第20卷第9期 控 制 与 决 策 2005年9月
Vol. 20 No. 9 Control and Decision Sept. 2005

文章编号: 1001-0920(2005)09-1026-04

基于危险理论的免疫识别新算法

于 瀛, 侯朝桢
(北京理工大学 自动控制系, 北京 100081)

摘 要: 为提高系统对"危害"的识别能力, 基于免疫学中的"危险理论"提出了一种新的免疫识别算法. 该算法以识别"危险"为核心思想, 增加了对抗原提呈细胞和不同危险信号的使用. 仿真实验结果表明, 该方法与传统方法相比, 具有更高的识别率.

关键词: 危险理论; 负选择; 免疫识别
中图分类号: TP301.6 **文献标识码**: A

A Novel Immune Discrimination Algorithm Based on Danger Theory

YU Ying, HOU Chao-zhen
(1. Department of Automatic Control, Beijing Institute of Technology, Beijing 100081, China. Correspondent: YU Ying, E-mail: yuying12@sohu.com)

Abstract: To improve the system's capability of discrimination against danger, a novel immune discrimination algorithm based on the newly developed "Danger Theory" in immunology is proposed. This algorithm is based on the thinking of "danger" discrimination and introduced antigen presenting cell and different danger signals. Experimental results show that new algorithm provides by higher discrimination ratio.

Key words: Danger theory; Negative selection; Immune discrimination

1 引 言

随着人们对生物免疫系统认识的不断深入, 基于生物免疫机制开发新的计算智能-人工免疫系统(AIS)的计算模型和应用已成为目前的研究热点. 人工免疫系统框架下的免疫算法主要包括: 以保护和检测侵害为目的的免疫识别算法和由系统内部学习机制的优化而产生的免疫优化算法[1]. 免疫识别算法是构建人工免疫系统平台的基础.

目前, 免疫识别算法大多是基于传统免疫学理论中"自我-非我"(SNS)识别的模式而建立的, 其中以Forrest建立的负选择算法(NSA)为代表[2]. 这类算法对"自我"集定义中如何保证其不包含任何"非我"及如何满足"自我"集的动态变化问题未能很好地解决.

随着免疫学理论的不断深入, 近年来在免疫学界又新兴起一种新的理论——危险理论(DT). 对象的危险性是构成其识别模式的基础[3]. 目前, 它已引起了研究者的关注[4,5]. 本文以此为基础对如何改进免疫识别算法, 从而提高系统对危害的识别能力进行了探讨.

2 危险理论的基本思想

生物免疫系统中面对侵害能够产生免疫应答的机制主要依赖于B淋巴细胞所分泌的可用于识别和响应抗原(Ag)激励的抗体(Ab). 以往认为抗体与抗原之间的这种识别主要基于"自我-非我"的识别模式, Matzinger在此基础上提出了危险理论, 其主要思想是免疫系统更关注那些由受侵害组织发出危险信号作用的危险实体, 而不是那些与自身不同的异己[3]. 如图1所示, 说明了这两种理论的主要区别和关系, 其中N代表"非我"集, S代表"自我"集, 且

收稿日期: 2004-10-15; **修回日期**: 2004-12-08.
作者简介: 于瀛(1977—), 女, 河北唐山人, 博士生, 从事人工免疫系统、进化计算等研究; 侯朝桢(1938—), 男, 四川自贡人, 教授, 博士生导师, 从事分布式控制、进化计算等研究.

$S \cap N = \varnothing, S \cup N = I$(全集). D 代表所有危险个体 d_i 组成的集合, 即 $D = \{d_i \mid i = 1, 2, \cdots, n\}$.

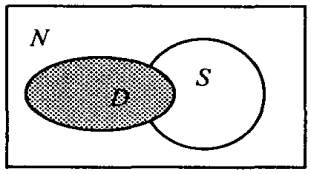

图 1　自我 - 非我理论与危险理论

由图 1 可见, 危险 d_i 存在于两种情况: 1) $D_S = D \cap S = \{d_i \mid d_i \in D \cap S\}$; 2) $D_N = D \cap N = \{d_i \mid d_i \in D \cap N\}$; 且 $D = D_S \cup D_N$. 这样, $\exists d_i \notin N$. 这说明并非所有的危险都来自于"非我", 也可能存在于"自我"集当中. 而在"自我 - 非我"模型中却存在 $D = N$ 的设定, 这样就忽略了两种情况: 1) D_S 存在的危险性; 2) 集合 $N - D_N$ 的无危害性. D_S 的存在使得由"自我"集定义而生成的抗体对这部分危害产生免疫耐受, 而对 $N - D_N$ 的响应则会增加系统的误报率.

根据文献[3], 危险理论模型以识别危险为核心, 把免疫系统中负责将抗原提呈给抗体的抗原提呈细胞(APC) 作为接收危险信号的主体. 受危害组织发出的危险信号 1 激活抗原提呈细胞, 当被激活的抗原提呈细胞在与抗体的结合中收到共同激励, 即危险信号 2 时, 抗体确认此抗原是危险的. 这样, 危险理论模型为建立AIS 中基于侵害检测的免疫识别算法(DADT) 提供了一种新思路.

3　DADT 算法设计

3.1　算法基本思想

从"自我 - 非我"识别的角度出发, DADT 算法应该完成系统对"某些自我"和"某些非我"的识别, 即危险理论中对危险的识别. 这种转变不是表达数据方式的改变, 而是表达和处理数据的变化, 图 2 表示了 DADT 算法的基本思想. 为表述方便, 以Ab 表示抗体, Ag 表示抗原, APC 表示抗原提呈细胞.

算法利用危险信号触发免疫应答, 危险信号辅助算法识别出需要的特征向量. 通过危险信号 1 和 2(Sig1, Sig2) 分别在Ag-APC, Ab-APC 之间建立两级不同的危险识别机制. 其中: Ag-APC 之间完成 APC 对Ag 的特征提取, 以 Sig1 激活APC 向Ab 对 Ag 的提呈; 而Ab-APC 之间要完成Ab 对由APC 提呈的信息进行识别, 由 Sig2 确认此Ag 是危险的, 由此产生免疫应答.

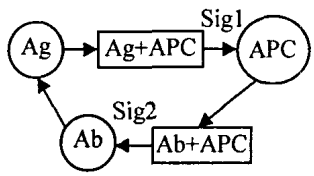

图 2　DADT 算法基本思想

3.2　算法描述

本文对 DADT 算法的描述建立在实赋范空间的基础上, 并提出了相关定理和定义, 算法以二进制位串为数据表示基础. 设 X^L 为由Ab 集 B、Ag 集 G 以及APC 集 P 组成的集合, 即 $X^L = B \cup G \cup P$, 其中 L 为其维数, 即数据长度. 集合 B, G, P 大小分别为 $m, n, r > 0$, 表示如下:

$$B = \{x_{bi} \mid i = 1, 2, \cdots, m\} = (x_{b1}, x_{b2}, \cdots, x_{bm})^{\mathrm{T}} = \begin{bmatrix} x_{b11} & x_{b12} & \cdots & x_{b1L} \\ x_{b21} & x_{b22} & \cdots & x_{b2L} \\ \vdots & \ddots & & \vdots \\ x_{bm1} & x_{bm2} & \cdots & x_{bmL} \end{bmatrix}, \tag{1}$$

$$G = \{x_{gi} \mid i = 1, 2, \cdots, n\} = (x_{g1}, x_{g2}, \cdots, x_{gn})^{\mathrm{T}} = \begin{bmatrix} x_{g11} & x_{g12} & \cdots & x_{g1L} \\ x_{g21} & x_{g22} & \cdots & x_{g2L} \\ \vdots & \ddots & & \vdots \\ x_{gn1} & x_{gn2} & \cdots & x_{gnL} \end{bmatrix}, \tag{2}$$

$$P = \{x_{pi} \mid i = 1, 2, \cdots, r\} = (x_{p1}, x_{p2}, \cdots, x_{pr})^{\mathrm{T}} = \begin{bmatrix} x_{p11} & x_{p12} & \cdots & x_{p1L} \\ x_{p21} & x_{p22} & \cdots & x_{p2L} \\ \vdots & \ddots & & \vdots \\ x_{pr1} & x_{pr2} & \cdots & x_{prL} \end{bmatrix}. \tag{3}$$

其中: $x_{bij}, x_{gij}, x_{pij} \in \{0, 1\}$.

定理 1　集合 X^L 为一实赋范线性空间.

证明　因为 $X = B \cup G \cup P$, 且集合 B, G, P 大小分别为 $m, n, r > 0$, 所以 X 为一非空集合, 可表示为

$$X = \{x \mid x \in \{x_{bi}, x_{gi}, x_{pi}\}\},$$

又因为 $x_{bij}, x_{gij}, x_{pij} \in \{0, 1\}$, 显然 x_{bi}, x_{gi}, x_{pi} 均为实向量, 所以 X 显然为一实线性空间.

对于 $\forall x, x' \in X$, 满足以下范数公理[6]:

1) $\| x \| \geq 0, \| x \| = 0 \Leftrightarrow x = \theta$;

2) $\| \alpha x \| = |\alpha| \| x \|, \alpha$ 是实数;

3) $\| x + x' \| \leq \| x \| + \| x' \|$.

所以 X 为一实赋范线性空间. □

在定理 1 的基础上将亲合度的计算扩展到实赋范空间进行讨论. 亲合度是人工免疫系统中广泛采

用的衡量细胞间识别程度的物理量. 目前普遍采用的海明距是亲合度计算中的实用方法之一, 以Ab识别Ag为例, 其定义如下:

$$D = \sum_{j=1}^{L} \delta_j, \tag{4}$$

其中 $\delta_j = 1$, 当 $x_{bij} \neq x_{gij}$ 时; 否则, $\delta_j = 0$.

在海明距的基础上, Ag-APC, Ab-APC 之间亲合度分别如下:

定理 2 Ag-APC 之间的亲合度 f^{GP} 为

$$f^{GP} = f^{Gp}(x_{gi}, x_{pj}) = \| x_{gi} - x_{pj} \|_1, \tag{5}$$

其中 $1 \leq i \leq n$, $1 \leq j \leq r$.

证明 在定理 1 的基础上, 由 $x = (x_1, x_2, \cdots, x_L) \in X^L$, 引入

$$\| x \|_1 = \sum_{i=1}^{L} |x_i|,$$

由海明距的定义式(4), 得

$$f^{GP} = f^{GP}(x_{gi}, x_{pj}) = \sum_{t=1}^{L} f_t^{GP} =$$

$$\sum_{t=1}^{L} |x_{git} - x_{pjt}| = \| x_{gi} - x_{pj} \|_1. \quad \square$$

定义 1 危险信号 1: 当Ag-APC之间的亲合度 $f^{GP} \leq T_1 (T_1 \geq 0)$ 时, APC 接收到的信号为危险信号 1, APC 同时被激活.

此时, 以该Ag为中心, T_1 为半径, 所构成的Ag的 T_1- 邻域内形成的该Ag的危险域.

在式(5)的定义下, Ag-APC 之间相似程度越高, 亲合度越小, 意味着APC对Ag信息提取程度越高, 形成的危险域越小. 这样, 收到危险信号 1 的 APC 数量会相应减少, 可以提高系统的使用率.

在APC被激活后, 需要将抗原信息提呈给Ab. 为了使APC能将Ag尽可能多的信息传递给Ab, APC将以概率α变异, 变异后的APC表示为APC*. α的选取应体现出APC对Ag特征的提取程度. 因此, 有

$$\alpha = \alpha(x_{gi}, x_{pj}) = 1 - e^{-\| x_{gi} - x_{pj} \|_1} = 1 - e^{-f^{GP}}. \tag{6}$$

定义 2 Ab-APC 之间的亲合度 f^{BP} 为

$$f^{BP} = f^{BP}(x_{bi}, x_{pj}) = 1 - \frac{\| x_{bi} - x_{pj} \|_1}{\max \| x_{bi} - x_{pj} \|_1}, \tag{7}$$

其中 $1 \leq i \leq m$, $1 \leq j \leq r$.

在式(7)的定义下, 不同的Ab与APC的亲和度越高, Ab对APC提呈信息的结合能力则越强. 这样, 采用式(5)计算的亲合度可直接体现出APC对Ag信息提取的能力, 而式(7)式的计算则可从另一角度表达出不同的Ab对APC提呈信息的识别程度.

定义 3 危险信号 2: 当Ab-APC之间的亲合度 $f^{BP} \geq T_2 (T_2 \geq 0)$ 时, Ab所接收到的信号为危险信号 2, 确认由APC提呈的信息是危险的, 产生免疫应答.

3.3 算法步骤

IDADT 算法框图如图 3 所示, 其主要步骤如下:

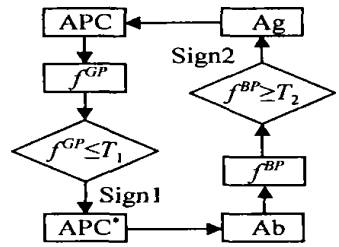

图 3 IDADT 算法

Step 1: 根据要识别的对象, 产生初始Ab, APC集合;

Step 2: 根据式(5), 计算Ag-APC之间的亲合度, 在危险域 T_1 内的APC将接收到危险信号 1, 同时APC被激活;

Step 3: 激活的APC按照式(6)进行变异;

Step 4: 根据式(7)计算Ab-APC* 之间的亲合度, 超过阈值 T_2 的Ab将接收到危险信号 2, 确认抗原的危险性, 产生免疫应答;

Step 5: 输出抗原;

Step 6: 重复 Step 2～ Step 5, 直到满足中止条件.

4 仿真实验

为了便于算法性能的比较, 本文采用文献[2]中负选择算法(NSA)所用的实验方法. 实验从算法检测失败率和执行时间两方面与负选择算法进行了比较. 检测失败率可体现出识别算法对抗原的检测效率; 执行时间因直接关系到算法执行的时效性, 故也是衡量识别算法检测效率的重要指标之一.

实验中每个算法均产生 10 组同样数量的抗体, 产生抗原的数目 $N_g = 500$. 抗体、抗原、抗原提呈细胞均采用随机生成的二进制位串表示, $L = 32$. 以 N_b 和 N_p 分别表示抗体和抗原提呈细胞生成的数目, 算法采用 $N_b = N_p$. 激活抗原提呈细胞和抗体的阈值分别表示为 T_1 和 T_2. P_f 表示检测失败的概率, 即 $P_f = N_g'/N_g$, 其中 N_g' 表示未检测到的抗原数目. T(ms) 代表算法完成每组实验的执行时间. 实验以抗原提呈细胞提取完所有抗原信息为结束条件. 实验结果如表 1 所示, 其中负选择算法中"自我"

表 1　NSA 算法与 IDADT 算法结果

算法	N_b	$T_2(T_1)$	N_g'	P_f	T/ms
NSA	8	16	38	0.076	78
	16	16	34	0.068	125
	24	16	42	0.084	167
	32	16	36	0.074	172
	40	16	38	0.076	225
	56	20	35	0.070	401
	64	20	34	0.068	703
	72	20	37	0.074	1 104
	96	20	41	0.082	1 964
	128	20	44	0.088	7 769
IDADT	8	0.5(16)	34	0.068	189
	16	0.5(16)	32	0.064	376
	24	0.5(16)	39	0.078	487
	32	0.5(16)	31	0.062	562
	40	0.5(16)	32	0.064	674
	56	0.625(12)	33	0.066	753
	64	0.625(12)	30	0.060	832
	72	0.625(12)	32	0.064	1 053
	96	0.625(12)	28	0.056	1 653
	128	0.625(12)	30	0.060	3 763

集大小为N_b, 激活抗体的阈值为T_2.

从表 1 的实验结果可以看出, 随着抗体和抗原提呈细胞数目的增多, T_1 随着危险域涉及范围而减小, 发出危险信号 1 更强, 表明抗原提呈细胞对抗原信息提呈能力增强. 随着 T_2 增大, 抗体与抗原的结合力更强, 表明抗体对抗原的识别能力在抗原提呈细胞发挥良好的信息提呈作用下不断增强. 在算法执行时间方面, IDADT 算法起初执行时间多于负选择算法. 这主要由于IDADT 算法中除了产生抗原提呈细胞需要占用额外的时间以外, 而且还增加了抗原提呈细胞对抗原信息的提取过程. 但随着抗体数量的增加, 负选择算法的执行时间以近乎指数的形式增长, 而IDADT 算法执行时间的增长量要小于负选择算法. 这主要由于算法中抗体的产生不需要由"自我"集的定义而产生, 因而不会伴随"自我"集的增加而指数增长. 从检测失败率和执行时间两方面来看, 改进的IDADT 算法建立的两级危险识别机制对抗原的整体检测率要高于负选择算法.

5　结　　论

本文在危险理论的基础上改进了现有的以"自我-非我"为基本模型的免疫识别算法. 新算法以识别的信息是否具有危险性为核心, 设立危险域, 通过不同的危险信号建立两级危险识别机制, 并由抗原提呈细胞发挥将危险信息特征提呈给抗体的作用. 仿真实验证明, 该算法是有效的, 且具有更强的危险识别能力.

参考文献(References)

[1] 蔡自兴, 龚涛. 免疫算法研究的进展[J]. *控制与决策*, 2004, 19(8): 841-846.
(Cai Z X, Gong T. Advance in Research on Immune Algorithms[J]. *Control and Decision*, 2004, 19(8): 841-846.)

[2] Forrest S, Perelson A S, Allen L, et al. Self-nonself Discrimination in a Computer[A]. *Proc of IEEE Symposium on Research in Security and Privacy*[C]. Oakland, 1994: 202-212.

[3] Matzinger P. The Danger Model: A Renewed Sense of Self[J]. *Science*, 2002, 296(12): 301-305.

[4] Aickelin U, Cayzer S. The Danger Theory and Its Application to AIS[A]. *Proc of the First Int Conf on AIS*[C]. England, 2002: 141-148.

[5] Emma Hart, Peter Ross. Improving SOSDM: Inspirations From the Danger Theory[A]. *Proc of the Second Int Conf on AIS*[C]. Berlin: Springer-Verlag, 2003: 194-203.

[6] 李大华. *应用泛函分析*[M]. 武汉: 华中科技大学出版社, 1999: 58-63.
(Li D H. *A Course in Applied Functional Analysis*[M]. Wuhan: Huzhong University of Science and Technology Publishing Company, 1999: 58-63.)