

第四部分 进化控制

一、生物进化与模拟进化计算

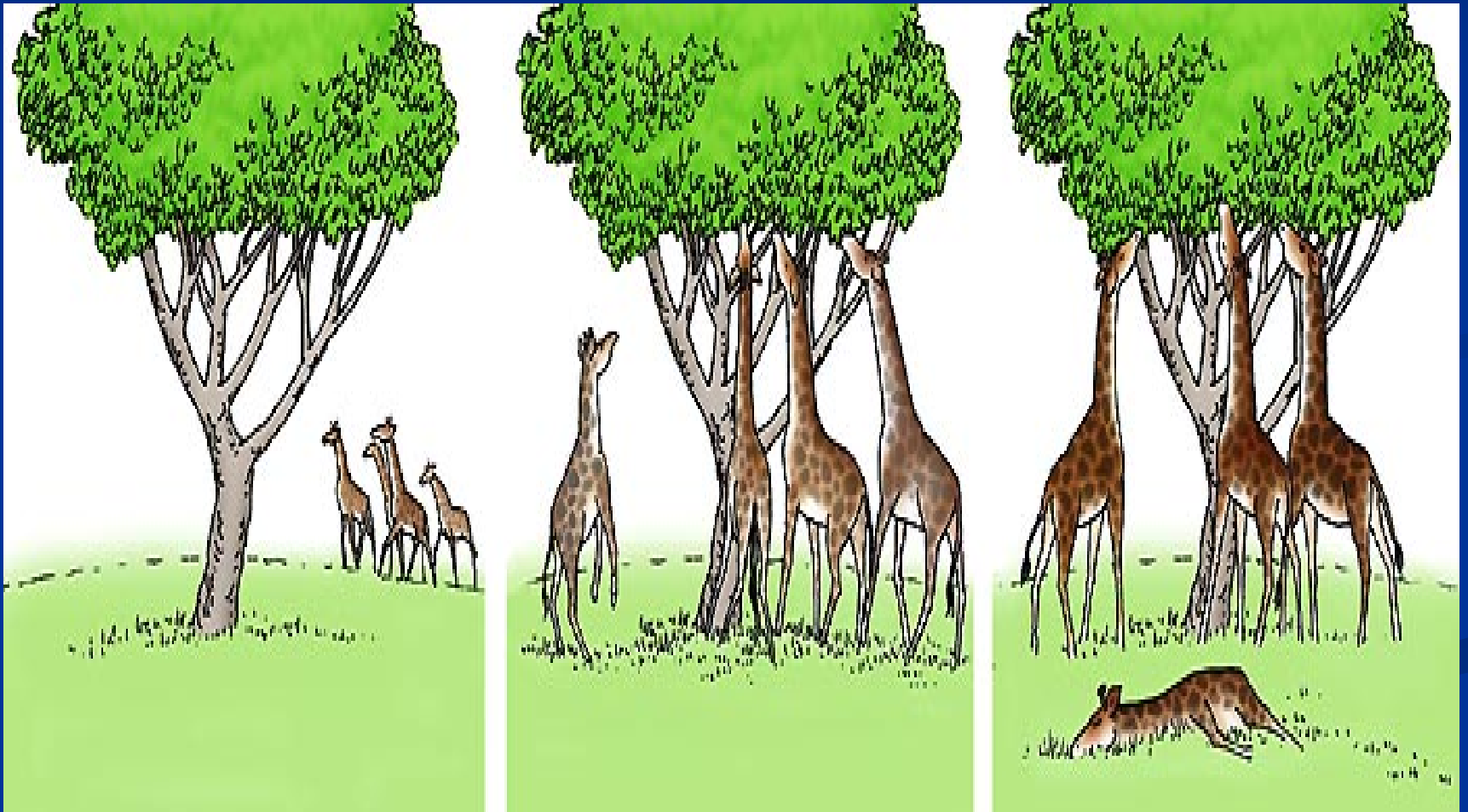
1. 达尔文进化论

自然界生态系统中的任何生物都是与其生活的环境密切相联系，并且与之相适应的。

达尔文(Darwin)发现的自然进化规律的核心思想是“物竞天择、适者生存”。生物在其生存斗争的过程中，优胜劣汰，使其物种品质不断得到改良，以适应其生存环境，体现为适应性的进化(evolution)。

自然选择—在生存斗争中，适者生存、不适者被淘汰。

自然选择的结果—使生存下来的生物产生适应性。



实例

长颈鹿祖先发生变异

个体有长颈长前肢的也有短颈短前肢的

缺乏青草的环境下

长颈长前肢生存

逐代选择

现代长颈鹿

生物进化过程

变异是不定向

个体差异

生存斗争

适者生存，
不适者淘汰

逐代遗传积累

适应环境新类型

自然选择

①选择的原材料——生物个体所包含的变异

②选择的因素——自然环境

③选择的手段——
(进化的动力)
生存斗争

④选择的结果——适应环境

现代进化的理论

种群是生物进化的单位

突变和基因重组产生进化的原材料

自然选择决定生物进化的方向

隔离导致物种形成

现代进化论所揭示的进化机制在本质上是一个自然的、并行发生的、稳健的优化过程，这一优化过程的目标是使生命体达到适应环境的最佳结构和性能，生物种群则通过优胜劣汰及遗传变异来达到优化的目的。

这一自然界的有效的优化机制是与人类自己发明的任何解析方法截然不同的。生物在进化过程中所解决的许多问题具有混沌、偶然、暂态和非线性相互作用等特点，这样的问题正是传统优化方法难以解决的。

2. 模拟进化计算

自然界为我们提供了可供借鉴的十分有效的适应性机制，它为许多传统优化方法难以解决的问题提供了新的解决手段。模拟进化计算技术就是模拟自然界生物进化机制求解优化和搜索问题的一类自组织、自适应的人工智能技术。

◆ 进化计算模型

从抽象层次看，进化计算技术所处理的问题可描述为：

$$\text{Max } f(X)=f(x_1, x_2, x_3, \cdots, x_n) \quad (1)$$

$$X=(x_1, x_2, x_3, \cdots, x_n) \in \Omega$$

其中， Ω 为某个 n 维空间上的一个子集， $f: \Omega \rightarrow \mathbf{R}^1$ 是一个也许无法解析表达的函数。（1）式也可解释为最优化问题，包括各类数学规划、组合优化、随机优化等优化模型。

作为求解（1）式问题的模拟进化计算技术，其基本原理可概括为：

- ① 将目标函数 f 对应为一物种群体在给定环境下的适应性度量；
- ② 将优化变量 \mathbf{X} 对应为种群中的个体；
- ③ 将寻优过程对应为相应种群在给定环境下的进化过程。

定义1（染色体编码） 一个具有下列性质的有限长字符串 $a_1 a_2 a_3 \dots a_L$ 称为优化变量 X 的一个染色体编码。

$$A=a_1 a_2 a_3 \dots a_L \Rightarrow X=(x_1, x_2, x_3, \dots, x_n)$$

其中， L 称为 X 的编码长度， $A=a_1 a_2 a_3 \dots a_L$ 称为 X 的编码。

把 a_i 看作是一个遗传基因，其所有可能取值称为等位基因。一个 A 即为由 L 个基因组成的一个染色体。对于给定问题， L 选取后通常固定不变。

根据不同的需要，等位基因可以是一组整数，或某一范围内的实数，或一组符号。最简单常用的等位基因由数字0和1组成，相应的染色体则为一个 L bit长的二进制串。染色体编码 A 称为个体的基因型，而与它对应的 X 则称为个体的表现型。

定义2（适应值函数） 设 R_0^+ 为正实数集。一个映射 $F: \Omega \rightarrow R_0^+$ 称为（1）式的一个适应值函数，如果 F 与 f 有相同的全局极大值点，且满足：

$$f(X_1) \geq f(X_2) \Rightarrow F(X_1) \geq F(X_2), \quad X_1, X_2 \in \Omega$$

个体 A 的适应度就用其表现型的适应值函数值表示，其表现型 X 越接近目标函数 f 的全局极大值点，其适应度就越大；反之，其适应度就越小。

模拟进化计算求解（1）式问题就是通过搜索个体空间来完成的。

与生物一代一代地进化相似，模拟进化计算的过程也是一个反复迭代产生一代代新种群的过程。

种群不断地经历遗传和进化，每次都按照优胜劣汰的规则将适应性强的个体更多地遗传到下一代，这样，在种群中形成越来越优秀的个体。当求解要求满足后，最后一代中的最优个体的表现型就作为（1）式问题的解。

生物的进化机制主要是通过选择、染色体之间的交叉和基因变异来完成。



模拟进化算法使用选择算子(selection operator)、交叉算子(crossover operator)和变异算子(mutation operator)来模拟自然界的进化机制。

- ① **选择算子**: 模拟自然选择操作, 体现“优胜劣汰”规律。它按一定的方法, 依据个体的适应度从当前种群中选择出一些优良个体(作为母本或让其遗传到下一代)。
- ② **交叉算子**: 模拟有性繁殖的基因重组操作。它将选择算子所选择的每一对母体, 依一定概率(称为交叉概率, crossover rate)交换它们之间的部分基因。
- ③ **变异算子**: 模拟基因突变的操作。它对种群中的每一个个体以一定概率(称为变异概率, mutation rate)改变某一个或某一些基因座上的基因值为其它等位基因。

交叉算子和变异算子的作用在于生成新的个体, 它们统称为繁殖算子。

- 模拟进化算法的一般框架:

- **Step 1 (初始化)** 确定种群规模 N 及终止准则(如最大进化代数或近似解精度); 随机生成 N 个个体作为初始种群 $P(0)$; 置进化代数计数器 $t:=0$ 。

- **Step 2 (个体评估)** 评估计算种群 $P(t)$ 中每一个个体的适应度。

- **Step 3 (种群进化)**

- 选择: 将选择算子作用于种群 $P(t)$;

- 繁殖: 分别将交叉算子和变异算子作用于选择后的种群, 生成下一代种群 $P(t+1)$ 。

- **Step 4 (终止判别)** 如果 $P(t+1)$ 满足终止准则, 则输出 $P(t+1)$ 中具有最大适应度的个体作为最优解, 终止计算; 否则置 $t:=t+1$, 转Step 2。

遗传算法(Genetic Algorithm, 简称GA)就是一类现在得到广泛应用的典型的模拟进化算法, 它由美国密执安大学的**Holland**教授提出, 后经由**De Jong**、**Goldberg**等归纳总结而成。Holland将进化算法用于自然和人工系统的自适应行为的研究, 并出版了其开创性的著作《**Adaptation in Natural and Artificial Systems**》。