



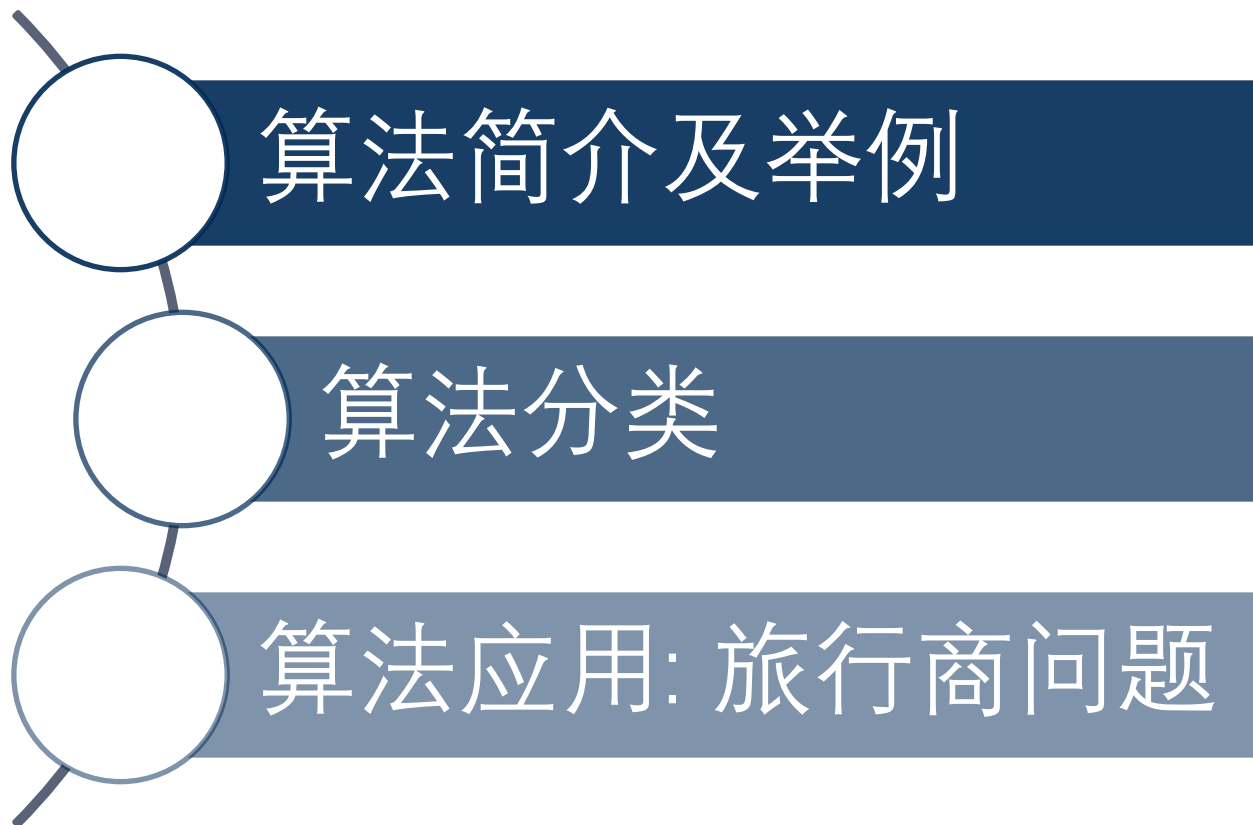
清华大学
Tsinghua University

ESTIMATION OF DISTRIBUTION ALGORITHM

分布估计算法

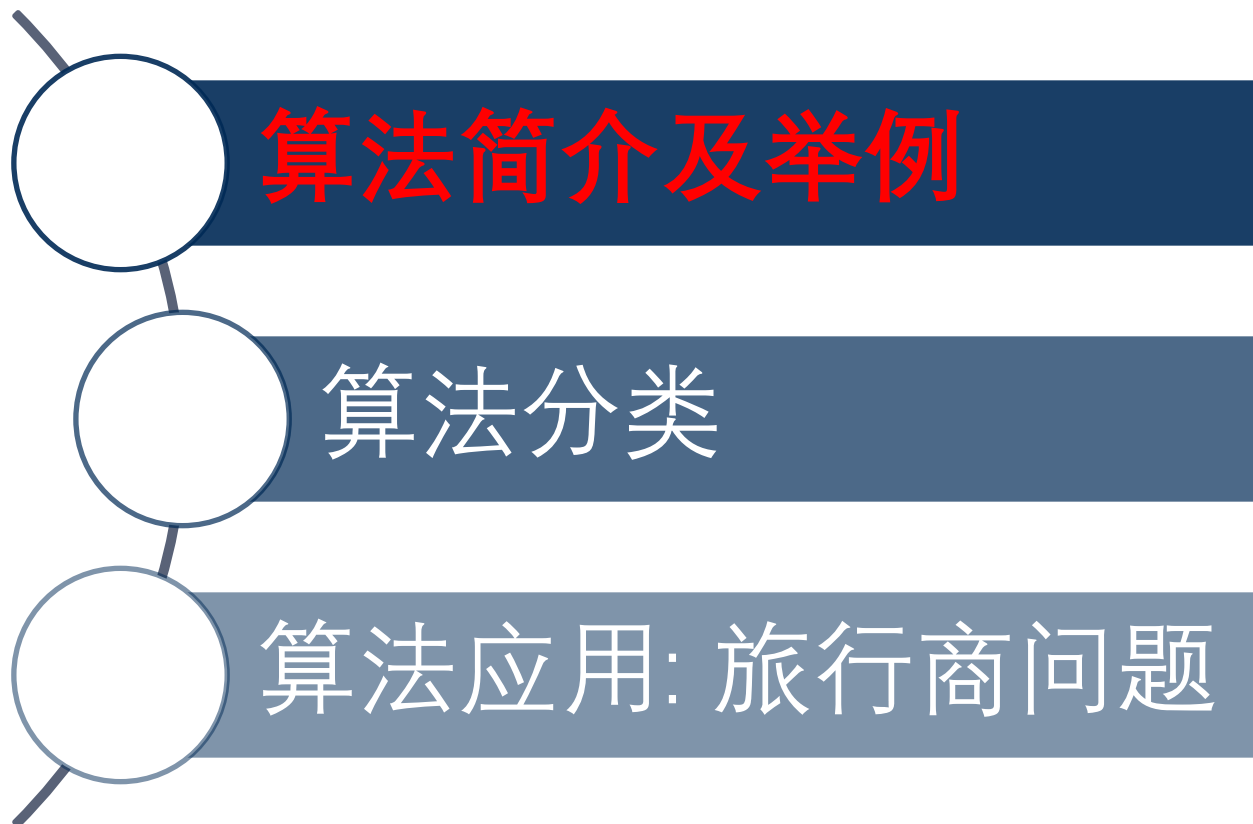


目录





目录





算法介绍

- 分布估计算法(Estimation of Distribution Algorithm, EDA)的概念最早在1996年提出,并迅速发展,目前相关理论研究和工程应用都取得了不少成果.
- EDA是一种基于统计学原理的随机群体进化算法,它提出了一种全新的进化模式,与GA相比,EDA中没有交叉、变异等针对个体的进化操作,取而代之的**概率模型**的采样、学习和更新。



举例说明^[1]

$$\max f(x) = \sum_i x_i$$

Subject to:

$$x_i \in \{0,1\}^n, n = 3$$

概率模型?

$$p = (p_1, p_2, \dots, p_n), p_i \in [0,1]$$

p_i 表示 $x_i = 1$ 的概率;
 $1 - p_i$ 表示 $x_i = 0$ 的概率.



举例说明

$$p = (p_1, p_2, \dots, p_n), p_i \in [0, 1]$$

STEP 1: 均匀初始化概率模型 $p^0 = (0.5, 0.5, 0.5)$, 对概率模型**采样**初始化种群 $B_0(N=8)$, 计算适应值函数如下所示.

	x_1	x_2	x_3	$f(x)$
1	0	0	1	1
2	1	1	0	2
3	0	0	0	0
4	0	1	1	2
5	0	1	0	1
6	1	0	0	1
7	1	0	1	2
8	1	1	1	3



举例说明

$$p^0 = (0.5, 0.5, 0.5)$$

STEP 2: 选择适应值较高的4个个体作为优势种群, **更新**概率模型 p . 考虑优势个体 X_s , 将优势种群频率作为新一代概率值($p_i^1 = P(x_i = 1|X_s)$)

	x_1	x_2	x_3	$f(x)$
1	0	0	1	1
2	1	1	0	2
3	0	0	0	0
4	0	1	1	2
5	0	1	0	1
6	1	0	0	1
7	1	0	1	2
8	1	1	1	3

	x_1	x_2	x_3	$f(x)$
2	1	1	0	2
4	0	1	1	2
7	1	0	1	2
8	1	1	1	3

$$avgf(x) = 1.5$$

$$p_1 = P(x_1 = 1|X_s) = 0.75$$

$$p_2 = P(x_2 = 1|X_s) = 0.75$$

$$p_3 = P(x_3 = 1|X_s) = 0.75$$

$$\Rightarrow p = (0.75, 0.75, 0.75)$$



举例说明

$$p^1 = (0.75, 0.75, 0.75)$$

STEP 3: 采样产生新一代个体, 可见群体适应度显著提高.

	x_1	x_2	x_3	$f(x)$
1	1	1	1	3
2	1	1	0	2
3	1	1	0	2
4	0	1	1	2
5	1	1	1	3
6	1	0	1	2
7	0	1	1	2
8	1	0	1	2

$$avgf(x) = 2.25$$

显然, 概率向量 \mathbf{p} 描述了各可能解在空间的分布情况, 产生任意解 $\mathbf{b} = (b_1, b_2, \dots, b_n)$ 的概率是?

$$\begin{aligned} P(\mathbf{b}) &= P(x_1 = b_1, x_2 = b_2, \dots, x_n = b_n) \\ &= \prod_{i=1}^n P(x_i = b_i) \\ &= \prod_{i=1}^n |1 - b_i - p_i| \end{aligned}$$

$$P^0(1,1,1) = 0.5 \times 0.5 \times 0.5 = 0.125$$



$$P^1(1,1,1) = 0.75 \times 0.75 \times 0.75 = 0.42$$



举例说明

$$p^1 = (0.75, 0.75, 0.75)$$

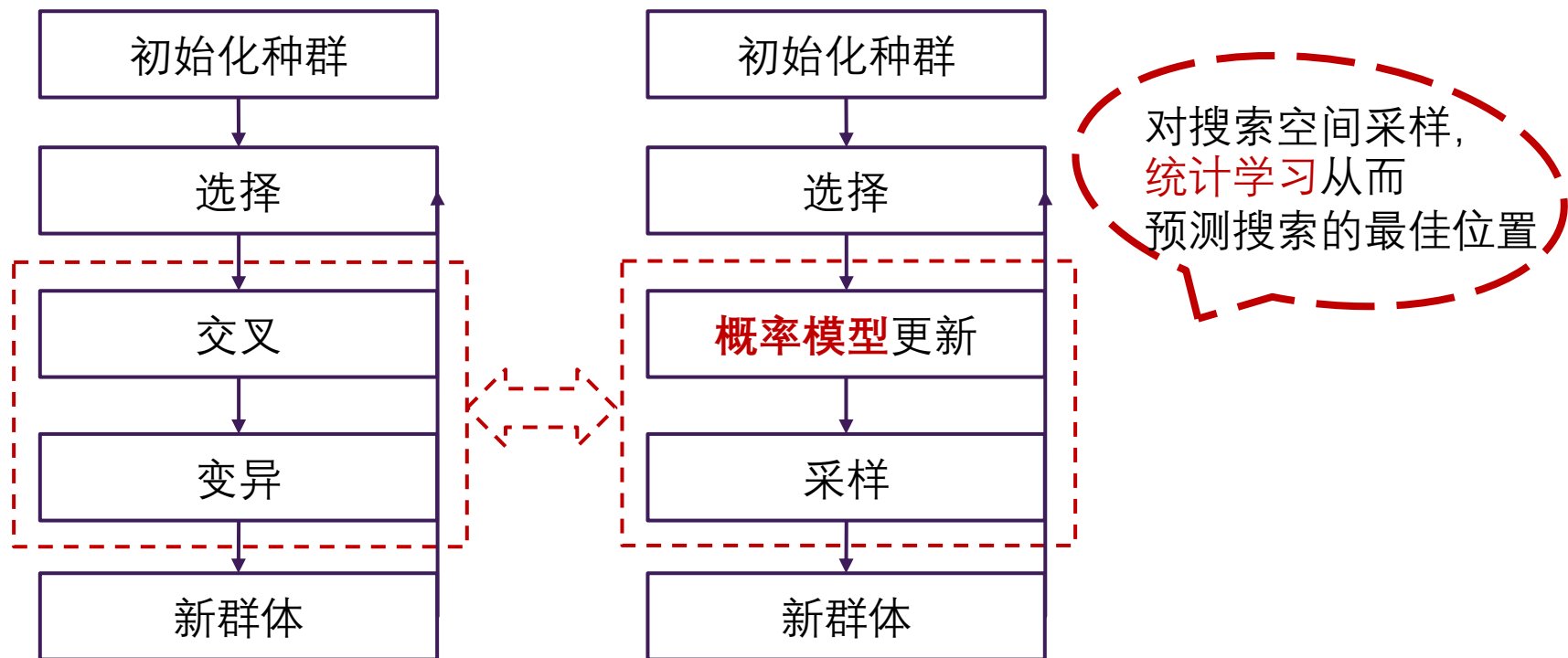
STEP 2: **选择**适应值较高的4个个体作为优势种群, **更新**概率模型 p .

	x_1	x_2	x_3	$f(x)$
1	1	1	1	3
2	1	1	0	2
3	1	1	0	2
4	0	1	1	2
5	1	1	1	3
6	1	0	1	2
7	0	1	1	2
8	1	0	1	2

	x_1	x_2	x_3	$f(x)$
1	1	1	1	3
2	1	1	0	2
3	1	1	0	2
5	1	1	1	3

通过以上步骤, 采样得到适应值高的个体的概率变大, 算法迭代进化后, 最终概率模型收敛, 得到问题的最优解.

对比GA



基于基因的微观层面进化
GA

基于搜索空间的宏观层面进化
EDA



算法流程

步骤1（初始化）：定义概率模型，初始化种群：

$D(0) \leftarrow N$ 个初始个体；

步骤2（进化）：

- 步骤2.1（选择）：选择精英种群： $D^{Se}(g) \leftarrow$

选择 $Q = N \cdot \eta\%$ 个精英个体；

- 步骤2.2（更新）：基于精英个体更新概率模型：

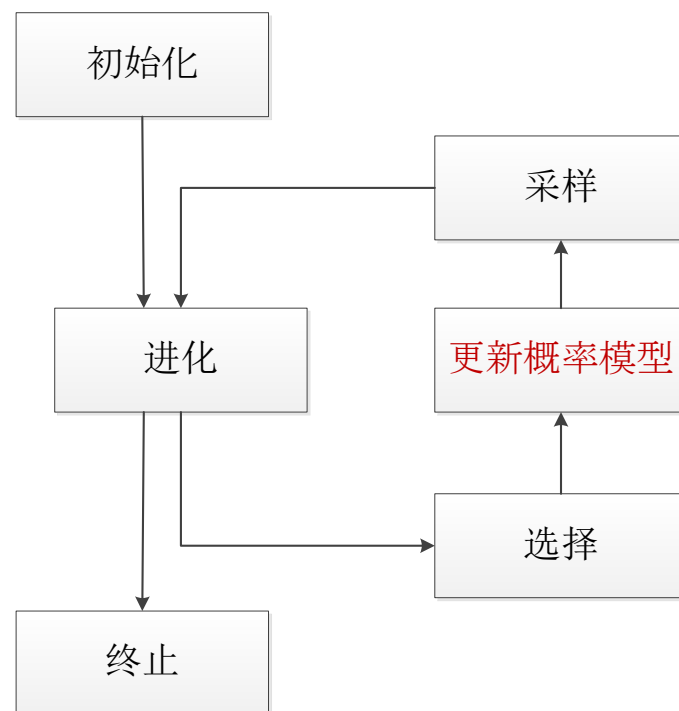
$p_g(x) = p(x|D^{Se}(g-1))$;

- 步骤2.4（采样）：随机采样生成新种群：

$D(g) \leftarrow$ 基于采样生成新种群；

步骤3（判断终止）：判断是否满足终止准则；若

是，则输出优化结果，否则转到步骤2.



[1] Larranaga P, Lozano J A. Estimation of distribution algorithms: A new tool for evolutionary computation[M]. Springer, 2002.

[2] 王凌, 王圣尧, 方晨. 分布估计调度算法 [M]. 清华大学出版社: 北京. 2017.

A decorative vertical line on the left side of the slide, consisting of three white circles connected by a dark blue line, with short diagonal lines at the top and bottom.

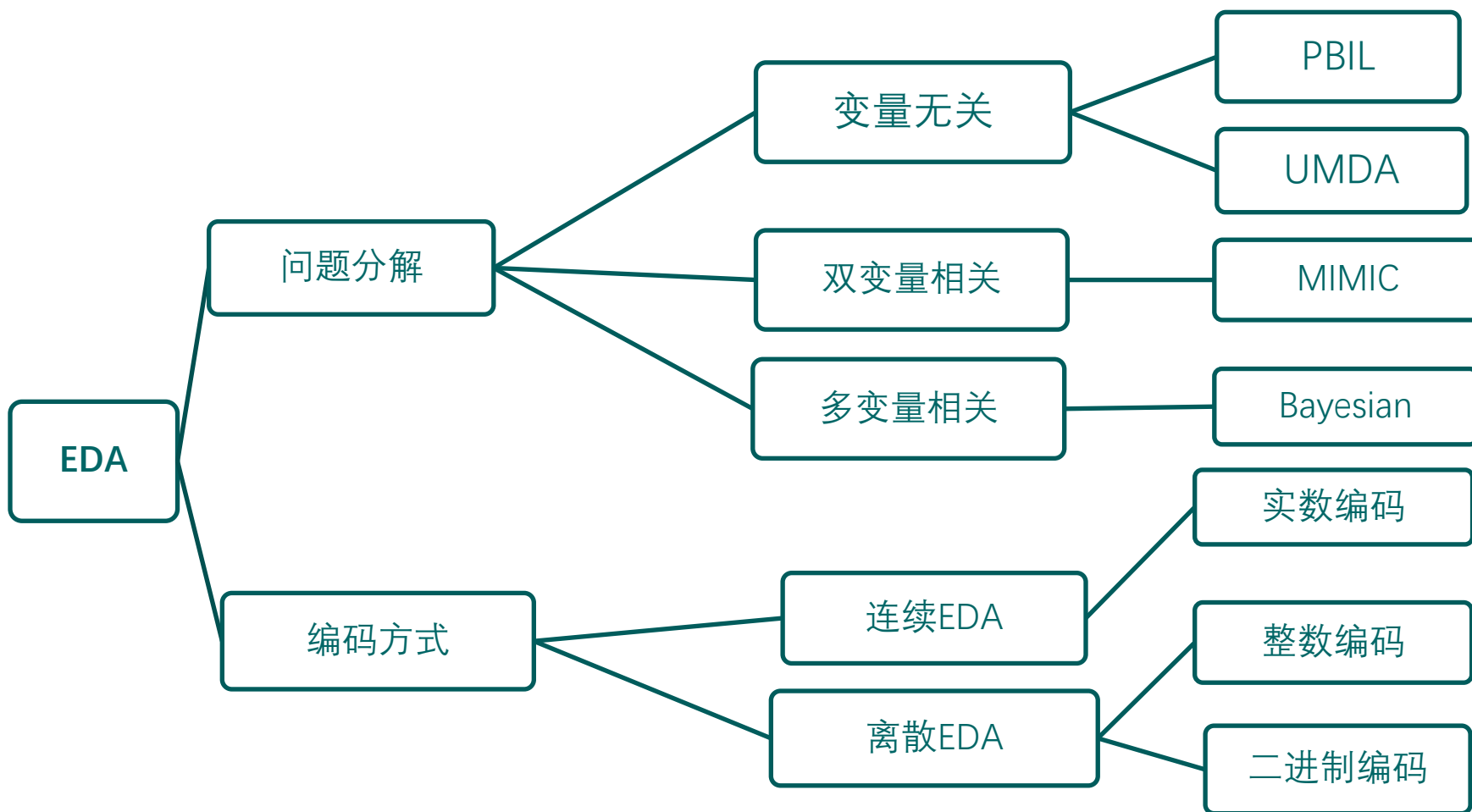
算法简介及举例

算法分类

算法应用: 旅行商问题



算法分类



问题分解方式^[3]

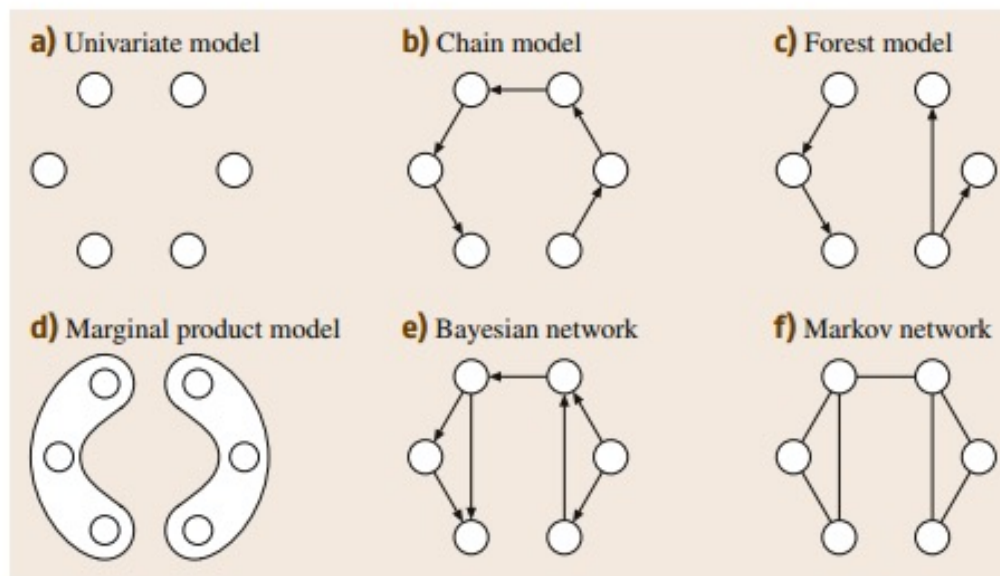


Fig. 45.2a–f Illustrative examples of graphical models. Problem variables are displayed as circles and dependencies are shown as edges between variables or clusters of variables. **(a)** Univariate model. **(b)** Chain model. **(c)** Forest model. **(d)** Marginal product model. **(e)** Bayesian network. **(f)** Markov network

变量无关模型:
$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i) .$$

双变量相关模型:
$$P(X_1, X_2, \dots, X_n) = \prod_{X_i \in R} P(X_i) \times \prod_{X_i \in X \setminus R} P(X_i | \text{parent}(X_i))$$

多变量相关模型:
$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{parent}(X_i))$$



变量无关模型 PBIL(Population Based Incremental Learning)

- PBIL算法是由美国卡耐基梅隆大学的Baluja在1994年提出的，用以解决二进制编码的优化问题。虽然在1994年“分布估计算法”这一概念尚未在学术界提出，但是PBIL算法公认是最早的EDA。

- 基本步骤:

STEP 1: 初始化概率向量，采样产生 N 个个体作为初始种群;

STEP 2: 选择 E 个优势个体($E < N$)更新概率向量:

$$p_{i,j}(g+1) = (1 - \alpha) \times p_{i,j}(g) + \alpha \times \sum_k I_{i,j}^k(g) / E$$

STEP 3: 根据更新后的概率向量采样生成新的个体;

STEP 4: 若终止准则未满足, $g++$, 否则输出最优解。

UMDA



双变量相关MIMIC (Mutual Information Maximization for Input Clustering)

- PBIL算法与UMDA算法都没有考虑变量之间的相互关系，算法中任意解向量的联合概率密度可以通过各个独立分量的边缘概率密度相乘得到
- 实际问题中，变量并不是完全独立的。在EDA的研究中，最先考虑变量相关性的算法是假设最多有两个变量相关，如由De Bonet等于1997年提出MIMIC算法
- 在MIMIC算法中，变量之间的相互关系是一种**链式关系**，描述解空间的概率模型如下：

$$p_{\pi}(x) = p(x_{i_1}|x_{i_2}) \cdot p(x_{i_2}|x_{i_3}) \cdots p(x_{i_{n-1}}|x_{i_n}) \cdot p(x_{i_n})$$

其中， $\pi = (i_1, i_2, \dots, i_n)$ 表示变量 (x_1, x_2, \dots, x_n) 的一个排列

- 每一次循环，根据选择的优势群体构造概率模型 p_{π} ，并由 p_{π} 采样产生新个体。
对第 i_n, i_{n-1}, \dots, i_1 个变量依次采样，即首先根据 i_n 的概率分布 $p(x_{i_n})$ 随机采样产生第 i_n 个变量，然后根据第 $x_{i_{n-1}}$ 的条件概率分布 $p(x_{i_{n-1}}|x_{i_n})$ 随机采样产生第 i_{n-1} 个变量.....

A decorative vertical line on the left side of the slide, consisting of three white circles connected by a dark blue line, with short diagonal lines at the top and bottom.

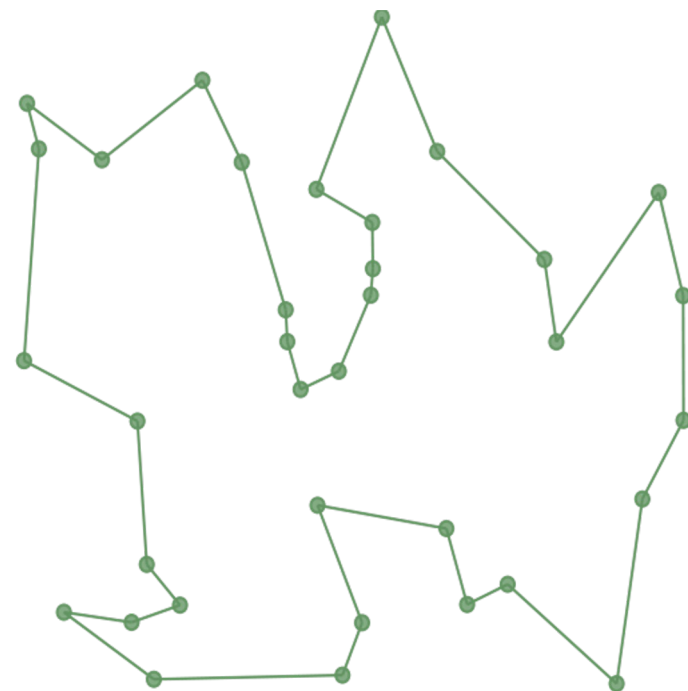
算法简介及举例

算法分类

算法应用: 旅行商问题

应用：旅行商问题

- 推销员到 n 个城市推销商品，求包含 n 个城市的最短环路
- 解空间大小 $(n-1)!$, NP-hard问题





应用：旅行商问题

- 城市编号 $\{1, 2, \dots, n\}$, 用长度为 n 的数字串 $C_1 C_2 \dots C_n$ 表示旅行顺序, 其中 C_i 表示第 i 个到达的城市编号;
- 路径长度(评价值函数)计算:

$$dis = \sum_i D(C_i, C_{(i+1)\%n})$$

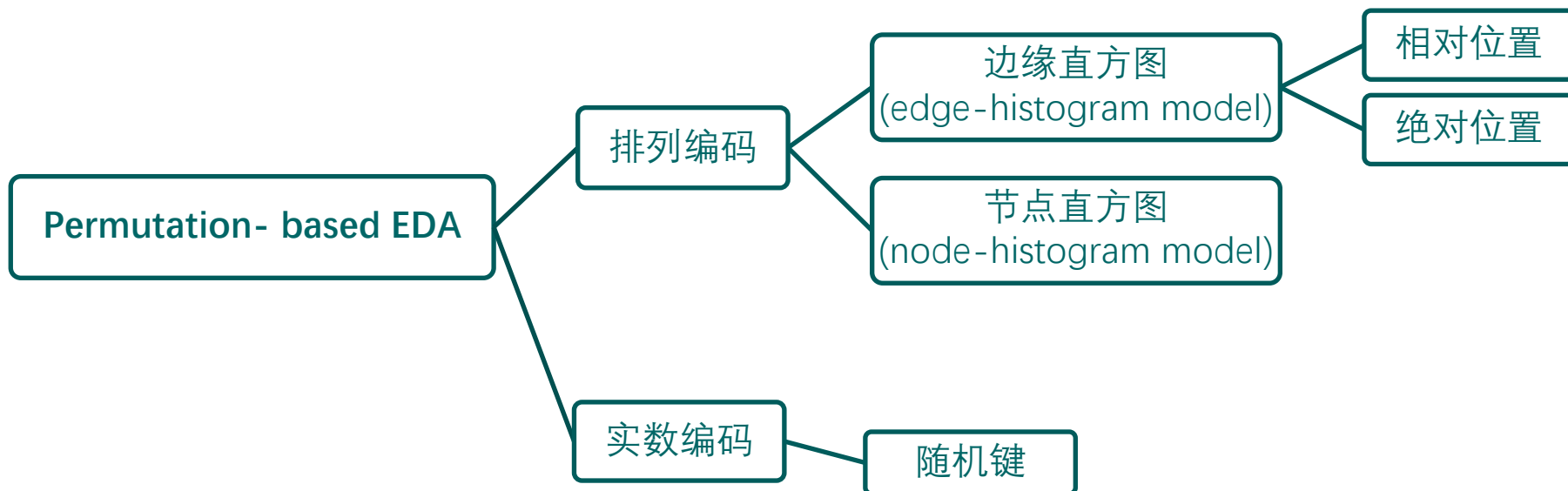
其中, $D(i, j)$ 表示城市 i 和 j 之间的距离。

如何建立概率模型？



求解排列优化问题的EDA

概率模型的建立与编码相关

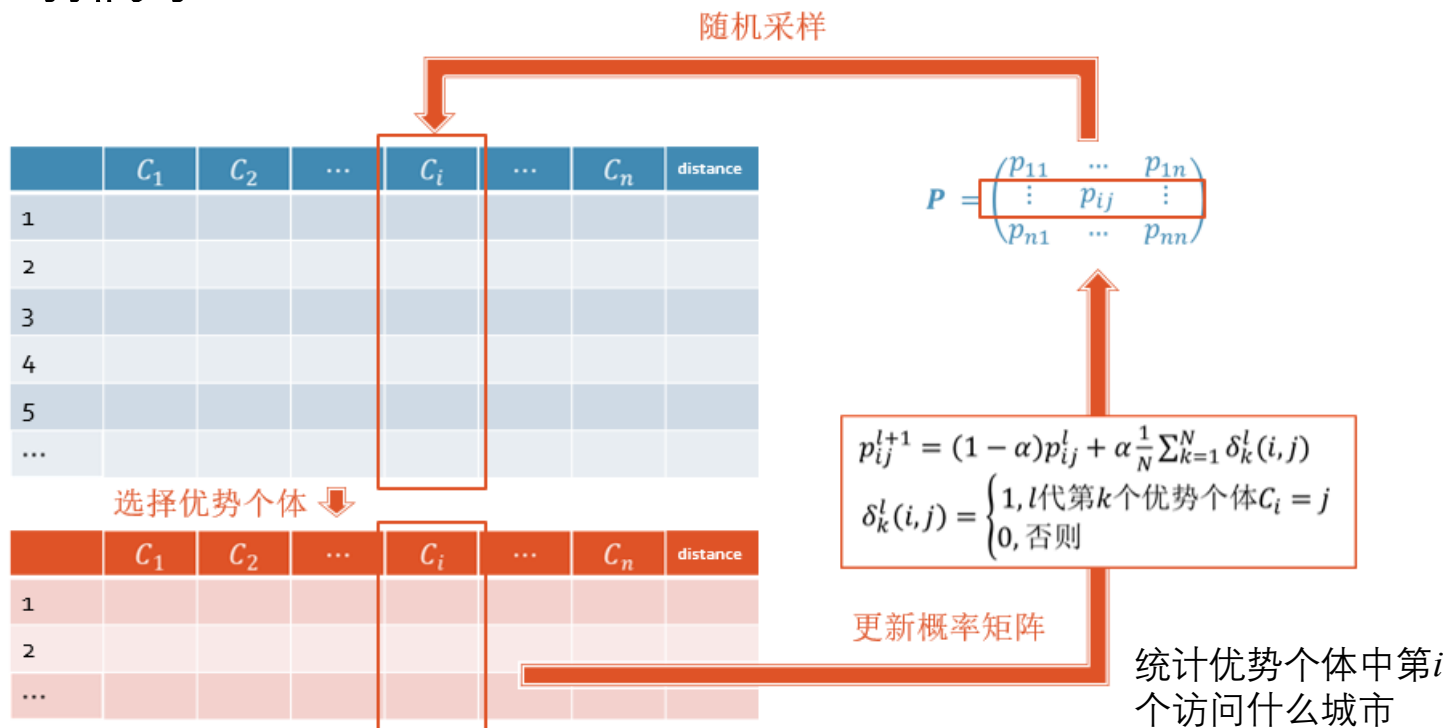


变量无关：绝对位置模型

➤ $P = \{p_{ij}\}_{n \times n}$

➤ 其中 p_{ij} 表示 $C_i = j$ (第 i 个到达的城市是 j) 的概率

➤ 采样得到排列？

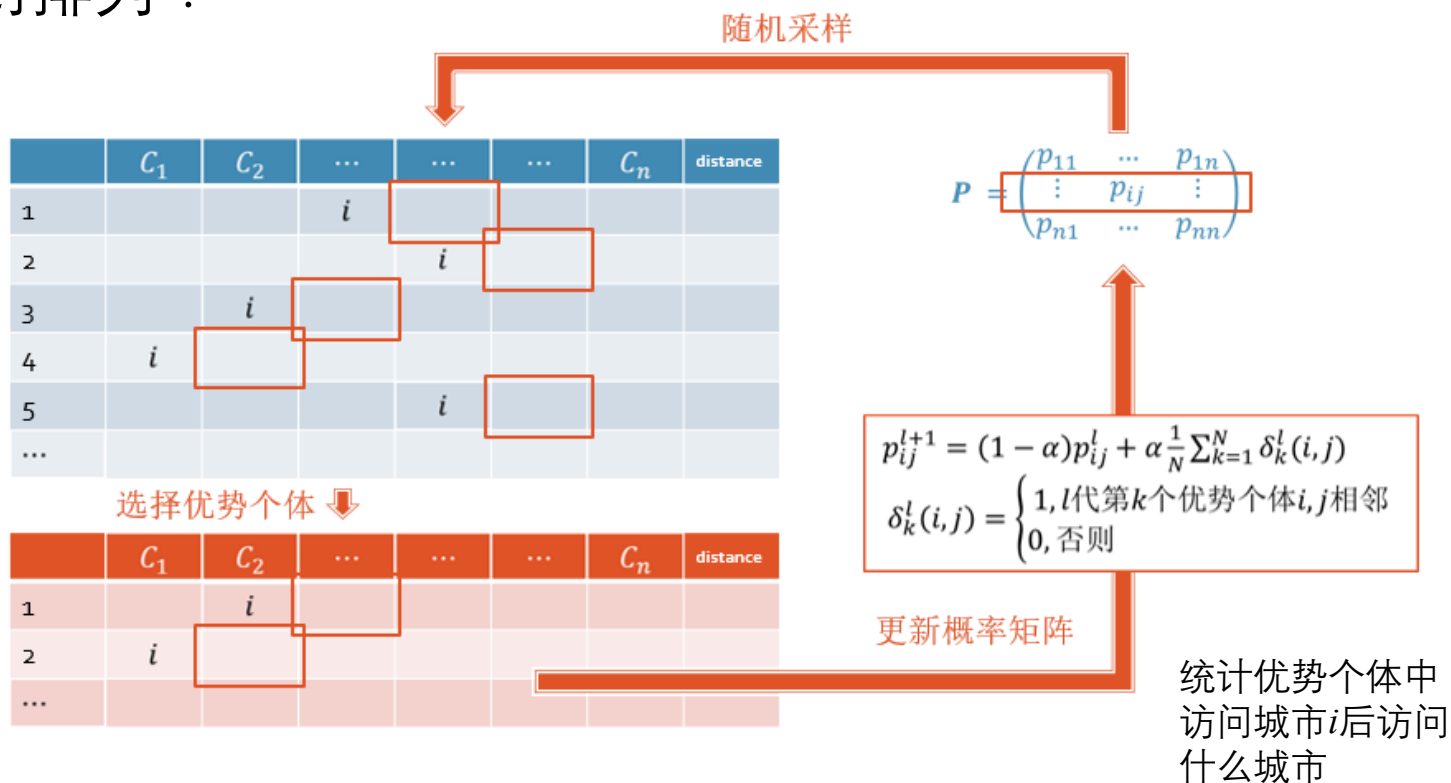


双变量相关：相对位置模型

➤ $P = \{p_{ij}\}_{n \times n}$

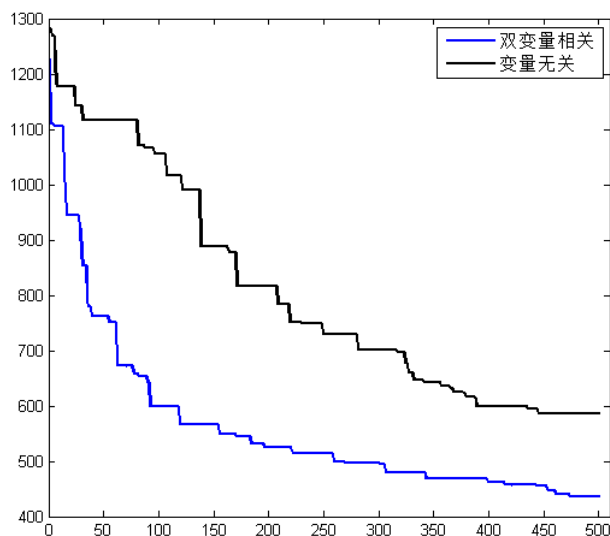
➤ 其中 p_{ij} 表示 $C_k = i, C_{k+1} = j$ (城市 i 之后访问城市 j 的概率);

➤ 采样得到排列？

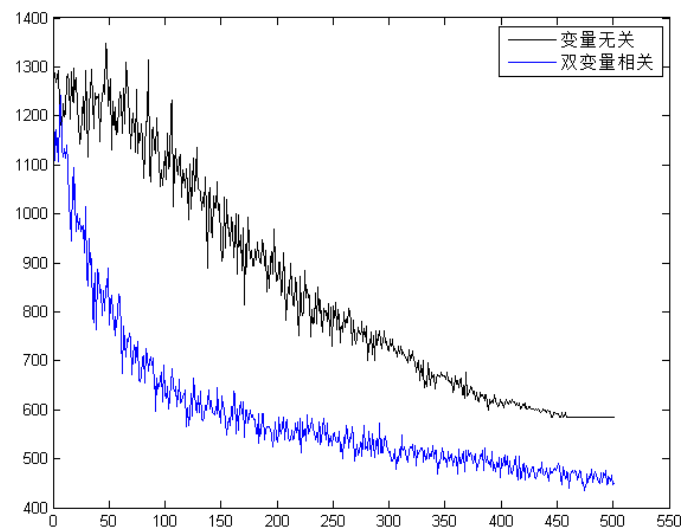


概率模型性能比较

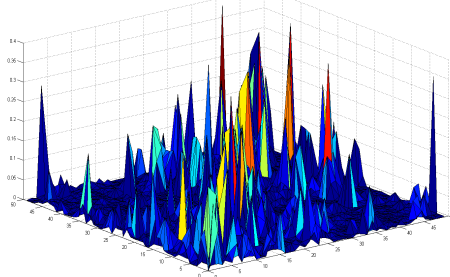
最优解的变化曲线



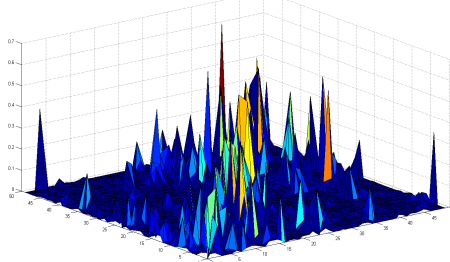
当前种群最优解的变化曲线



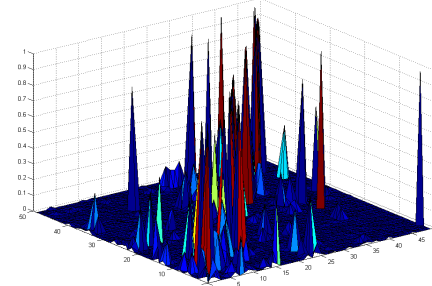
进化100代



进化300代



进化500代



相对位置概率模型收敛过程



总结

- 针对不同类型的优化问题，需要设计不同的概率模型来描述解空间的分布。
EDA在解决非线性和变量耦合的优化问题时能够利用问题信息，从而产生更好的个体。
- EDA基于群体的宏观进化方式使得其可以利用解空间的全局信息和进化过程中的历史信息，从而具有更强的全局搜索能力和更快的收敛速度。
- EDA对解空间的分布进行估计并采样产生新个体的方法，更容易使其作为一种手段和框架与其它算法混合，增强寻优性能。



Reference

- [1] 周树德, 孙增圻. 分布估计算法综述[J]. 自动化学报, 2007, 33(2): 113-124.
- [2] 王圣尧, 王凌, 方晨, 等. 分布估计算法研究进展 [J]. 控制与决策, 2012, 27(7): 961-966.
- [3] Larranaga P, Lozano J A. Estimation of distribution algorithms: A new tool for evolutionary computation[M]. Springer Science & Business Media, 2002.
- [4] M. Pelikan, M. W. Hauschild, F.G. Lobo, Estimation of Distribution Algorithms, in: J. Kacprzyk, W. Pedrycz, Handbook of Computational Intelligence, Berlin Heidelberg: Springer, 2015.
- [5] 王凌, 王圣尧, 方晨. 分布估计调度算法 [M]. 清华大学出版社: 北京. 2017.
- [6] J. Ceberio, E. Irurozki, A. Mendiburu, et al, A review on estimation of distribution algorithms in permutation-based combinatorial optimization problems, Prog. Artif. Intell. 1 (1) (2012) 103–117.



清华大学
Tsinghua University

谢谢大家！