

# Review of Natural Computing

## 课堂练习

### 自然计算的定义是什么？有什么特点？主要研究分支有哪些？

绪论中有。定义：以自然界（包括...），尤其是生物体的功能、特点和作用机理为基础，研究其中丰富的信息处理机制，抽取相应的计算模型，设计相应的算法，并在各种领域加以应用。

特点（特征）：**比喻性**地使用**自然系统**潜在的概念、原理和机制。有记忆性；自学习、自组织、自适应  
主要研究分支：进化计算、人工神经网络、模拟退火算法、群体智能（蚁群粒子群）、免疫计算（人工免疫算法）、混沌优化、分布估计算法

### 进化计算具有什么特点，他通常应用于哪些领域

特点7点：有指导搜索、自适应搜索，渐进式寻优、并行式搜索、黑箱式结构、全局最优解、通用性。

应用：人工智能、人工生命、复杂问题的优化、结构性优化。

### 试画出简单遗传算法的流程图并描述求解过程

编码、初始化种群、迭代（计算适应度、选择、交叉、变异）、判断是否满足终止条件

### 什么叫模拟退火算法？组合优化与模拟退火有哪些相似性

模拟固体退火过程的一种拟物智能算法。受物理退火启发，模拟固体的升温、等温、冷却过程，结合Metropolis准则进行科学计算的一种启发式随机寻优算法。

解-粒子状态

最优解-最低能量状态

控制参数-温度

目标函数-能量

初始参数-熔化

Metropolis随机抽样-等温

控制参数的下降-冷却

### Metropolis准则与冷却进度表在模拟退火算法中起到什么作用

Metropolis准则：接受准则判断是否接受新解，抽样准则判断是否达到平稳分布，是否进入下一个状态。

冷却进度表：记录了一系列重要参数。初始温度、温度衰减、马尔科夫链长度、结束条件等。

### 试画出并论述分布估计算法的基本流程

初始化种群和分布、适应度计算、选择优势个体、估计优势个体概率分布建立概率模型、对模型采样得到新一代种群

## **群智能算法与进化计算的异同点有哪些**

同：

1. 都是仿生算法
2. 都是基于种群
3. 一种随机搜索优化算法
4. 根据个体适配信息进行搜索，不受函数的限制（通用）
5. 隐含并行计算
6. 全局最优
7. 可能无法收敛

异：

1. 群智能保留记忆在个体中，进化计算记忆随种群改变
2. 群智能的信息共享是单向的，进化计算种群共享信息
3. 群智能没有编码和遗传操作，参数更少、原理更简单、更易实现

## **试论述蚁群算法求解TSP问题的基本流程**

初始化蚁群和分布，添加禁忌表，禁忌表未满则一直循环，根据信息素和启发式信息选择路径，添加禁忌表，禁忌表满，更新信息素，判断结果是否满足要求，不满足清空继续循环。

## **粒子群算法的基本原理和特点是什么**

鸟类觅食，追逐最优个体，一群随机粒子，根据惯性（自身运动的信任）、个体最优（个人经验）和全局最优（群体经验）进行搜索。

特点：有记忆性；自适应、自学习、自组织；灵活、稳健、自我组织；全局最优、并行、随机搜索、根据个体适配搜索，不依赖函数；简单参数少。

## **模拟退火算法的三大支柱**

新解接受准则和随机数产生器（Metropolis算法）

冷却进度表

新解产生器和邻域结构（邻域结构由新解产生器决定）

## **分布估计算法中没有遗传操作，取而代之的是**

建立概率模型和采样

## **蚂蚁行进时会释放\_\_\_\_作为群体内通讯物质**

信息素

## **免疫算子由\_\_\_\_和\_\_\_\_两个操作完成**

接种疫苗、免疫选择

## 什么是遗传算法？基本原理和流程图

遗传算法是一类借鉴生物界自然选择和自然遗传机制的随机化搜索算法

### 生物免疫系统的两种应答方式

初次应答和二次应答

## 1 绪论

人工智能是指给定问题、约束、目标和知识库，系统通过学习（获取信息、完善知识、创建策略）解决问题的能力。**人工智能以模拟自然计算为基础，利用人工机器模拟人的智能。**

- 研究自然现象尤其是生命体的功能、特点和作用机理，建立相应的计算模型，服务于人类社会的工程应用，是自然计算的宗旨（师法与自然，作用于自然）。
- 自然计算的研究有利于揭示生命现象，增进人们对自然界的认识。与此同时，来源于自然界的计算模型也能很好地服务于人类的生产实践活动，在优化、控制、网络安全、创造性设计等领域具有很好的用途，对推动社会发展有着积极的意义。

- Leiden自然计算研究中心的研究者们认为，自然计算是一种表示由自然启发的计算的一般性术语，而自然计算的研究就包括从理论上和经验上对由自然启发的计算的理解。
- 自然计算的特征就是比喻性地使用自然系统潜在的概念、原理和机制。
  - 进化算法使用生物学中变异、基因重组和自然选择的概念；
  - 而神经网络由大脑和神经系统中相互高度连接的神经结构启发而来；
  - 分子计算使用DNA分子和酶来设计算法并实现生物硬件，等等。
- 一般而言，自然计算是指以自然界（包括生命、生物及生态系统、物理、化学、经济以及社会文化系统等），特别是生物体的功能、特点和作用机理为基础，研究其中所蕴含的丰富的信息处理机制，抽取相应的计算模型，设计相应的算法，并在各相关领域加以应用。

以自然界的 功能、特点、作用机理为基础→研究信息处理机制→抽取模型→设计算法→加以应用

交叉学科（自然科学+计算机科学），尤其是生命科学和计算机科学

目前，自然计算模式的研究主要包括：

- ◆ 进化计算、神经网络计算、群体智能（蚁群算法、微粒群算法、蜂群算法等）、生物启发式计算、量子计算、光子计算、自组织迁移算法、差分进化、分布估计算法、生态计算、分子计算、DNA计算、元胞自动机、免疫计算（人工免疫系统、克隆选择等），模拟退火算法、化学计算、人工内分泌系统、文化算法、社会认知算法、Memetic计算、混沌优化、情感计算及其他相关复杂自适应系统等，

表 1.1 自然计算启发源与要素

启发源属性		自然计算模式	对应的自然系统	要素
物质层次	物理	模拟退火算法	退火	原子
		量子计算	原子	量子
		光子计算	光	光子
		混沌优化	混沌系统(物理现象)	物质
		云计算	大气	云
	化学	化学(反应)计算	化学反应过程	分子
		分子计算	分子体系	分子

启发源属性		自然计算模式	对应的自然系统	要素
生命层次	生命系统	进化计算	遗传系统	基因积木块(染色体)
		差分进化	生物进化系统	种群个体
		神经计算	神经网络	神经元
		免疫计算	免疫系统	抗体
		内分泌计算	内分泌系统	分子
		元胞自动机	元胞动态系统	细胞
		膜计算	细胞膜	分子
		DNA 计算	DNA 双螺旋	分子
		蛋白质计算	蛋白质	分子
		进化计算	遗传系统	基因积木块(染色体)
	生物(群)	差分进化	生物进化系统	种群个体
		蚁群算法	蚁群	蚂蚁
		微粒群算法	鸟群	鸟
		蜂群算法	蜂群	蜜蜂
		鱼群算法	鱼群	鱼
		细菌觅食算法	菌群	细菌
		群搜索优化	生物群	鸟、鱼、狮子
		细菌(群)趋药性算法	细菌(群)	细菌
	生态系统	搜索者优化	人群	人
		生物地理学优化	物种、种群	生物
		自组织迁移算法	种群	生物
		人口迁移算法	人群	人口

- 自然计算其本质是借鉴自然界的[功能与作用机理](#)抽象出的计算模型，其研究必然涉及现代自然科学的方方面面，相关领域非常广泛。

## 进化计算

真正的"进化计算" (evolutionary computation)  
这一术语是在20世纪 90年代初被提出的，它是模拟生物进化过程中的[自然选择机制](#)和[信息遗传规律](#)的算法的总称。

模拟自然选择机制和信息遗传规律的算法的总称

## 复杂适应系统

包括受化学、物理现象启发的智能算法，如模拟退火、混沌优化

# 2 自然计算的研究综述与统一模型

## 研究方向

### 进化计算

遗传算法 (GA)  
进化策略 (ES)  
进化规划 (EP)  
遗传编程 (GP)

### 群体智能

蚁群算法 (ACO)  
粒子群算法 (PSO)

## 免疫计算

人工免疫系统 (AIS)

克隆选择

## 模拟退火算法

模拟退火算法 (SA)

## 复杂适应性系统

经济、生态、神经、计算机网络等系统行为的一般规律

## 进化计算

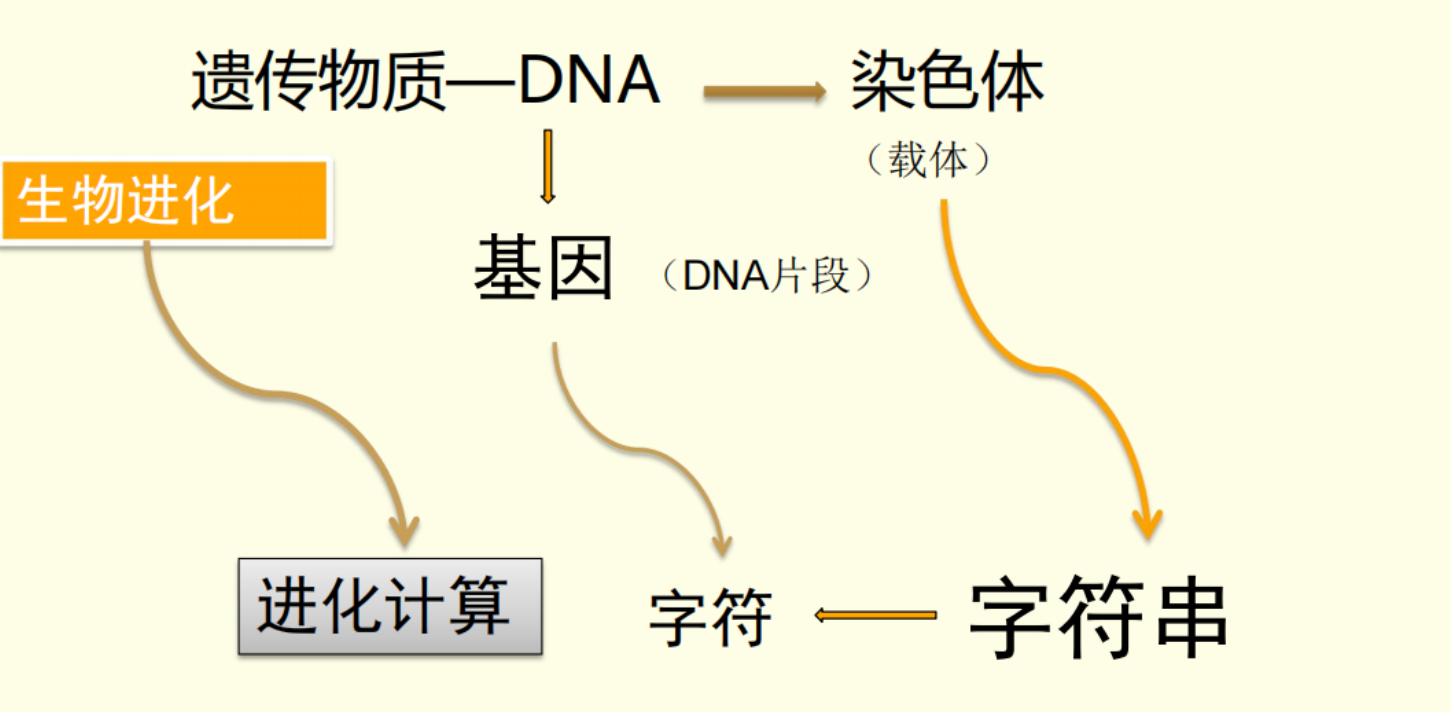
- 高鲁棒性和广泛适应性的全局优化方法，
- 具有自组织、自适应、自学习的特性，能够不受问题的限制，有效地处理传统优化算法难以解决的复杂问题。

多点并行分布式全局优化搜索算法

- 生物群体的生存过程普遍遵循达尔文的物竞天择、适者生存的进化准则，生物通过个体间的**选择、交叉、变异**来适应大自然环境。
- 在生物繁殖过程中，大多数生物通过遗传使物种保持相似的后代，部分生物由于变异，后代具有明显的差别，甚至形成新物种。正是由于生物的不断繁殖，生物数目大量增加，而自然界中生物赖以生存的资源却是有限的，因此，为了生存，生物就需要竞争。生物在生存竞争中，根据对环境的适应能力，**适者生存，不适者消亡**。自然界的生物，就是根据这种优胜劣汰的原则，不断地进行进化。

进化计算就是借鉴生物进化的规律，通过繁殖—竞争—再繁殖—再竞争，实现优胜劣汰，一步一步地逼近问题的最优解。进化计算中的“进化”就由此得名。

映射关系



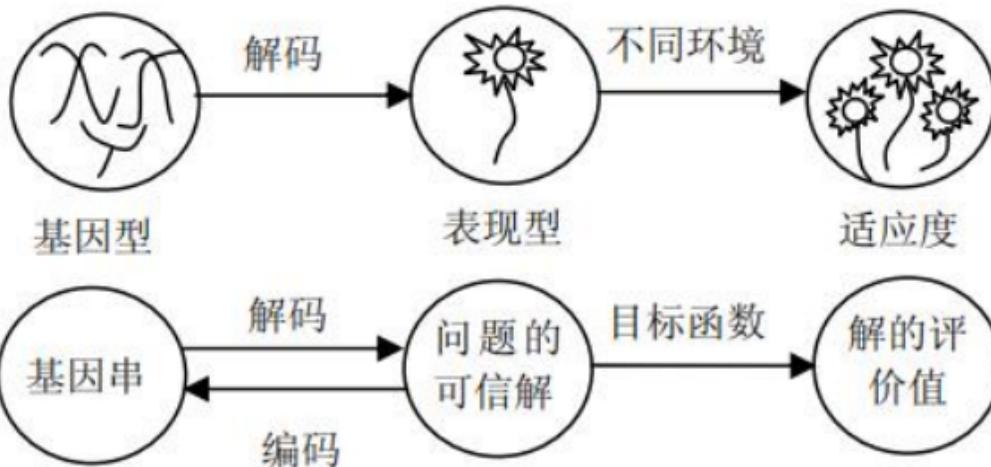
遗传学	GA
染色体 (Chromosome)	数据、数组、位串
基因 (gene)	位
等位基因 (allele)	特性值
基因座 (locus)	串中位置
基因型 (genotype)	解结构 (基因空间)
表现型 (phenotype)	解结构 (问题空间)

## 历史和分支

### 遗传算法

60年代 最早提出

## 自然进化



## 遗传规划 (了解一下就行)

- 20世纪90年代美国斯坦福大学的J.R.Koza教授提出了遗传规划
- 基于群体的、具备随机和定向搜索特征的迭代过程
- 优点：不需要给出具体的函数形式  
在初始群体足够大而且交叉和变异概率设置合理的情况下，不会陷入局部优化
- 和遗传算法不同点：遗传算法的个体是一个定长的字符串，GP的个体是一个函数表达式--二叉树来表示

## 进化策略

### 定义

模拟生物进化过程中的自然选择机制和信息遗传规律的算法的总称

### 进化计算的特点

数学上的实质为搜索寻优算法

## □ 特征

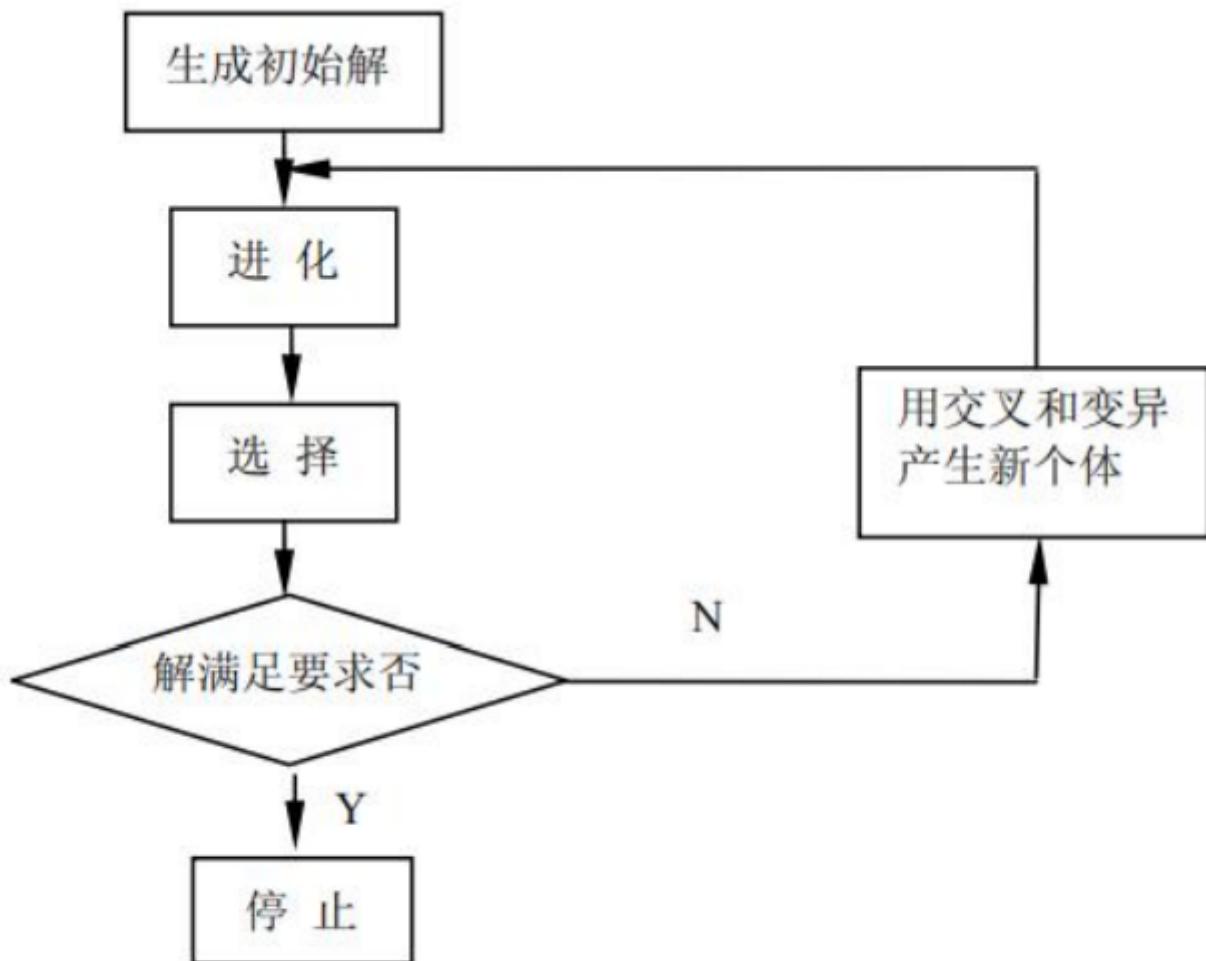
- 1) 有指导搜索——依据是适应度，也就是它的目标函数
- 2) 自适应搜索——无需添加任何额外的作用
- 3) 渐进式寻优——反复迭代
- 4) 并行式搜索——每代运算都是针对一组个体同时进行
- 5) 黑箱式结构——用字符串表达 问题 及 适应度
- 6) 全局最优解——多点并行搜索
- 7) 通用性强——框架式算法，简单的原则要求，在过程中无需干预

有指导、自适应

渐进式寻优、并行式搜索、黑箱式结构

全局最优、通用

# 流程



进化算法工作流程

# 应用

## 3.1.4 进化计算的应用

- 进化计算是一种新型的优化技术，它仿效生物界的进化与遗传，弥补传统优化技术的不足。

### 1) 人工智能

- 继模糊数学、专家系统、人工神经网络之后，成为处理人工智能的又一个有力工具。
- 机器学习中的规则优化，分类器系统中知识的自动获取、筛选，进化计算与人工神经网络的融合，图像识别

2) 结构性优化----工程技术的优化包括结构优化和参数优化。

3) 复杂问题的优化----非线性、多峰值、不确定性

4) 人工生命----人工生命是研究用人工的方法模拟自然生命的特有行为，而基于进化计算的模型是研究人工生命的主要基础理论之一。

5) 综合应用----交叉渗透，相互促进

主要还是搜索寻优问题

## 简单遗传算法SGA

### 解码与编码

编码方式：

色体，有时也叫个体。

二进制编码、格雷编码、实数  
编码、符号编码、排列编码等

长度（位数）为 L 的 0/1 字符串

$$x = x_{\min} + \frac{x_{\max} - x_{\min}}{2^L - 1} Dec(y)$$

$x_{\min}$ 、 $x_{\max}$  —— 变量  $x$  所取的最小及最大的十进制数值；

$y$  —— 对应于  $x$  的二进制数；

Dec —— 将二进制数转换为十进制数

$$\delta = \frac{x_{\max} - x_{\min}}{2^L - 1}$$

[例] 优化问题中，精度要求 $\delta = 10^{-6}$ ，对于 x

$$2^L = \frac{x_{\max} - x_{\min}}{\delta} + 1 = \frac{2+1}{10^{-6}} + 1 = 3000001$$

$$L = \lg 3000001 / \lg 2 \approx 21.5165$$

本例的二进制编码长度至少需要 22 位

## 产生初始群体

随机方法产生初始种群M

M越大，搜索范围越广，但遗传操作速度也越慢

一般取50-100

## 计算适应度

### 选择

轮赌选择法

- 选择操作也叫复制操作，根据个体的适应度函数值所度量的优、劣程度决定它在下一代是被淘汰还是被遗传。
- 简单遗传算法采用J.H.Holland教授推荐的赌轮选择法，赌轮选择法的基本思想是个体被选中的概率取决于个体的相对适应度：

$$p_i = f_i / \sum f_i$$

## 交叉操作

参数pc

- 交叉是遗传算法区别于其他进化算法的重要特征，它在遗传算法中起关键作用，是产生新个体的主要方法。
- 从一对父辈产生一对子辈
- 单点交叉
- 交叉概率  $p_c$ 
  - typically in range [0.5, 1.0]
- Crossover 是GA最本质的遗传算子：
  - 在进化中，能加速搜索
  - 能形成schemas的有效组合(subsolutions on different chromosomes)

GA区别于其他进化算法的重要特征，最本质的遗传算子

整个选择后的新群体中，究竟有多少个体要进行交叉操作需要一个参数来控制，控制被交换个体数目参数  $M_c$  可表达为：  $M_c = p_c \cdot M$

示例中，假设  $p_c$  取 0.3，由于群体规模  $M = 30$ ，故有  $M_c = p_c \cdot M = 9$ ，即要有 9 个个体要进行交叉操作，然而 9 为奇数，所以需要随机选择增加或删除一个个体，假定选择的结果是删除，那么就有 8 个个体要进行交换。

交换前	$v_2: 11011100010101   00001100$ $v_{30}: 11000100010101   00001110$
交换后	$v'_2: 11011100010101   00001110$ $v'_{30}: 11000100010101   00001100$

## 变异

参数  $pm$

- 在变异操作中，用变异概率  $p_m$  确定需要变异的个体。因为变异是针对字符进行的，所以变异概率也是针对字符而言，即：

$$p_m = \frac{B}{M \cdot L}$$

- 通常， $pm$  约为 0.005 ~ 0.01。

## 终止条件

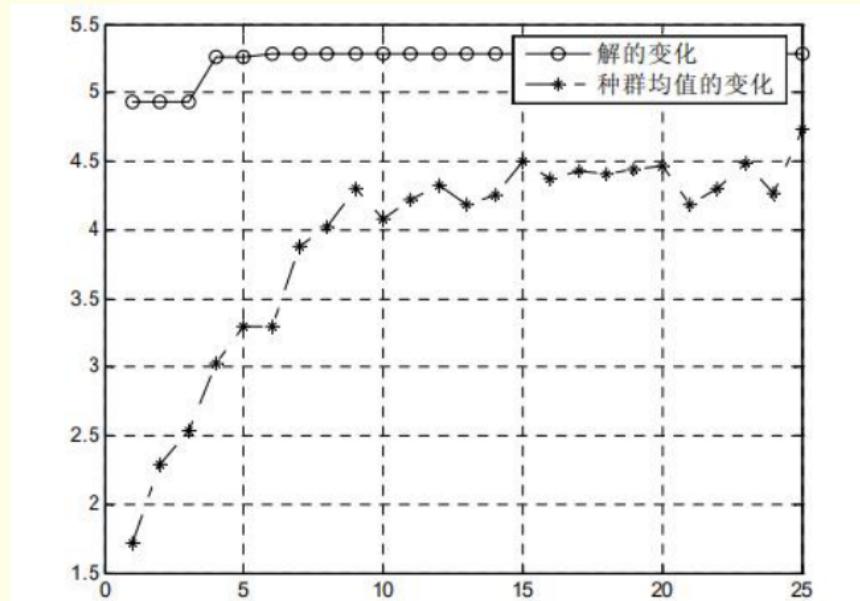
- 遗传算法是一个反复迭代的过程，每次迭代期间，要执行适应度计算、选择、交叉、变异等操作，直至满足终止条件。
- 常用的遗传算法终止条件主要有：
  - 1) 规定最大迭代次数  $N$ 
    - ◆ 当某种操作的迭代次数达到  $N$ ，操作停止并输出结果。一般  $N$  取  $200 \sim 500$  次。由于遗传算法中有许多随机因素的影响，所以，最后一代的个体不一定含有最优个体，所以要记录每代的最优个体以便于最后比较。
  - 2) 规定最小的偏差  $\delta$ 
    - 对于适应度目标已知的遗传算法，如用方差作为适应度计算的曲线拟合问题，可用最小的偏差  $\delta$  作为终止条件，即：

$$|f_{\max} - f'| \leq \delta$$

### ■ 3) 观察适应度的变化趋势

- ◆ 在遗传算法初期，最优个体的适应度以及群体的平均适应度都较小，以后随着选择、交叉、变异等操作，适应度值增加，到了算法后期，这种增加趋缓甚至停止，在这种情况下，应终止遗传算法。

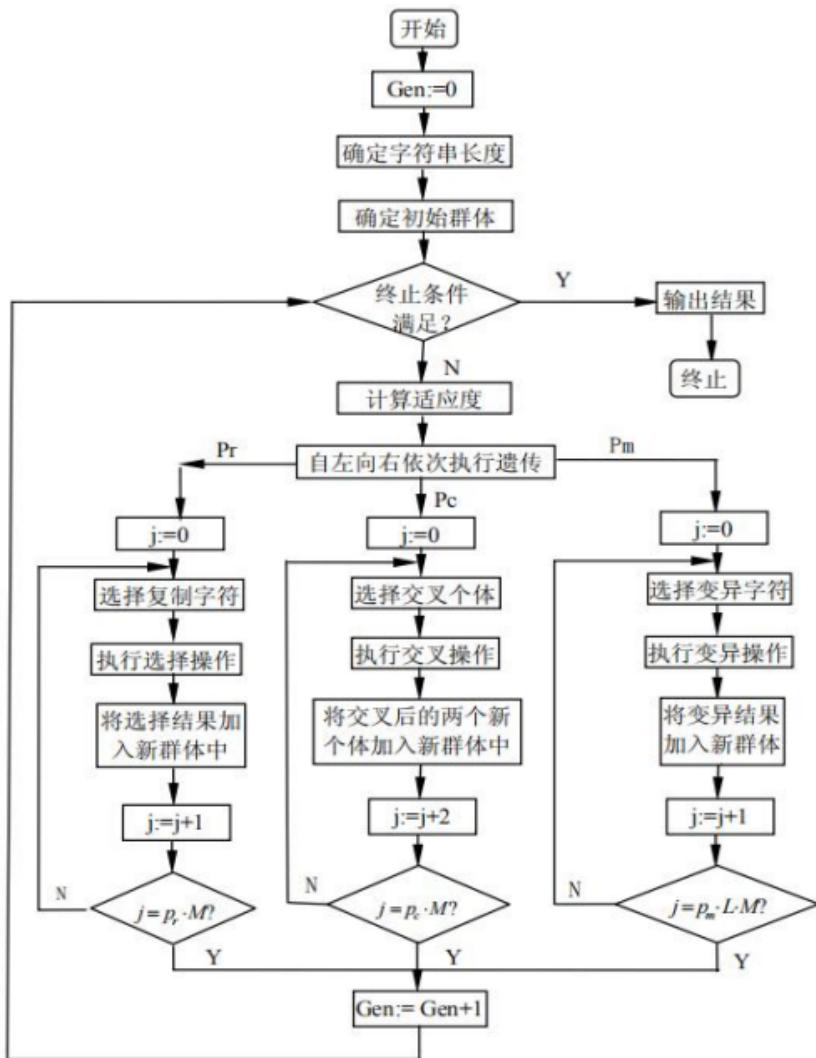
本节示例的计算机演算，得到本问题的最优解  $\max f(x) = 5.2786$ ，迭代过程中种群的变化和解的变化见图



# 流程

主要步骤包括：

- (1) 初始化群体；
- (2) 计算群体中每个个体的适应度值；
- (3) 由个体适应度值所决定的某个规则选择将进入下一代的个体；
- (4) 按概率  $P_c$  进行交叉操作；
- (5) 按概率  $P_m$  进行变异操作；
- (6) 没有满足某种停止条件，则转第 (2) 步，否则进入 (7)；
- (7) 输出种群中适应度值最优的染色体作为问题的满意解或最优解。



37

## 理论基础

### 模式理论

实质上是模式的运算

隐并行性

仅适用于二进制编码，无法证明收敛性，不能对算法参数提供实用指导

### 有限状态马尔科夫链模型

随机抽样，估计最优解在某区域存在的概率

研究收敛性

SGA不收敛到全局最优解

改进：EGA，选择最大适应度个体，可收敛到全局最优解

# 遗传规划

与GA类似，但是在种群初始化和编码方式上有区别

采用分层式编码，分支结点为运算符，叶子结点为终止符

## 进化策略

父代变异产生子代，然后进行选择

□ Rechenberg还提出著名的“1/5成功法则”：

- 对所有的变异，成功变异的比率应该是1/5。如果大于1/5，增加变异算子的方差；否则，减少其方差。

$$\sigma^{t+1} = \begin{cases} C_d \cdot \sigma^t & \text{若 } \varphi < 1/5 \\ \sigma^t & \text{若 } \varphi = 1/5 \\ C_i \cdot \sigma^t & \text{若 } \varphi > 1/5 \end{cases}$$

式中， $\varphi$ ——经历k次迭代后变异成功的次数与总变异次数之比，通常  $k > 10$ ；  
 $C_d$ ——小于1的系数，通常取0.82； $C_i$ ——大于1的系数，通常取1/0.82。

## 3.4.2 多元进化策略

### (1) 初始化。

随机生成 $k$ 个个体，构成第1代群体的父代。其中每个个体的不同是由个体的基因型的不同而决定的。

### (2) 变异。

每个父代个体将自己的基因经过变异后产生 $\lambda$ 个子代，此 $\lambda$ 个子代个体与其各自的父代个体在基因上有所不同。

### (3) 选择。

每个个体由于对环境的适应性不同而表现出不同的生命力或适应度，它由遗传基因决定。在 $\lambda$ 个个体中 $((\mu, \lambda)\text{-ES})$ 或者 $\mu + \lambda$ 个个体中 $((\mu + \lambda)\text{-ES})$ 选择生命力较强的 $\mu$ 个个体成为下一代的父辈个体。

**如果终止条件满足，则算法结束。否则回到步骤2）。**

变异为主要算子，交叉算子不必需。选择是确定性选择，精英选择，固定选择一定数目的最优个体到下一代

## 进化规划

考虑表现型

## 进化计算与遗传算法的总结

### 进化计算

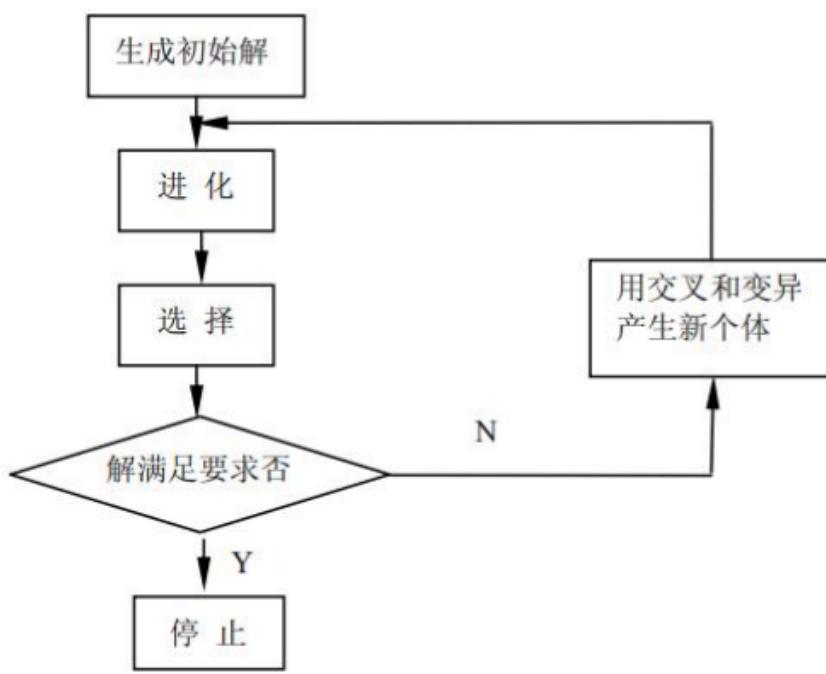
#### 定义

模拟生物进化过程中的自然选择机制和信息遗传规律的算法的总称

#### 特点

7点

## 流程



不要求所研究的问题是连续的、可导的，但是却可以很快地得出所要求的最优解

## 进化算法工作流程

## 应用

人工智能、人工生命、结构性优化、复杂问题优化

## 遗传算法

### 定义

遗传算法是一类借鉴生物界**自然选择**和**自然遗传机制**的**随机化搜索算法**

### 特点

交叉和变异

## 流程

上面有

## 原理

就是步骤那几个，编码、初始化、适应度、选择、交叉、变异、终止条件，简要说明一下各自是什么。

# 模拟退火算法

## 定义

模拟退火算法 (simulated annealing, SA) 是源于**固体退火原理**的一种**拟物智能算法**。

其出发点是受物理退火过程启发，模拟**固体加温、等温、冷却**等物理过程，结合**Metropolis抽样准则**进行科学计算的一种**启发式随机寻优算法**。

## 特点

- 1) 与局部搜索算法相比，模拟退火算法可望在较短时间内求得最优近似。
- 2) 模拟退火算法允许任意选取初始解和随机数序列，又能得出最优近似解。因此该求解优化算法的前期工作量大大减少。
- 3) 模拟退火算法能应用于多种组合优化问题，为一个问题编制的程序可以有效地应用于其他问题的求解

快、任意选取-前期工作量少、通用

## 原理

物质总是趋于最低的能态

物理退火过程：加温、等温、降温

# 相似性比较

## 相似性比较

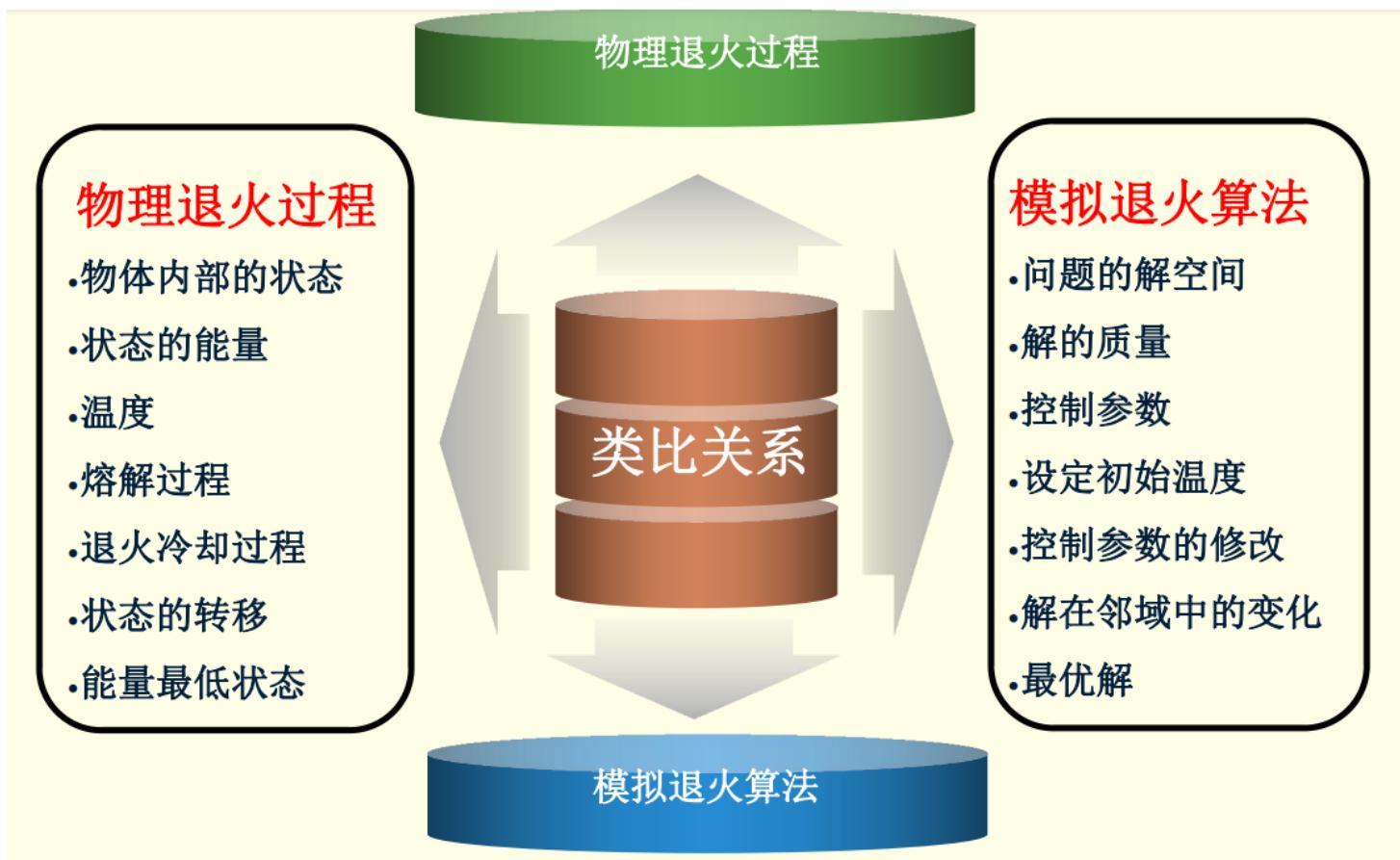
组合优化问题	金属物体
解	粒子状态
最优解	能量最低的状态
设定初温	熔解过程
Metropolis抽样过程	等温过程
控制参数的下降	冷却
目标函数	能量

## 设计与原理提出

模拟退火算法(SA) 就是这样一个将退火过程中**系统熵值**类比为**目标函数值F**，来模拟这个退火系统的算法。

Metropolis提出

# 与物理退火的比较



## 主要思想

在搜索区间（二维平面中）随机游走（即随机选择点），再以 Metropolis 抽样准则，使随机游走逐渐收敛于局部最优解。而**温度**即是 Metropolis 算法中的一个重要控制参数，可以认为这个参数的大小**控制了**随时过程向局部或全局最优解移动的**快慢**。

### 三大支柱

1. 接收准则和随机数产生器（Metropolis算法）
2. 冷却进度表
3. 新解产生器和邻域结构

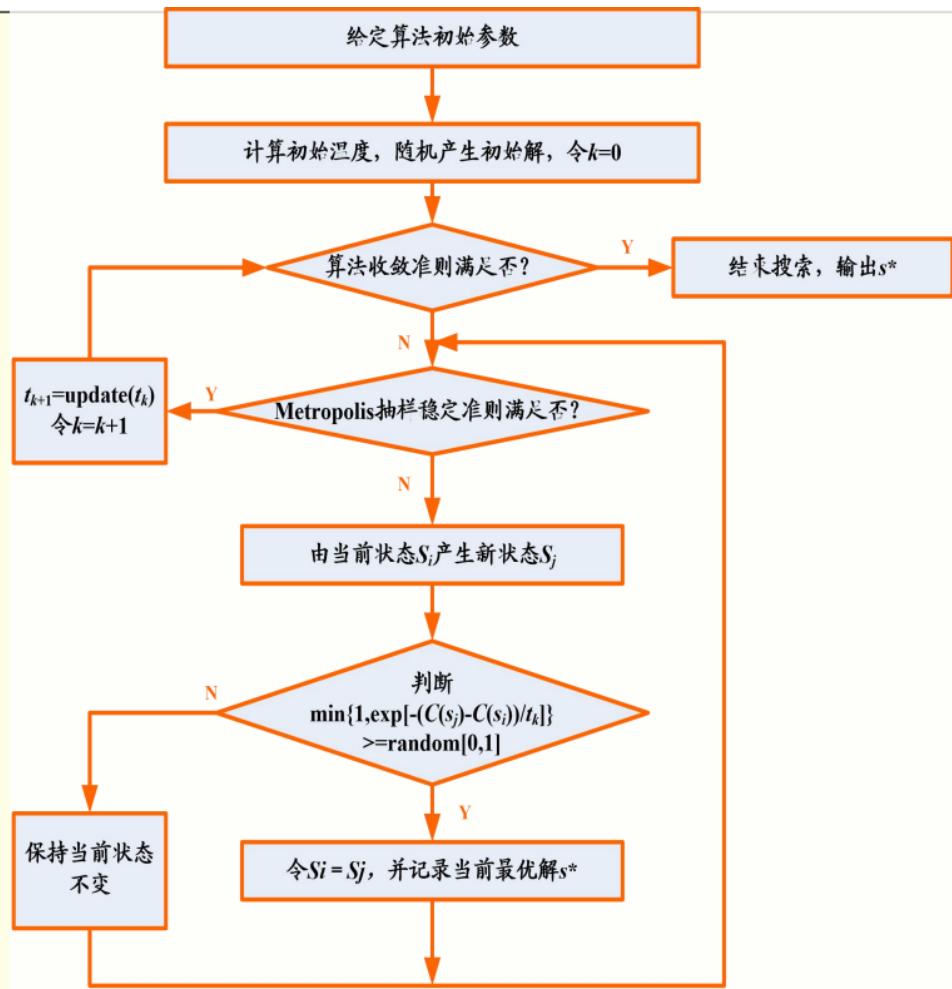
## 流程

### 模拟退火算法 (SA) 描述如下：

- (1) 初始化。任给一初始状态  $V_0, V_1 = V_0$ , 计算  $E(v_0)$ , 并为参数  $T$  设置初始温度值。
- (2) 产生一随机扰动  $\Delta v$ , 计算  $\Delta E = E(v_i + \Delta v) - E(v_i)$ 。
- (3) 若  $\Delta E < 0$ , 则转 (5), 否则在  $(0, 1)$  区间上产生一个均匀分布的随机数  $\xi$ 。
- (4) 若  $e^{-\Delta E / kT} \leq \xi$ , 则转 (2)。
- (5) 用  $v_i + \Delta v$  取代原来的  $v_i$ , 并令  $E = E + \Delta E$ 。
- (6) 在该  $T$  下, 检验系统是否稳定, 若不稳定则转 (2)
- (7) 以某一方式取  $T' < T$ , 令  $T = T'$ 。
- (8) 退火过程是否基本结束, 是就停止, 不是则转 (2)。

## 关键环节

- 1 初温、初始解
- 2 状态产生函数
- 3 状态接受函数
- 4 退温函数
- 5 抽样稳定准则
- 6 收敛准则



满足Metropolis接受准则，接受新解；满足Metropolis抽样稳定准则（达到平稳分布）更新状态，降低温度，否则就要一直尝试新解。

## Metropolis准则

以概率接受新解

$$e^{-\frac{\Delta E}{k_B T}}$$

## 实现思想

### 模拟退火算法的实现思想

- 用固体退火模拟组合优化问题，
- 将内能 E 模拟为目标函数值 f，
- 温度T 演化成控制参数 t，

固体退火-组合优化

内能E-函数f

温度T-控制参数t

## 冷却进度表

### 4.2.3 冷却进度表

- 我们称调整模拟退火法的一系列重要参数为冷却进度表。
- 一个冷却进度表应当规定下述参数：
  - ①控制参数t 的初值。
  - ②控制参数 t 的衰减函数。
  - ③马尔可夫链的长度  $L_k$  (即每一次随机游走过程，要迭代多少次，才能趋于一个准平衡分布，即一个局部收敛解位置)。
  - ④结束条件的选择。

一系列参数

温度、温度衰减、马尔科夫链长度、结束条件

#### 参数选择

##### 初始温度

大时易搜索到全局最优解，但算的慢

##### 衰减K

接近于1的常数

$$T^* = K$$

马尔科夫链长度 $L_k$

每次迭代产生的变换数

终止条件

## 新解的产生和邻域结构

- 新解的产生是由一个产生装置从当前解出发，在解空间中产生一个新解，以便于下面的计算和接受(这是算法中最耗时的工作)
  - 通常选择由当前解经过简单的变换即可产生新解的方法，如对构成解的全部或部分元素进行置换、互换或反演等。
- 新解的产生方法决定了当前解的邻域结构。

### 例：TSP问题

#### 例1 30城市TSP问题 ( $d^*=423.741$ by D B Fogel)

##### ● 参数设定

初始温度  $t_0=73150.8$ ,

截止温度  $t_f=0.01$ ;

退温系数  $\text{alpha}=0.90$ ;

内循环次数  $L=200*\text{CityNum}$ ;

这个L就是每个温度下要进行这么多的抽样，才会降低温度

虽不是最优解，只是近似全局最优，但效率高

# 分布估计算法 EDA

## 定义&本质

分布估计算法 (Estimation of Distribution Algorithms, EDA) 本质上是一种**基于概率模型的新型进化算法**，是自然计算的又一典型实现模式

分布估计算法又称为**基于概率模型的遗传算法 (Probabilistic Model Building Genetic Algorithms , PMBGA)**，是 20世纪90年代初提出的一种新型的启发式算法。它结合了**统计学习理论和遗传算法**的原理，通过构建概率模型、采样和更新概率模型等操作实现群体的进化。

分布估计算法的**思想起源于遗传算法**，但却有与遗传法截然不同的进化模式。从一定意义上说，遗传算法是在“微观”层面上对生物进化进行模拟，而分布估计算法则是在“宏观”的层面上来控制算法搜索。是一种全新的进化模型。

## 与遗传算法的关系

和遗传算法的算法结构相比，分布估计算法没有交叉算子和变异算子，取而代之的是**建立概率模型**和**采样**两大操作。

建立概率模型和采样是分布估计算法的核心步骤，也是 EDA 与 GA 的最大不同之处。

由于分布估计算法没有 交叉 和 变异 操作，因而通常不用**基因**来描述个体所包含的信息，取而代之的是**变量**(Variables)。

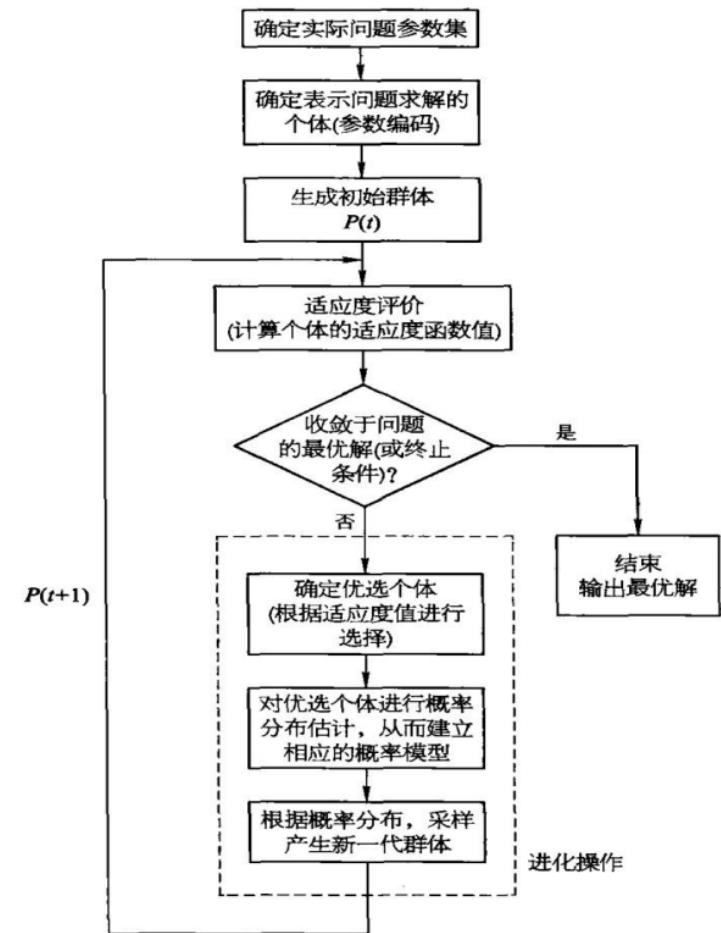
## 思想

分布估计算法的主要思想就是把**自然进化算法**和**构造性数学分析方法**相结合，以指导对问题空间的有效搜索。这种**将构造性概率模型引入进化算法的思想**形成了分布估计算法的理论依据。

分布估计算法的基本思想就是使用概率的方法描述和表示每一代群体

## 流程

- 步骤1：初始化群体，并对每一个个体进行估值（适应度值计算）；
- 步骤2：根据个体估值，按照一定的选择策略从群体中选择较优的个体（优选个体）；
- 步骤3：根据选择的个体估计概率分布，建立相应的概率模型；
- 步骤4：根据上一步估计得出的概率分布，采样产生新一代个体，并重新对每一个新个体进行适应度估值；
- 步骤5：如果某准则满足，则算法停止；否则，返回步骤2。



选择→估计分布→采样新一代个体

估计分布：建立概率模型，统计学习方法

采样：蒙特卡洛方法

## 经典的分布估计算法UMDA的执行步骤

Univariate marginal distribution algorithm(UMDA)

- ☒ 第一步：随机产生  $N$  个个体来组成一个**初始种群**，并评估初始种群中所有个体的适应度。
- ☒ 第二步：按**适应度从高到低**的顺序对种群进行排序，并从中**选出最优的  $Se$  个个体**( $Se \leq N$ )。
- ☒ 第三步：分析所选出的  $Se$  个个体所包含的信息，估计其**联合概率分布  $p(x)$** 。

$$p(x) = p(x | D^S) = \prod_{i=1}^n p(x_i)$$

- ☒ 其中， $n$  为解的维数， $p(x_i)$  为每维变量的边缘分布。

- ☒ 第四步：从构建的概率模型  $p(x)$  中采样，得到  $N$  个新样本，**构成新种群**。此时，若达到算法终止条件则结束，否则执行第二步。
- ☒ 注意：UMDA 在估计概率模型时，认为变量之间是**独立不相关的**。

## 高斯概率模型

建模连续变量

每一个变量满足一维高斯分布

第一步：随机产生初始种群并计算种群中所有个体的适应度，同时根据公式初始化概率模型。

$$\left\{ \begin{array}{l} \mu_j = \frac{\sum_{i=0}^{N-1} V[i][j]}{N} \\ \sigma_j = \sqrt{\frac{\sum_{i=0}^{N-1} (V[i][j] - \mu_j)^2}{N}} \end{array} \right.$$

其中： $V[i][j]$  为种群中个体  $i$  的变量  $x_j$  的值，  
 $N$  为种群的规模。

第二步：采用线性学习方式更新高斯分布的均值。设在第  $t$  代时， $x_j$  对应的高斯分布的均值为  $\mu_j^t$ ，则有：

$$\mu_j^{t+1} = (1 - \alpha) \cdot \mu_j^t + \alpha \cdot (x_j^{\text{best},1} + x_j^{\text{best},2} - x_j^{\text{worst}})$$

其中  $\alpha$  为学习因子， $x^{\text{best}1}$ 、 $x^{\text{best}2}$  和  $x_j^{\text{worst}}$  分别表示种群中最优个体、次优个体和最差个体所对应的  $x_j$  值。

- 第三步：采用线性学习方式更新高斯分布的方差。
- 设在第  $t$  代时， $x_j$  对应的高斯分布的方差为  $\sigma_j^t$ ，则有：

$$\sigma_j^{t+1} = (1 - \alpha) \cdot \sigma_j^t + \alpha \cdot \sqrt{\left( \sum_{k=0}^{K-1} (x_{jk} - \bar{x})^2 \right) / K}$$

- 其中  $x_{jk}$  为种群按适应度从大到小排名第( $k-1$ )的个体的  $x_j$  值， $\bar{x}$  为选出的  $K$  个较优种群的  $x_{jk}$  的均值，即：

$$\bar{x} = \left( \sum_{k=0}^{K-1} x_{jk} \right) / K$$

- 第四步：根据更新的高斯分布产生新样本。

均值更新：种群最优、次优和最差

方差更新：用选择后的种群计算方差然后增量式更新

## 应用

函数优化、组合优化、生物信息学

生物信息学：基因结构分析、DNA分类聚类、蛋白质结构预测和设计

## 群智能算法

### 特点

指无智能的主体通过合作表现出智能行为的特性，在没有集中控制且不提供全局模型的前提下，为寻找复杂的分布式问题求解方案提供了基础

### 优点

1. 灵活性：群体可以随时适应环境变化
2. 稳健性：即使个体失败，群体也能完成任务
3. 自我组织：活动即不受中央控制，也不受局部监管

# 蚁群算法

信息素（派洛蒙）

□ 作为蚁群算法的蚂蚁个体，其运动和通信的简单规则，只需包含以下几个主要方面：

- (1) 搜索范围---搜索参数半径
- (2) 局部环境---感知它周围的局部环境信息，并且该局部环境中的信息素是按一定速度消失的。
- (3) 觅食规则---只在其能感知的范围内进行信息探索和留存；总体上说，其搜寻的效率和正确性会通过其他蚂蚁的行为反馈加以调整。
- (4) 移动规则---都朝信息素最多的方向移动当周围没有信息素指引的时候，蚂蚁会按照自己原来运动的方向惯性地运动下去，并且，在运动的方向上有一个随机的小的扰动，以保留原来的运动记忆。如果发现有其已经经过的地点，则以较大概率进行避让。
- (5) 避障规则---将移动的方向上存在障碍物，则它会随机选择另一个方向。或者按照信息素的指引继续其觅食行为。
- (6) 通信规则---信息素

搜索范围&局部环境（全局和局部）

觅食-移动-避障-通信

## 蚁群算法解决TSP

- 每只蚂蚁都可以是具有下列特征的简单智能体：
  - 其选择城市的概率是城市之间的距离和连接支路上所包含的当前信息素余量的函数；
  - 为了强制蚂蚁进行合法的周游，直到一次周游完成时，才允许蚂蚁游走已访问过的城市（这可由禁忌表来加以控制）；
  - 当完成一次周游，它在每条访问过的支路  $(i, j)$  上留下一种叫信息素的物质；
- **解决TSP问题**
  - 在算法的初始时刻，将  $m$  只蚂蚁随机放到  $n$  座城市；
  - 将每只蚂蚁  $k$  的禁忌表  $tabu_k(s)$  的第一个元素  $tabu_k(1)$  设置为它当前所在城市；
  - 设各路径上的信息素  $\tau_{ij}(0)=C$  ( $C$  为一较小的常数)；

## ■ 解决TSP问题

每只蚂蚁根据路径上的**信息素和启发式信息**（两城市间距离）独立地选择下一座城市：

在时刻 $t$ , 蚂蚁 $k$ 从城市 $i$ 转移到城市 $j$ 的概率为

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}(t)]^\beta}{\sum_{s \in J_k(i)} [\tau_{is}(t)]^\alpha [\eta_{is}(t)]^\beta}, & j \in J_k(i) \\ 0, & j \notin J_k(i) \end{cases}$$

下一步允许的城市的集合

$\alpha$ 、 $\beta$ 分别表示信息素和启发式因子的相对重要程度。

$$J_k(i) = \{1, 2, \dots, n\} - tabu_k, \quad \eta_{ij} = 1/d_{ij}$$

自然计算 (natural computation)

16

不能访问已经访问过的城市，当j在禁忌表中时访问概率为0  
概率为信息素和启发式的加权，归一化

## ■ 解决TSP问题

当所有蚂蚁完成一次周游后，各路径上的信息素将进行更新：

$$\tau_{ij}(t+n) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij}$$

$$\Delta\tau_{ij} = \sum_{k=1}^m \Delta\tau_{ij}^k, \quad \Delta\tau_{ij}^k = \begin{cases} \frac{Q}{L_k}, & \text{若蚂蚁 } k \text{ 在本次周游中经过边 } ij \\ 0, & \text{否则} \end{cases}$$

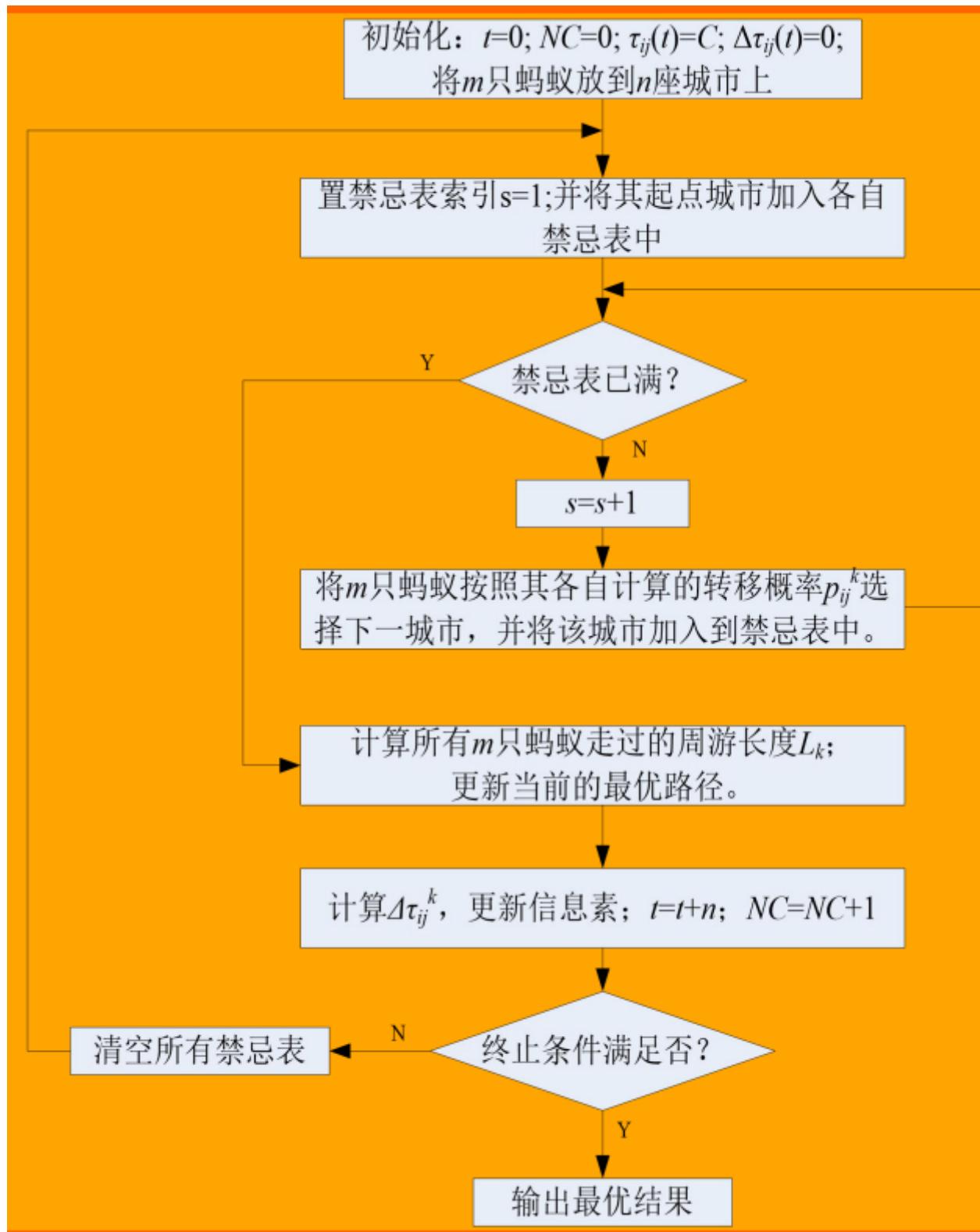
其中， $\rho$  ( $0 < \rho < 1$ ) 表示路径上信息素的蒸发系数，  
 $Q$  为正常数， $L_k$  表示第  $k$  只蚂蚁在本次周游中所走  
过路径的长度。

自然计算 (natural computation)

17

每只蚂蚁在一条路径上留下信息素  $\frac{Q}{L_k}$ ，所有蚂蚁的信息素总和加上蒸发后的原有信息素得到这条路径上的信息素

## 流程



## 优缺点分析

### □ 优点

- (1) 通用性
- (2) 鲁棒性
- (3) 群体性
- (4) 并行性

### □ 不足之处

- (1) 需要较长的计算时间，容易出现停滞现象。
- (2) 候选解并非都是最优解，这样计算信息素的增量会导致错误的引导信息，从而造成大量的无效搜索，使系统出现停滞现象。
- (3) 信息素均匀分配策略，即对已搜索路径中的所有路段采用同样的信息素增量，与路段的重要性无关。

计算时间长、停滞、无效搜索、信息素分配没有考虑路段重要性

### ■ 优点

较强的鲁棒性；  
分布式计算；  
易于与其他方法结合。

### ■ 缺点

搜索时间较长；  
容易出现停滞现象。

## 参数影响分析

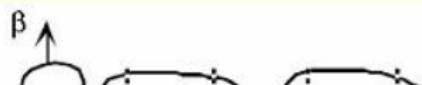
### 参数 $\alpha$ 、 $\beta$ 对算法性能的影响

停滞现象（Stagnation behavior）：所有蚂蚁都选择相同的路径，即系统不再搜索更好的解。

原因在于较好路径上的信息素远大于其他边上的，从而使所有蚂蚁都选择该路径。

### 参数 $\alpha$ 、 $\beta$ 对算法性能的影响

$\alpha$ 取较大值时，意味着在选择路径时，路径上的信息素非常重要；当 $\alpha$ 取较小值时，变成随机的贪婪算法。



#### ■ 蚂蚁数 $m$ 对算法的影响

$m \approx n$ 时，*ant-cycle*可以在最少迭代次数内找到最优解。

## 蚂蚁的初始分布

两种情况实验：

- (1) 所有蚂蚁初始时刻放在同一城市；
- (2) 蚂蚁分布在不同城市中。

第(2)中情况可获得较高性能。

# 粒子群算法

## 定义

源于对鸟群捕食行为的研究，是基于迭代的方法

## 原理

初始化为一群随机粒子，通过迭代找到最优。每次迭代中，粒子通过跟踪“个体极值 (pbest)”和“全局极值(gbest)”来更新自己的位置。

### 粒子速度和位置的更新

假设在 $D$ 维搜索空间中，有 $m$ 个粒子；

其中第 $i$ 个粒子的位置为矢量  $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{iD})$

其飞翔速度也是一个矢量，记为  $\vec{v}_i = (v_{i1}, v_{i2}, \dots, v_{iD})$

第 $i$ 个粒子搜索到的最优位置为  $\vec{p}_i = (p_{i1}, p_{i2}, \dots, p_{iD})$

整个粒子群搜索到的最优位置为  $\vec{p}_{gbest} = (p_{gbest1}, p_{gbest2}, \dots, p_{gbestD})$

第 $i$ 个粒子的位置和速度更新为：

$$v_{id}^{k+1} = w v_{id}^k + c_1 rand() (p_{id} - x_{id}^k) + c_2 rand() (p_{gbest} - x_{id}^k)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad i = 1, 2, \dots, m; \quad d = 1, 2, \dots, D$$

速度是原速度（惯性）、个体极值方向、全局极值方向的加权

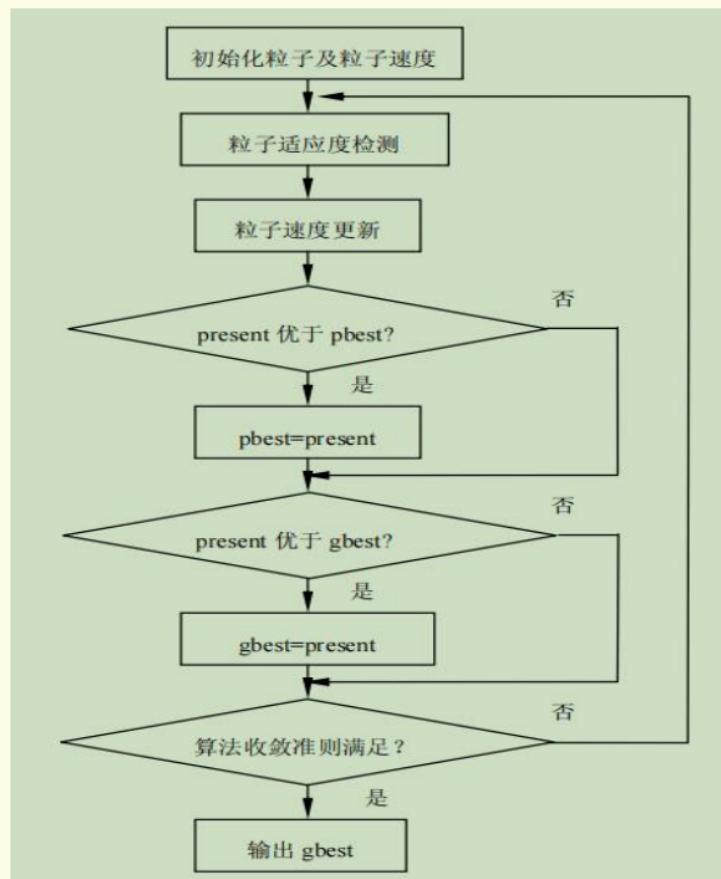
w: 惯性权重 c: 加速因子

惯性、认知、社会

## 流程

# 粒子群优化算法的框架

- (1) 初始化粒子群，包括群体规模，每个粒子的位置和速度；
- (2) 计算每个粒子的适应值；
- (3) 对每个粒子，用它的适应值和个体极值  $pbest$  比较，如果较好，则替换  $pbest$ ；
- (4) 对每个粒子，用它的适应值和全局极值  $gbest$  比较，如果较好，则替换  $gbest$ ；
- (5) 根据公式更新粒子的速度和位置；
- (6) 如果满足结束条件（误差足够好或者到达最大循环次数）退出，否则回到 (2)



## 参数分析

### 惯性权重w

w利于搜索空间的扩展，大w利于跳出局部最优，小w利于算法收敛

### 加速因子c

小的c使粒子在目标区域附近徘徊，大的c使粒子突然冲向或越过目标

## 加速常数 $c_1$ 和 $c_2$

将 $c_1$ 和 $c_2$ 统一为一个控制参数， $\varphi = c_1 + c_2$

如果 $\varphi$ 很小，微粒群运动轨迹将非常缓慢；

如果 $\varphi$ 很大，则微粒位置变化非常快；

实验表明，当 $\varphi=4.1$ （通常 $c_1=2.0$ ,  $c_2=2.0$ ）时，具有很好的收敛效果。

- **粒子数**

一般取20 ~ 40，对较难或特定类别的问题可以取100 ~ 200。

- **最大速度 $v_{max}$**

决定粒子在一个循环中最大的移动距离，通常设定为粒子的范围宽度。

- **终止条件**

最大循环数以及最小错误要求。

## 与遗传算法比较

### ■ 共性

- (1) 都属于仿生算法;
- (2) 都属于全局优化方法;
- (3) 都属于随机搜索算法;
- (4) 都隐含并行性;
- (5) 根据个体的适配信息进行搜索，因此不受函数约束条件的限制，如连续性、可导性等；
- (6) 对高维复杂问题，往往会遇到早熟收敛和收敛性能差的缺点，都无法保证收敛到最优点。

### 差异

- (1) PSO有记忆，所有粒子都保存较优解的知识，而GA，以前的知识随着种群的改变被改变；
- (2) PSO中的粒子是一种单向共享信息机制。而GA中的染色体之间相互共享信息，使得整个种群都向最优区域移动；
- (3) GA需要编码和遗传操作，而PSO没有交叉和变异操作，粒子只是通过内部速度进行更新，因此原理更简单、参数更少、实现更容易。

共性：

1. 仿生
2. 全局优化
3. 随机搜索

- 4. 并行
- 5. 根据适配信息进行搜索
- 6. 可能无法收敛

其实就是遗传的几个特点

差异：

- 1. PSO有记忆，记忆保存在粒子中，但GA保存在种群中，随种群改变
- 2. PSO单向共享，GA相互共享
- 3. PSO没有编码和遗传，原理更简单、参数更少、更易实现

## 应用

神经网络训练、参数优化、组合优化

## 群智能算法总结

### ■ 共同特点

基于概率计算的随机搜索进化算法，在结构、研究内容、方法以及步骤上有较大的相似性；

### ■ 存在的问题

(1) 数学理论基础相对薄弱；

(2) 参数设置没有确切的理论依据，对具体问题和应用环境的依赖性大；

- 存在的问题
  - (3) 比较性研究不足，缺乏用于性能评估的标准测试集；
  - (4) 不具备绝对的可信性，存在应用风险；
- 进一步的工作
  - (1) 进一步研究真实群居动物的行为特征；
  - (2) 进一步研究算法的收敛性；

缺乏理论依据，研究不足，不具备绝对可信性，存在应用风险。

## 人工免疫计算

### 生物免疫的两种类型

特异性免疫、非特异性免疫

# 免疫生物学的基本概念

## ① 抗原

是指能够刺激和诱导机体的免疫系统使其产生免疫应答，并能与相应的免疫应答产物在体内或体外发生特异性反应的物质。

## ② 抗体

是指免疫系统受抗原刺激后，免疫细胞转化为浆细胞并产生能与抗原发生特异性结合的免疫球蛋白，该免疫球蛋白即为抗体。

# 免疫系统的主要功能

## ① 免疫防御

即机体防御病原微生物的感染；

## ② 免疫（自身）稳定

即机体通过免疫功能经常消除那些损伤和衰老的细胞以维持机体的生理平衡；

## ③ 免疫监视

即机体通过免疫功能防止或消除体内细胞在新陈代谢过程中发生突变的和异常的细胞。

## 免疫系统

两个部分：固有免疫系统和自适应免疫系统

自适应免疫又分体液免疫和细胞免疫

## 免疫应答

初次应答和二次应答

学习和记忆抗原

## 免疫+遗传

### 抗原

所有可能错误的基因，即非最佳个体的基因。

### 疫苗

根据进化环境或待求问题的先验知识，所得到的最佳个体基因的估计。

### 抗体

根据疫苗修正某个个体的基因所得到的新个体。

## 免疫算子有两种类型：



**全免疫**

**目标免疫**

**非特异性免疫**

**特异性免疫**

即：群体中的每个个体在进化算子作用后，对其每一环节都进行一次免疫操作的免疫类型；

即：在进行了进化操作后，经过一定的判断，个体仅在作用点处发生免疫反应的一种类型。

免疫算子由**接种疫苗**和**免疫选择**两个操作完成，(免疫选择又包括免疫检测和退火选择)

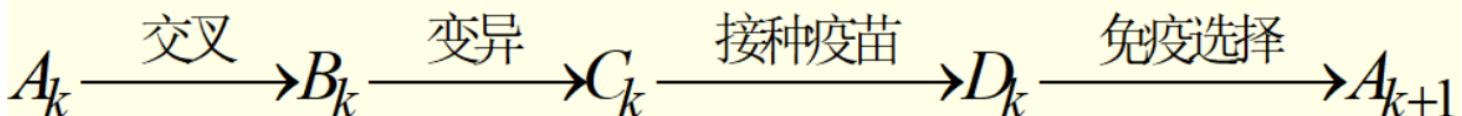
# 免疫算法

## 流程

- 随机产生初始父代种群  $A_1$ , 根据先验知识抽取疫苗;
- 若当前群体中包含最佳个体, 则算法停止运行并输出结果; 否则, 继续;
- 对当前第  $k$  代父本种群  $A_k$  进行交叉操作, 得到种群  $B_k$ ;
- 对  $B_k$  进行变异操作, 得到种群  $C_k$ ;
- 对  $C_k$  进行接种疫苗操作, 得到种群  $D_k$ ;
- 对  $D_k$  进行免疫选择操作, 得到新一代父本  $A_{k+1}$ , 转至第二步。

相较于遗传算法多了抽取疫苗和后续对变异完的种群进行接种疫苗和免疫选择操作 (免疫选择代替了选择)

状态转移过程示意图:



定 义: 如果对于任意的初始分布均有

$$\lim_{k \rightarrow \infty} \sum_{s_i \cap S^* \neq \emptyset} P\{A_k^i\} = 1$$

则称算法收敛。

定 理: 免疫算法是收敛的。

## 免疫规划

免疫+进化规划

## 免疫策略

免疫+遗传策略

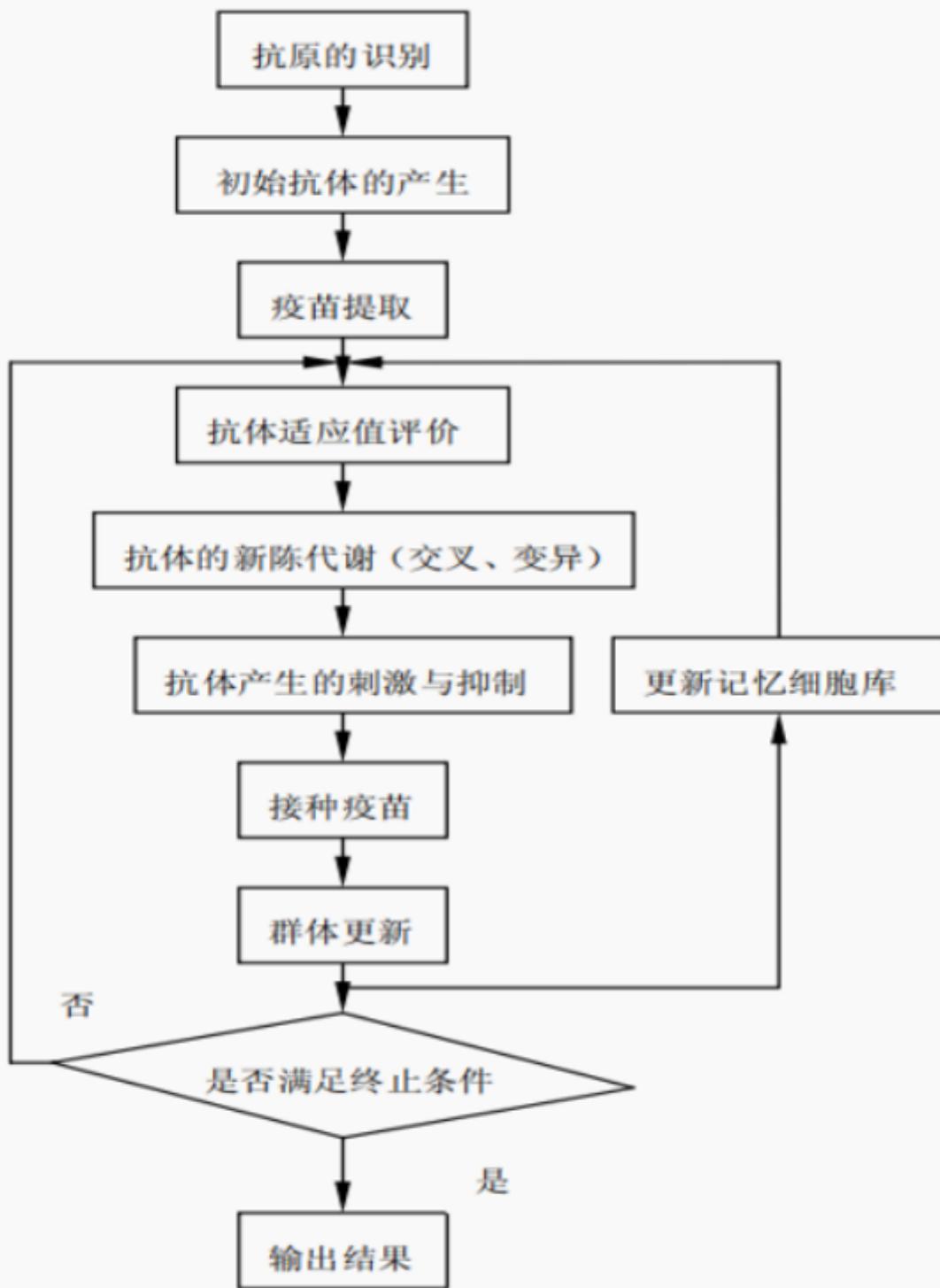
### 机理

# 免疫算子的机理

定 理: 

在免疫选择作用下，若疫苗使抗体适应度得到提高，且高于当前群体的平均适应度，则疫苗所对应的模式将在群体中呈指数级扩散；否则，它将被遏制或呈指数级衰减。

# 免疫算法流程



免疫算法流程图

多了一个疫苗提取和接种疫苗

# 混沌

## 定义

### 混沌的定义

□ 设 $V$ 为一个集合,  $f: V \rightarrow V$ , 称为在 $V$ 上是混沌的, 如果:

- $f$  对初始条件的敏感依赖性;
- $f$  是拓扑传递的;
- 周期点在 $V$ 中是稠密的;

混沌 —— 指确定性系统产生的一种对初始条件具有敏感依赖性的回复性非周期运动。

对初值敏感

### The general concept of chaos

- ⌚ 如果一个接近实际而没有系统内在随机性的模型仍然具有貌似随机的行为, 就可以称这个真实的物理系统是混沌的.
- ⌚ 例如: 钟摆的摆动。如果科学家忽略任何外来的随机扰动和气流, 该模型应是线性的、周期性的系统。

广义随机性和缺乏确定性是一个意思（广义随机：不完全的随机，比如洗牌，依赖上一次洗牌的结果）

## 混沌

- 混沌可以说他是确定性的行为；
- 或者，若考虑他出现在稍微有点随机性的实际系统中，也可以说他是近似与确定性的，然而却不是看起来像确定性的。
- 在某些动力系统中，两个几乎一致的状态经过充分长的时间后会变得毫无一致性。

每个状态可以计算，但是微小扰动就会看起来随机

敏感地依赖于初始条件

混沌可以理解为貌似随机的确定性。

貌似随机的确定性

## 特性

### □ 基于混沌现象的独特性质

- ◆ 遍历性
- ◆ 随机性
- ◆ 规律性
- ◆ 初值敏感性

## 混沌优化算法

利用混沌变量进行搜索，容易跳出局部最优

### 算法模型

#### □ 搜索过程分为两个阶段：

1、基于确定性迭代式产生的遍历性轨道对整个解空间进行考察，确定最佳状态解；

2、以第一阶段结果为中心，通过附加小幅度扰动进行局部区域细搜索，直至算法终止准则满足。

选择用于载波的混沌变量，使用Logistic映射

$$x_{n+1} = ax_n(1 - x_n) \quad (8.1)$$

其中  $a$  是控制参量，取  $a = 4$ 。

设  $0 \leq x_n \leq 1$ ,  $n = 0, 1, 2 \dots$  可知系统完全处于混沌状态

- 混沌优化方法直接采用混沌变量进行搜索，搜索过程按混沌运动自身的规律进行，比较模拟退火法更容易跳出局部最优点，搜索效率高。

## 分形几何

Koch曲线，有界但无限，自相似性

### 维数

- 从上面的等式中可以看到,  $r$  的幂次实际就是该几何体的空间维数,可以表示为:

$$n \cdot r^D = 1$$

- 对上式两边取对数得:

$$D = -\frac{\ln n}{\ln x}$$

- 显然,  $D$ 具有维数的概念.

- 在第n步时,其等长折线段总数为 $4^n$ ,每段的长度为 $\left(\frac{1}{3}\right)^n$
- 则Koch曲线的维数为:

$$D = -\frac{\ln 4^n}{\ln \left(\frac{1}{3}\right)^n} = \frac{\ln 4}{\ln 3} = 1.26186$$

- 英国海岸线的维数为 $D=1.25$  (Mandelbrot)

n是分成的小块的数目, x是每个小块的长度  
维数描述了复杂性、无序性、填充空间的能力

### 随机分形搜索算法

## 算法模型

- 将微粒类比为寻优个体, 扩散过程类比为优化过程。
- 个体的位置  $\longleftrightarrow$  待优化问题的解
- 优化搜索  $\longleftrightarrow$  高斯游走模型在个体当前位置扩散
- 更新操作  $\longleftrightarrow$  每个个体基于其他个体位置进行更新

## 流程

- Step 1 设置算法参数，并进行种群初始化；  
Step 2 计算种群中每个个体的适应度值，并找到当前最佳点BP；  
Step 3 执行扩散过程，根据所选择的高斯游走方式生成个体的新位置；  
Step 4 根据式(8.5)和式(8.6)进行群体第一次更新；  
Step 5 根据式(8.5)和式(8.7)进行群体第二次更新；  
Step 6 判断是否满足算法停止条件，若满足，则算法结束，并输出结果；否则，执行Step 2.

第一次更新：

$$\text{定义寻优个体的优化性能级别 } P_{a_i} = \frac{\text{rank}(P_i)}{N} \quad (8.5)$$

$$\text{更新方法: } P'_i(j) = \begin{cases} P_r(j) - \varepsilon_3 \cdot (P_l(j) - P_i(j)) & \text{if } P_{a_i} < \varepsilon_4 \\ P_i(j) & \text{否则} \end{cases} \quad (8.6)$$

$$\text{第二次更新: } P''_i = \begin{cases} P'_i - \hat{\varepsilon} \cdot (P'_l - BP) & \text{if } \varepsilon_5 < 0.5 \\ P'_i - \hat{\varepsilon} \cdot (P'_l - P'_r) & \text{否则} \end{cases} \quad (8.7)$$