

西安电子科技大学
XIDIAN UNIVERSITY

人工智能概论

刘若辰
西安电子科技大学人工智能学院

西安电子科技大学
XIDIAN UNIVERSITY

第七章 遗传算法

- 7.1 遗传算法 (GA)
- 7.2 遗传算法实例

人工智能学院

西安电子科技大学
XIDIAN UNIVERSITY

7.1. 遗传算法

1. 遗传算法的生物学基础
2. 遗传算法的组成

人工智能学院

西安电子科技大学
XIDIAN UNIVERSITY

引言

- 生物种作为复杂系统，具有奇妙的自适应、自组织和自优化能力，这是一种生物在进化过程中体现的智能。
- 遗传算法：着眼于对生物群体进化过程的模拟。

1975年约翰·霍兰德提出

- 应用领域：
 - 函数优化（连续，组合）
 - 机器人路径规划
 - 图像处理与模式识别
 - 机器学习
 - 智能控制
 - 等

约翰·霍兰德 (1929-)

人工智能学院

遗传算法的生物学基础

达尔文的进化论-适者生存原理

- 每一种物种在发展中越来越适应环境。物种每个个体的基本特征由后代所继承，但后代又会产生一些异于父代的新变化。在环境变化时，只有那些能适应环境的个体特征才能保留下来。

孟德尔的遗传学说-基因遗传原理

- 遗传以密码方式存在细胞中，并以基因形式包含在染色体内。每个基因有特殊的位置并控制某种特殊性质；所以，每个基因产生的个体对环境具有某种适应性。基因突变和基因杂交可产生更适应于环境的后代。经过优胜劣汰的自然淘汰，适应性高的基因结构得以保存下来。

“适者生存+劣者淘汰”

达尔文 (1809-1882)
孟德尔 (1822-1884)

人工智能学院

遗传算法的生物学基础

遗传学基本概念与术语

- 基因重组/交叉：在有性繁殖过程中，控制不同形状的染色体基因重新组合；
- 基因突变：染色体的某些基因位的组成或数目发生改变。

生物进化过程

基因重组

基因突变

人工智能学院

遗传算法的生物学基础

遗传算法 (Genetic Algorithm, GA) 是一类借鉴生物界的进化规律 (适者生存，优胜劣汰遗传机制) 演化而来的随机化搜索方法。

生物进化过程

通过繁殖、变异、竞争、选择来实现的。

人工智能学院

遗传算法的基本思想

- 遗传算法GA把问题的解表示成“染色体”，在算法中即是以一定方式编码的串。
- 在执行遗传算法之前，给出一群“染色体”，也即假设解。
- 把这些假设解置于问题的“环境”中，并按适者生存的原则，从中选择出较适应环境的“染色体”进行复制，再通过交叉，变异过程产生更适应环境的新一代“染色体”群。

这样，一代一代地进化，最后就会收敛到最适应环境的一个“染色体”上，它就是问题的最优解。

进化计算: 遗传算法、进化策略、进化规划、遗传规划

人工智能学院

遗传算法的组成

一般的遗传算法由**四个部分组成**:

- 编码机制
- 种群初始化
- 适应度函数
- 遗传算子（选择、交叉、变异）

遗传算法模型

人工智能学院

遗传算法的组成

编码机制

- 用遗传算法解决问题时，首先要对解决问题的模型结构和参数进行编码
- 编码：对于解空间的解用特定符号按一定顺序排成的串，其中每一个码以及相应的位置对应的就是基因以及基因位
- GA中的编码方法：二进制编码、浮点数编码方法、格雷码编码、符号编码等。

解空间

其中一个码和码所在的位置即基因/基因位

编码

一个解的编码即：染色体

人工智能学院

遗传算法的组成

二进制编码：它所构成的个体的基因型是一个由0或1组成的编码串

假设解的取值范围为: $[X_{min}, X_{max}]$

某一个体的编码是: $b_L b_{L-1} b_{L-2} \dots b_2 b_1$, 则对应的解码公式为:

$$X = X_{min} + \frac{X_{max} - X_{min}}{2^L - 1} (\sum_{i=1}^L b_i 2^{i-1})$$

串长L取决于求解精度: $\delta = \frac{|X_{max} - X_{min}|}{2^L - 1}$

德

人工智能学院

遗传算法的组成

决策变量（整数或实值）用二进制字符串表示。

例如，对于**三维空间**的优化问题，设一个染色体的编码为：

| | | |
|---------------|---------------|---------------|
| x_1 | x_2 | x_3 |
| 1 0 1 1 | 0 1 1 0 | 1 1 1 0 |

解码

- 染色体中有**12个基因**
- 对于**整数**决策变量:
 $x_1 = 1 \times 2^0 + 1 \times 2^1 + 0 \times 2^2 + 1 \times 2^3 = 11$
 $x_2 = 0 \times 2^0 + 1 \times 2^1 + 1 \times 2^2 + 0 \times 2^3 = 6$
 $x_3 = 0 \times 2^0 + 1 \times 2^1 + 1 \times 2^2 + 1 \times 2^3 = 12$
 $(11, 6, 12)$

解码

- 对于**实值**决策变量，如果它们都在 $[-1, 1]$ 之间：
 $x_1 = -1 + (1 - (-1)) \times 11 / (2^4 - 1) = 7/15$
 $x_2 = -1 + (1 - (-1)) \times 6 / (2^4 - 1) = -3/15$
 $x_3 = -1 + (1 - (-1)) \times 12 / (2^4 - 1) = 9/15$
 $(7/15, -3/15, 9/15)$

人工智能学院

 西安电子科技大学
XIDIAN UNIVERSITY

练习

1. 利用遗传算法求解最大化问题 $f(x) = 5x^2 + 2$, $x \in [-14, 48]$, 个体采用二进制编码, 编码长度为 5, 则该二进制编码的精度是_____.
 (A) 1 (B) 2 (C) 3.

2. 遗传算法中, 设某参数的取值范围为 $[-1, 30]$, 对其采用编码长度为 5 的二进制编码, 则对二进制符号串“00010”进行解码, 得到的数值为_____.
 (A) 2 (B) 1 (C) 0

13 人工智能学院

 西安电子科技大学
XIDIAN UNIVERSITY

遗传算法的组成

➤ 种群初始化: 随机初始化

| 解的编号 | 解的编码 | 个体的编号 | 个体的编码 |
|------|-------|-------|------------|
| 1 | ○○○○○ | 1 | 0001100000 |
| 2 | ○○○○○ | 2 | 0101111001 |
| 3 | ○○○○○ | 3 | 0000000101 |
| 4 | ○○○○○ | 4 | 1001110100 |
| ... | ... | ... | ... |
| N | ○○○○○ | N | 0001010011 |

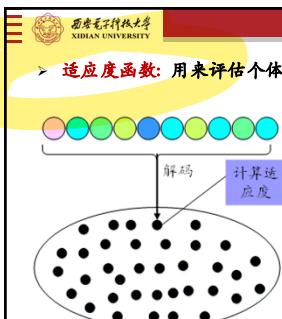
群体规模

人工智能学院

 西安电子科技大学
XIDIAN UNIVERSITY

遗传算法的组成

➤ 适应度函数: 用来评估个体优、劣的标准, 适应度值越大的个体越优



| 编号 | 个体的编码 | 解码 | 适应度值 |
|-----|------------|-------|----------|
| 1 | 0001100000 | X_1 | $f(X_1)$ |
| 2 | 0101111001 | X_2 | $f(X_2)$ |
| 3 | 0000000101 | X_3 | $f(X_3)$ |
| 4 | 1001110100 | X_4 | $f(X_4)$ |
| ... | ... | ... | ... |
| N | 0001010011 | X_N | $f(X_N)$ |

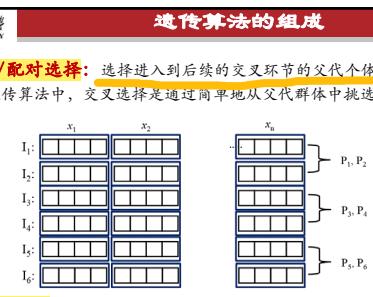
人工智能学院

 西安电子科技大学
XIDIAN UNIVERSITY

遗传算法的组成

➤ 交叉选择/配对选择: 选择进入到后续的交叉环节的父代个体的过程。

- 在典型的遗传算法中, 交叉选择是通过简单地从父代群体中挑选个体对来完成的。



- 配对选择方法: 轮盘赌选择、锦标赛选择、随机遍历抽样、局部选择、截断选择

人工智能学院

遗传算法的组成

➤ **交叉操作:** 选中的两个父代个体交换某些基因位形成子代个体的过程

- **交叉概率** P_c : 在种群中, 个体被选择出进行交叉的概率; (随机数 $< P_c$, 可交叉)
- 交叉方式: 单点交叉、两点交叉、多点交叉、均匀交叉等;
- 单点交叉: 随机产生一个有效的交叉位置, 染色体交换位于该交叉位置后的所有基因。

交叉概率 P_c

父代个体 1: 紫色圆圈
父代个体 2: 绿色圆圈

交叉点 cross

子代个体 1: 紫色圆圈, 绿色圆圈, 绿色圆圈, 绿色圆圈, 绿色圆圈, 绿色圆圈
子代个体 2: 紫色圆圈, 紫色圆圈, 绿色圆圈, 绿色圆圈, 绿色圆圈, 绿色圆圈

在GA中起关键作用, 是产生新个体的主要操作
单点交叉

遗传算法的组成

➤ **变异操作:** 编码按小概率扰动产生的变化, 类似于基因的突变

- 变异概率 P_m : 控制算法中变异操作的使用频率。 (随机数 $< P_m$, 可变异)
- 变异方式: 单点变异、均匀变异、高斯变异等。
- 单点变异: 如果对于某一基因位, 产生的随机数小于 P_m , 则改变该基因的取值。否则该基因不发生变异, 保持不变。

变异概率 P_m

父代个体

mutate

子代个体

随机改变染色体某些基因位上的遗传信息
变异操作只是GA中的次要操作

遗传算法的组成

➤ **环境选择:** 经过交叉、变异后, 得到一个子代种群, 从子代种群和父代种群中选择出下一代种群的过程, 称为环境选择。

- 选择方法: 非概率选择:
 - 代替换, 部分替换、精英选择、 $(\mu+\lambda)$ -选择、 (μ, λ) -选择
 - 概率选择: 轮盘赌选择、锦标赛选择。

代替换

部分替换

遗传算法的组成

➤ **轮盘赌选择 (roulette wheel selection):** 依据个体的适应度值计算每个个体在下一代中出现的概率, 并按照此概率随机选择个体构成子代种群。

$$p(I_i) = \frac{f(I_i)}{\sum_{i=1}^P f(I_i)}$$

其中 $f(I_i)$ 是个体 I_i 的适应度值; P 是种群大小

锦标赛选择(tournament selection): 每次从种群中取出一定数量个体, 然后选择其中最好的一个进入子代种群。重复该操作, 直到新的种群规模达到原来的种群规模

