

第3章 免疫算法

生物免疫系统与脑神经系统、遗传系统并称为基于生物的三大信息处理系统。生物免疫系统的智能性和复杂性堪比大脑,有“第二大脑”之称。免疫算法是在免疫细胞的水平上模拟人体免疫系统信息处理过程中的识别、记忆、学习、正反馈、适应、负反馈、优化等功能。人工免疫算法没有统一的模型和算法结构。本章主要介绍生物免疫学的基本概念、免疫系统的组织结构、适应性免疫应答、克隆选择理论、克隆选择算法及其实现步骤。

3.1 免疫系统的基本概念

生物免疫系统的主要功能是识别“自己”与“非己”成分,并能破坏和排斥“非己”成分,而对“自己”成分能免疫耐受,不发生排斥反应,以维持机体的自身免疫稳定。下面先给出免疫学中的有关基本概念。

- (1) 免疫应答:指免疫系统识别并消灭侵入机体的病原体的过程。
- (2) 抗原(Ag):指能够诱导免疫系统发生免疫应答,并能与免疫应答的产物在体内或体外发生特异性反应的物质。
- (3) 表位:指抗原分子表面的决定抗原特异性的特殊化学基团,又称为抗原决定簇。
- (4) 淋巴细胞:指能够特异地识别和区分不同抗原决定簇的细胞,主要包括 T 细胞和 B 细胞两种。
- (5) 受体:指位于 B 细胞表面的可以识别特异性抗原表位的免疫球蛋白。
- (6) 抗体(Ab):指免疫系统受到抗原刺激后,识别该抗原的 B 细胞转化为浆细胞并合成和分泌可以与抗原发生特异性结合的免疫球蛋白。
- (7) 匹配:指抗原表位与抗体或 B 细胞受体形状的互补程度。
- (8) 亲和力:指抗原表位与抗体或 B 细胞受体之间的结合力,抗原表位与抗体或 B 细胞受体匹配得越好,二者之间的亲和力越高。
- (9) 免疫耐受:指免疫活性细胞接触抗原性物质时所表现的一种特异性无应答状态。
- (10) 免疫应答成熟:指记忆淋巴细胞比初次应答的淋巴细胞具有更高亲和力的现象。

3.2 免疫系统的组织结构

免疫器官是免疫细胞发生发育和产生效应的部位,免疫细胞主要在骨髓和胸腺中产生,从其产生到成熟并进入免疫循环,需要经历一系列复杂变化。免疫细胞主要包括淋巴细胞、吞噬细胞,淋巴细胞又分为 B 淋巴细胞和 T 淋巴细胞。

B 淋巴细胞是由骨髓产生的有抗体生成能力的细胞,其受体是膜结合抗体,抗原与这些膜

抗体分子相互作用可引起 B 细胞活化、增殖,最终分化成浆细胞以分泌抗体。

T 淋巴细胞产生于骨髓,再迁移到胸腺并分化成熟。T 细胞可分为辅助性 T 细胞(T_H)和细胞毒性 T 细胞(CTL),它们只识别暴露于细胞表面并与主要组织相容性复合体(MHC)相结合的抗原肽链,并进行应答。在对抗原刺激的应答中, T_H 细胞分泌细胞因子,促进 B 细胞的增殖与分化,而 CTL 则直接攻击和杀死内部带有抗原的细胞。

吞噬细胞起源于骨髓,成熟和活化后,产生形态各异的细胞类型,能够吞噬外来颗粒(如微生物、大分子,甚至损伤或死亡的自身组织)。

3.3 免疫系统的免疫机制

免疫系统的功能是免疫细胞对内外环境的抗原信号做出免疫应答反应。免疫应答是指免疫活性细胞对抗原分子的识别、活化、增殖、分化,以及最终发生免疫效应的一系列复杂的生物学反应过程,包括先天性免疫应答和适应性免疫应答两种。先天性免疫应答是生物在种系发展和进化过程中逐渐形成的天然防御机制,包括吞噬细胞对侵入机体的细菌和微生物的吞噬作用,以及皮肤、机体内表皮等生理屏障。适应性免疫应答不是天生就有的,而是个体在发育过程中接触抗原后发展而形成的,只对该特异抗原起作用而对其他抗原不起作用的免疫力,包括体液免疫和细胞免疫。

20 世纪 50 年代,著名的免疫学家 Burnet 提出了关于抗体形成的克隆选择学说,该学说得到了大量的实验证明,合理地解释了适应性免疫应答机理。克隆选择理论认为抗原的识别能够刺激淋巴细胞增殖并分化为效应细胞。受抗原刺激的淋巴细胞的增殖过程称为克隆扩增。B 细胞和 T 细胞都能进行克隆扩增,不同的是 B 细胞在克隆扩增中要发生超突变,即 B 细胞受体发生高频变异,并且其效应细胞产生抗体,而 T 淋巴细胞不发生超突变,其效应细胞是淋巴因子、 T_K 或 T_H 细胞。B 淋巴细胞的超突变能够产生 B 细胞的多样性,同时也可以产生与抗原亲和力更高的 B 细胞。B 细胞在克隆选择过程中的选择和变异导致了 B 细胞的免疫应答具有进化和自适应的性质。

当抗原侵入机体时,B 细胞的适应性免疫应答能够产生抗体,如图 3.1 所示。如果抗原表位与某一 B 细胞受体的形状互补,则二者之间会产生亲和力而相互结合,在 T_H 细胞发出的第二信号作用下,该 B 细胞被活化。活化 B 细胞进行增殖(分裂),增殖 B 细胞要发生超突变,一方面产生了 B 细胞的多样性,另一方面也可以产生与抗原亲和力更高的 B 细胞。

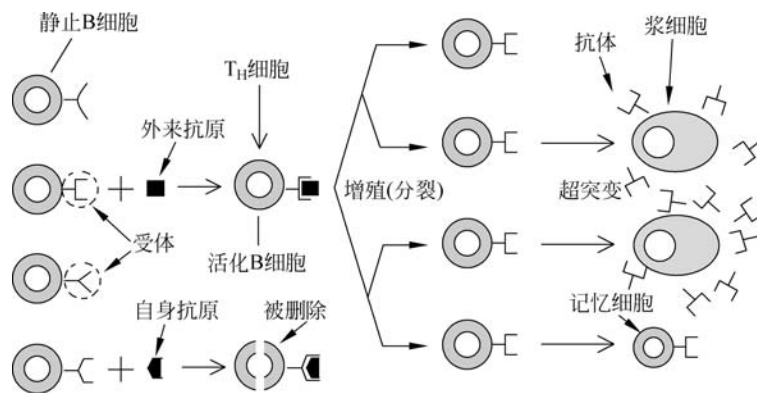


图 3.1 B 细胞的克隆选择过程

免疫系统通过若干世代的选择和变异来提高 B 细胞与抗原的亲和力。产生的高亲和力 B 细胞进一步分化为抗体分泌细胞,即浆细胞,浆细胞产生大量的活性抗体用以消灭抗原。同时,高亲和力 B 细胞也分化为长期存在的记忆细胞。记忆细胞在血液和组织中循环但不产生抗体,当与该抗原类似的抗原再次侵入机体时,记忆细胞能够快速分化为浆细胞以产生高亲和力的抗体。记忆细胞的亲和力要明显高于初始识别抗原的 B 细胞的亲和力,即发生免疫应答成熟。

3.4 免疫系统的学习及优化机理

从信息处理的观点看,生物免疫系统是一个并行的分布自适应系统,具有多种信息处理机制,它能够识别自己和非己,通过学习、记忆解决识别、优化和分类问题。

1. 免疫应答中识别、学习、记忆的机理

免疫系统中的每个 B 细胞的特性由其表面的受体形状唯一决定。体内 B 细胞的多样性极其巨大,可达 10^8 数量级。若 B 细胞的受体与抗原结合点的形状可用 L 个参数来描述,则每个 B 细胞可表示为 L 维空间中的一点,整个 B 细胞库都分布在这 L 维空间中,称此空间为形状空间。抗原在形状空间中用其表位的互补形状来描述。

形状空间的示意图如图 3.2 所示,B 细胞与抗原的亲和力大小可用它们在形状空间的距离定量表示。B 细胞与抗原距离越近,B 细胞受体与抗原表位形状的互补程度越大,于是二者之间的亲和力就越高。对于某一侵入机体的抗原 Ag_1 与 Ag_2 ,当体内的 B 细胞与它们的亲和力达到某一门限时才能被激活。被激活的 B 细胞大约为 B 细胞总数的 $(1\sim 10)$ 万分之一。这些被激活的 B 细胞分布在以抗原为中心,以 ϵ 为半径的球形域内,这一球形域称为该抗原的刺激球。

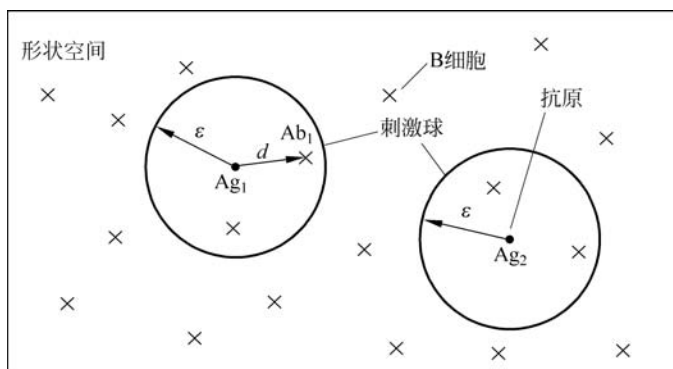


图 3.2 形状空间的示意图

根据 B 细胞和抗原的表达方式的不同,形状空间可分为欧几里得形状空间和海明形状空间。假设抗原 Ag_1 与 B 细胞 Ab_1 分别用向量 $(ag_1, ag_2, \dots, ag_L)$ 和 $(ab_1, ab_2, \dots, ab_L)$ 描述,若每个分量为实数,则它们所在的形状空间为欧几里得(以下简称“欧氏”)形状空间,抗原与 B 细胞之间的亲和力可表示为

$$\text{Affinity}(Ag_1, Ab_1) = \sqrt{\sum_{i=1}^L (ab_i - ag_i)^2} \quad (3.1)$$

若 B 细胞和抗原的每个分量为二进制数,则它们所在的形状空间为海明形状空间,B 细胞

和抗原之间的亲和力可表示为

$$\text{Affinity}(Ag_1, Ab_1) = \sum_{i=1}^L \delta, \text{其中 } \delta = \begin{cases} 1 & ab_i \neq ag_i \\ 0 & \text{其他} \end{cases} \quad (3.2)$$

生物适应性免疫应答中蕴含着学习与记忆原理,可通过 B 细胞和抗原在形状空间中的相互作用来说明。如图 3.3(a)所示,对于侵入机体的抗原 Ag_1 ,其刺激球内的 B 细胞 Ab_1 、 Ab_2 被活化。如图 3.3(b)所示,被活化 B 细胞进行克隆扩增,产生的子 B 细胞发生变化以寻求亲和力更高的 B 细胞,经过若干世代的选择和变化,产生了高亲和力 B 细胞,这些 B 细胞分化为浆细胞以产生抗体消灭抗原。因此,B 细胞是通过学习过程来提高其亲和力的,这一过程是通过克隆选择原理实现的。如图 3.3(c)所示,抗原 Ag_1 被消灭后,一些高亲和力 B 细胞分化为记忆细胞,长期保存在体内。当抗原 Ag_1 再次侵入机体时,记忆细胞能够迅速分化为浆细胞,产生高亲和力的抗体来消灭抗原,这称为二次免疫应答。如图 3.3(d)所示,若侵入机体的抗原 Ag_2 与 Ag_1 相似,并且 Ag_2 的刺激球包含由 Ag_1 诱导的记忆细胞,则这些记忆细胞被激活以产生抗体,这一过程称为交叉反应应答。由此可见,免疫记忆是一种联想记忆。

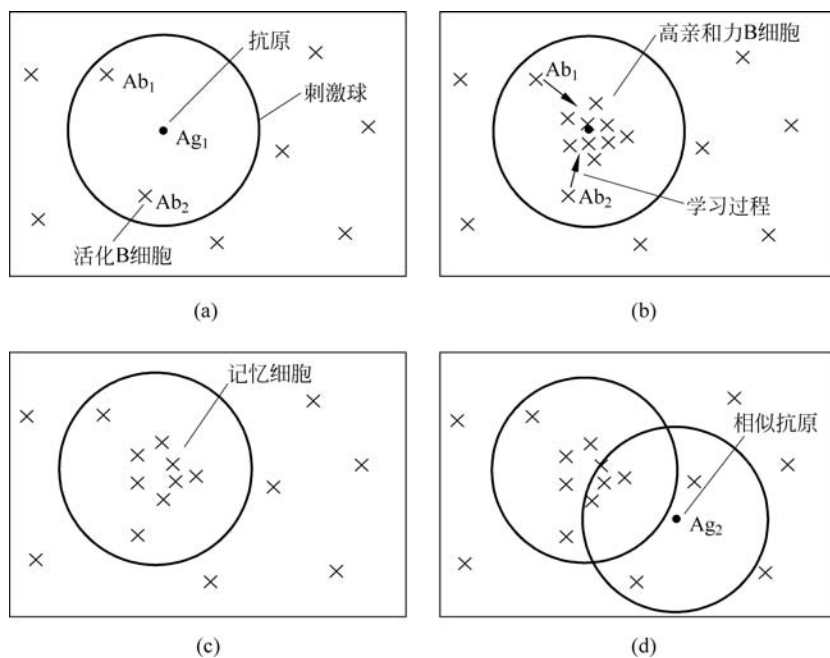


图 3.3 生物适应性免疫应答中蕴含着学习与记忆原理

2. 免疫应答中的优化机理

免疫系统通过 B 细胞的学习过程产生高亲和力抗体。从优化的角度来看,寻求高亲和力抗体的过程相当于搜索给定抗原的最优解,这主要是通过克隆原理和变异机制实现的。B 细胞的变异机制除了超突变外,还有受体修饰,即超突变产生的一些亲和力低的或与自身反应的 B 细胞受体被删除并产生新受体。B 细胞群体通过选择、超突变和受体修饰来搜索高亲和力的 B 细胞,进而产生抗体消灭抗原,这一过程如图 3.4 所示。

为便于说明,假设 B 细胞受体的形状只需一个参数的一维形状空间来描述。图 3.4 中横坐标表示一维形状空间,所有 B 细胞均分布在横坐标上,纵坐标表示形状空间中 B 细胞的亲和力。在初始适应性免疫应答中,如果 B 细胞 A 与抗原的亲和力达到某一门限值而被活化,

则该 B 细胞进行克隆扩增。在克隆扩增的同时 B 细胞发生超突变,使得子 B 细胞受体在母细胞的基础上发生变异,这相当于在形状空间中母细胞的附近寻求亲和力更高的 B 细胞。如果找到亲和力更高的 B 细胞,则该 B 细胞又被活化而进行克隆扩增。经过若干世代后,B 细胞向上“爬山”找到形状空间中局部亲和力最高的点 A' 。

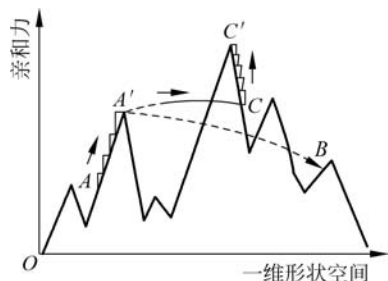


图 3.4 免疫应答中的优化搜索过程

如果 B 细胞只有超突变这一变化机制,那么适应性免疫应答只能获得局部亲和力最高的抗体 A' ,而不能得到具有全局最高亲和力的抗体 C' 。B 细胞的受体修饰可以有效避免以上情况的发生。如图 3.4 所示,受体修饰可以使 B 细胞在形状空间中发生较大的跳跃,在多数情况下产生了亲和力低的 B 细胞(如 B 点),但有时也产生了亲和力更高的 B 细胞(如 C 点)。产生的低亲和力 B 细胞或与自身反应的 B 细胞被删除,而产生的高亲和力 B 细胞 C 则被活化而发生克隆扩增。

经过若干世代后,B 细胞从 C 点开始,通过超突变找到形状空间中亲和力最高的 B 细胞 C' 。B 细胞 C' 进一步分化为浆细胞,产生大量高亲和力的抗体以消灭抗原。

因此,适应性免疫应答中寻求高亲和力抗体是一个优化搜索的过程,其中超突变用于在形状空间的局部进行贪婪搜索,而受体修饰用来脱离或避免搜索过程中陷入形状空间中的局部最高亲和力的点。

3.5 免疫算法及克隆选择算法的实现步骤

免疫算法大多将 T 细胞、B 细胞、抗体等功能合而为一,统一抽象出检测器概念,主要模拟生物免疫系统中有抗原处理的核心思想,包括抗体的产生、自体耐受、克隆扩增、免疫记忆等。在用免疫算法解决具体问题时,首先需要将问题的有关描述与免疫系统的有关概念及免疫原理对应起来,定义免疫元素的数学表达,然后再设计相应的免疫算法。

1. 免疫算法的基本步骤

免疫算法没有统一的形式,已提出的多种形式有反向选择算法、免疫遗传算法、克隆选择算法、基于免疫网络的免疫算法、基于疫苗的免疫算法等。

免疫算法一般由以下基本步骤组成。

(1) 定义抗原。将需要解决的问题抽象成符合免疫系统处理的抗原形式,抗原识别则对应问题的求解。

(2) 产生初始抗体群体。将抗体的群体定义为问题的解,抗体与抗原之间的亲和力对应问题解的评估:亲和力越高,表明解越好。类似于遗传算法,首先产生初始抗体群体,对应问题的一个随机解。

(3) 计算亲和力。计算抗原与抗体之间的亲和力。

(4) 克隆选择。与抗原具有较高亲和力的抗体优先得到繁殖,抑制浓度过高的抗体(避免局部最优解),淘汰低亲和力的抗体。为获得多样性(追求最优解),抗体在克隆时经历变异(如高频变异等)。在克隆选择中,抗体促进和克隆删除对应优化解的促进与非优化解的删除等。

(5) 评估新的抗体群体。若不能满足终止条件,则转向步骤(3),重新开始;若满足终止

条件,则当前的抗体群体为问题的最优解。

2. 克隆选择算法的实现步骤

基于生物免疫系统克隆选择原理的克隆选择算法,模拟免疫系统的克隆选择过程进行优化与学习。克隆选择算法的步骤如下。

(1) 随机产生一个包含 N 个抗体的初始群体。

(2) 计算群体中每个抗体(相当于一个可行解)的亲合力(即可行解的目标函数值),根据抗体的亲合力,选出 n 个亲合力最高的抗体。

(3) 被选出的每个抗体均进行克隆,每个抗体克隆出若干新抗体,抗体的亲合力越高,其克隆产生的抗体越多。可以通过以下方法实现:将这些抗体按其亲合力的高低降序排列(假设有 n 个抗体),则这 n 个抗体克隆产生抗体的数目为

$$N_c = \sum_{i=1}^n \text{round}\left(\frac{\beta \cdot N}{i}\right) \quad (3.3)$$

其中, N_c 为总共产生的克隆抗体的数目; β 为一个因子,用以控制抗体克隆数目的大小; N 为抗体的总数; $\text{round}(\cdot)$ 表示取整操作。

(4) 这些新个体进行免疫应答成熟操作(即新个体发生变异以提升其亲合力),这些变异后的抗体组成下一代群体。

(5) 从群体中选出一些亲合力最高的个体加入记忆集合,并用记忆集合中的一些个体替换群体中的一些个体。

(6) 用随机产生个体替换群体中一部分个体。

(7) 返回步骤(2)循环计算,直到满足结束条件。

克隆选择算法的流程如图 3.5 所示。

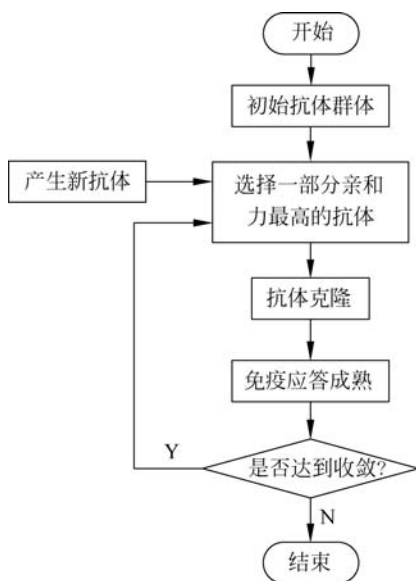


图 3.5 克隆选择算法的流程图

上述克隆选择算法已用于函数优化、组合优化(解决 TSP 问题)、模式识别等问题。