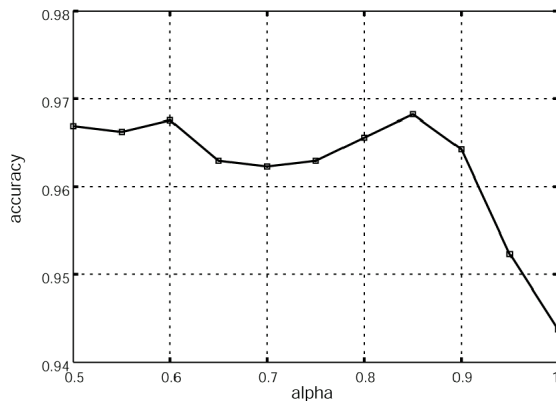
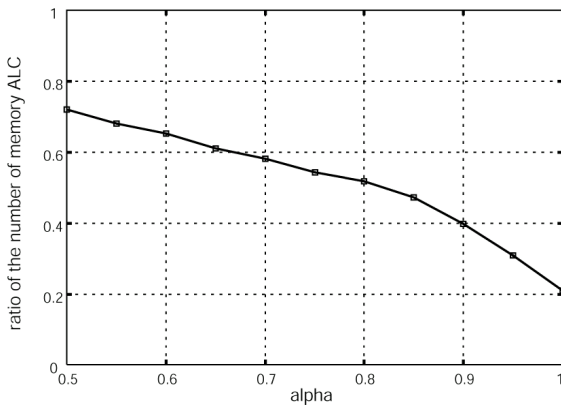Fig. 8: α vs. accuracy for Ionosphere data set

Fig. 9: α vs. rate of the number of memory ALCs' for the Ionosphere data set

was obtained by averaging the ten runs.

Table 3 shows the results that compare the accuracy of DAIS with one of the other systems for training and test sets. DAIS's accuracy is only 0.5% lower than the accuracy of the best system. In addition, DAIS's accuracy for training set is 100%. This is the unique characteristic of DAIS. Also, it is ensured that α is equal to or less than 1. This means that DAIS appropriately models the training mechanism in a thymus. DAIS's complete accuracy for a training set is different from over-training of neural network. Since DAIS shows high accuracy for a test set and considering the results of the training and test set, DAIS is the best classifier system for a satellite image data set.

Property: In this subsection, we discuss the characteristics of DAIS for the value of α. We focus on the accuracy and size of memory ALCs. It is important to decrease the memory size, because it makes

computational cost low when DAIS classifies unknown data. The cost is proportional to the memory size, since DAIS classifies data using all the memory ALCs.

The relationships between α and accuracy for the Ionosphere data set are depicted in Fig. 8. This shows that the accuracy falls when α is around 1. This is a common property for the other data sets, except in the case of the Diabetes data set. However, there is no other common property.

In contrast, the number of memory ALCs monotonically decreases with α (Fig. 9). The longitudinal axis represents the ratio of the number of memory ALC's to the number of training data. The reason for this decrease is that the large value of α makes a recognition domain of an ALC become large. Therefore, it is desirable for the value of α to be as large as possible. Therefore, the best value of α is a trade-off between the accuracy and the memory size.

These results mean that small α and large memory size do not always make the accuracy high and we must seek an optimized value of α for each problem.

CONCLUSION

This study introduced a new learning system based on the mechanism of the immune system. The results show that it is useful to apply the metaphor of self and non-self discrimination by T cell to classification problems. We prepare the method by deciding RDT with α. This is a novel algorithm in the field of AIS. The accuracy of DAIS is sufficiently high; if the value of α is equal to or less than 1, the DAIS's accuracy for training set is 100%. This shows that the prepared negative selection algorithm is acceptable for the T cell and training mechanisms.

DAIS is applicable to various classification problems. Unlike a neural network, DAIS does not need to optimize self-structure. It automatically determines the distribution of the memory ALCs. In addition, optimization of parameters is easy, because it only needs to optimize the value of α. These properties are important for classification. We think that DAIS is a general-purpose classification method. Thus, it can be said that the new negative algorithm and DAIS are useful for classification.

REFERENCES

1. De Castro, L. and J. Timmis, 2001. Artificial Immune Systems: A New Computational Intelligence Approach. Springer.

2. Hofmeyr, S. and S. Forrest, 1999. Immunity by Design; An artificial immune system. Proc. Genetic and Evolutionary Computation Conf., pp: 1289-1296.

3. Chao, D.L. and S. Forrest, 2002. Information filtering systems. Proc. Intl. Conf. on Artificial Immune Systems, pp: 132-140.

4. Timmis, J., M. Neal and J. Hunt, 2000. An artificial immune system for data analysis. BioSystems, 55: 143-150.

5. Potter, M.A. and K.A. De Jong, 1998. The coevolution of antibodies for concept learning. Proc. PPSN.

6. Watokins, A. and J. Timmis, 2004. Artificial immune recognition system (AIRS): An immune-inspired supervised learning algorithm. Genetic Programming and Evolvable Machines, 5: 291-317.

7. De Castro, L. and J. Timmis, 2002. Artificial immune systems: a novel paradigm for pattern recognition. L. Alonso, J. Corchado and C. Fyfe (Eds.), Artificial Neural Networks in Pattern Recognition, University of Paisley, pp: 67-84.

8. Graaff, A.J. and A.P. Engelbrecht, 2003. Using a threshold function to determine the status of lymphocytes in the artificial immune system. Proc. SAICSIT, pp: 286-274.

9. Graaff, A.J. and A.P. Engelbrecht, 2006. Optimised coverage of non-self with evolved lymphocytes in an artificial immune system. Intl. J. Computational Intelligence Res., 2: 127-150.

10. Blake, L. and C.J. Merz, 1998. UCI repository of machine learning databases. http://www.ics.uci. edu /~mlearn/ MLRepository. html.

11. Duch, 2000. Datasets used for classification: Comparison of results. http:// www. phys. uni. torun. pl/kmk/ projects/ datasets. html.