编码与解码

GA 的关键问题之一是把解编码为染色体， 也要能把染色体解码为解. 常用的编码方法有

编码与解码

GA 的关键问题之一是把解编码为染色体， 也要能把染色体解码为解. 常用的编码方法有

- 常规码，即二进制码
- 实数码
- 根据问题确定的编码

二进制码

就是 $\mathbf{0-1}$ 编码. 采用 $\mathbf{0-1}$ 码可以精确地表示整数.

二进制码

就是 $0-1$ 编码. 采用 $0-1$ 码可以精确地表示整数. 用 $0-1$ 编码精确表示 a 到 b 所有整数， 只需编码长度 n 满足 $\frac{b-a}{2^n}<1$, 即 $n>\log_2(b-a)$.

二进制码

就是 $0-1$ 编码. 采用 $0-1$ 码可以精确地表示整数. 用 $0-1$ 编码精确表示 a 到 b 所有整数， 只需编码长度 n 满足 $\frac{b-a}{2^n}<1$, 即 $n>\log_2(b-a)$. 如满足 $0\le x\le 31$ 的整数 x 只需 5 位数的二进制码，如

$$10000 \to 16 \quad 11111 \to 31 \quad 01001 \to 9 \quad 00010 \to 2$$

二进制码

就是 $0-1$ 编码. 采用 $0-1$ 码可以精确地表示整数. 用 $0-1$ 编码精确表示 a 到 b 所有整数， 只需编码长度 n 满足 $\frac{b-a}{2^n}<1$, 即 $n>\log_2(b-a)$. 如满足 $0\le x\le 31$ 的整数 x 只需 5 位数的二进制码，如

$$10000\to 16 \quad 11111\to 31 \quad 01001\to 9 \quad 00010\to 2$$

如 $0-1$ 背包问题， 因为决策变量 $x_i\in\{0,1\}$， 所以可按 $(x_1,x_2,\cdots,x_n)$ 的取值形成一个二进制码.

实数码

对于连续的实数变量， 可以采用实数编码. 实数编码可以用实数本身作为实数码， 也可以在解空间与码空间中作个对应关系.

实数码

对于连续的实数变量， 可以采用实数编码. 实数编码可以用实数本身作为实数码， 也可以在解空间与码空间中作个对应关系. 如, 设 (x_1, x_2, x_3) 是以下解空间中的向量

$$\begin{cases} x_1 + x_2^2 + x_3^3 = 1 \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0. \end{cases}$$

我们可以用以下码空间中的染色体 (v_1, v_2, v_3) 来对解编码

$$v_1 \geq 0, \quad v_2 \geq 0, \quad v_3 \geq 0.$$

那么编码与解码的过程可以体现在以下关系上

$$x_1 = \frac{v_1}{v_1 + v_2 + v_3}$$

$$x_2 = \sqrt{\frac{v_2}{v_1 + v_2 + v_3}}$$

$$x_3 = \sqrt[3]{\frac{v_3}{v_1 + v_2 + v_3}}$$

连续的实数变量在一定精度下也可以采用二进制编码. 对给定的区间 $[a, b]$, 设二进制编码的长为 n, 则变量

$$x = a + a_1\frac{b-a}{2} + a_2\frac{b-a}{2^2} + \cdots + a_n\frac{b-a}{2^n}$$

与二进制码 $a_1a_2\cdots a_n$ 相对应. 二进制码与实际变量的误差为 $\frac{b-a}{2^n}$.

根据问题确定的编码

如 TSP 问题， 可用城市的一个序列来表示可行解.

根据问题确定的编码

如 TSP 问题， 可用城市的一个序列来表示可行解.

例：约束机器排序问题

n 个加工量为 $\{d_i \mid i = 1, 2, \cdots, n\}$ 的产品在一台机器上加工， 机器在第 t 个时段的工作能力为 c_t， 求出所有产品得以加工的最少时段数.

它的数学模型为

$$
\begin{cases}
\min \ T \\
\text{subject to} \\
\qquad \sum\limits_{t=1}^{T} x_{it} = 1, \ i = 1, 2, \cdots, n \\
\qquad \sum\limits_{i=1}^{n} d_i x_{it} \le c_t, \ t = 1, 2, \cdots, T \\
\qquad x_{it} \in \{0, 1\}, i = 1, 2, \cdots, n, t = 1, 2, \cdots, T.
\end{cases}
$$

其中 x_{it}, T 为决策变量， x_{it} 表示第 t 时段加工产品 i.

对于约束机器排序问题， 用 $0-1$ 编码是不太有效的. 一个有效的编码是给产品 $(1, 2, \cdots, n)$ 一个加工序 $i_1, i_2, \cdots, i_n)$， 由加工序按能力约束以最小的剩余能力一次安排时段加工， 在加工的过程中不允许改变产品的加工序.

例

设产品的加工量为： $d_1 = 5$, $d_2 = 3$, $d_3 = 10$, $d_4 = 4$；机器在前 5 个时段的工作能力为 $c_1 = 3$, $c_2 = 9$, $c_3 = 10$, $c_4 = 5$, $c_5 = 20$.

例

设产品的加工量为： $d_1 = 5$, $d_2 = 3$, $d_3 = 10$, $d_4 = 4$;
机器在前 5 个时段的工作能力为 $c_1 = 3$, $c_2 = 9$,
$c_3 = 10$, $c_4 = 5$, $c_5 = 20$.

若按 $(1, 2, 3, 4)$ 顺序加工， 则各时段加工的产品为：
第一时段不加工产品，
第二时段加工产品 1 和2，
第三时段加工产品 3，
第四时段加工产品 4.

例

设产品的加工量为： $d_1 = 5$, $d_2 = 3$, $d_3 = 10$, $d_4 = 4$;
机器在前 5 个时段的工作能力为 $c_1 = 3$, $c_2 = 9$,
$c_3 = 10$, $c_4 = 5$, $c_5 = 20$.

若按 $(1, 2, 3, 4)$ 顺序加工， 则各时段加工的产品为：
第一时段不加工产品，
第二时段加工产品 1 和2，
第三时段加工产品 3，
第四时段加工产品 4.

而最优加工序为 $(2, 1, 4, 3)$.