演化计算理论研究与自适应演化算法

秦霄羽

xxq896@cs.bham.ac.uk

英国伯明翰大学

计算机科学学院

2021 年 04 月 21 日

目录

秦霄羽[1]，英国伯明翰大学计算机科学专业在读博士生，主要研究方向为演化计算理论（*Theory of Evolutionary Computation*）、自然启发式算法（*Nature-inspired Algorithms*）。博士期间研究课题为演化算法中的自适应机制（*Self-adaptation in Evolutionary Algorithms*）。

- **教育经历**：

山西大学， *2014.09 - 2018.07*
计算机科学与技术工学学士

英国拉夫堡大学， *2017.10 - 2018.07*
计算机科学专业访问学生

英国伯明翰大学， *2018.09 - 2019.09*
高级计算机科学理学硕士

英国伯明翰大学， *2020.02 - 2024.02*
计算机科学哲学博士（在读）

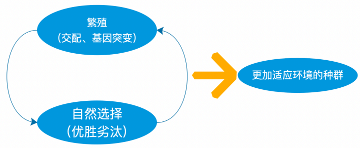
[1] 个人主页：cs.bham.ac.uk/∼xxq896/

伯明翰大学

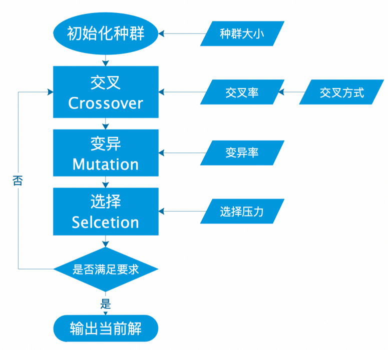
进化论与演化计算

生物进化论：

http://www.environmentalgraffiti.com/featured/amazing-insect-camouflage/14128



演化计算：



图：演化算法框架

一个有趣的应用

图: 使用演化算法设计软体机器人 [2]

[2] 来源： https://youtu.be/z9ptOeByLA4

演化算法的优势

- 高效性
  - 启发式：在合理时间内可得到高质量的解
  - 可并行：基于种群
- 实用性
  - 可多目标优化
  - 可全局搜索：容易跳出局部最优
- 鲁棒性
  - 可在不确定环境下优化

演化计算应用

解决组合优化问题：

- TSP 问题
- 投资组合问题
- 航班调度问题
- …

解决连续数值优化问题：

- 工业设计问题
- 进化神经网络
- 强化学习
- …

演化算法的分类

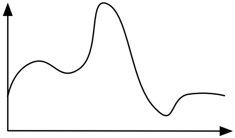
遗传算法

Genetic Algorithm

$$4321 \Rightarrow 1100010010001$$

进化策略

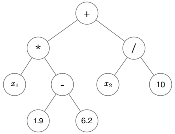
Evolutionary Strategy



进化编程

Evolutionary Programming

$$x_1 \times (1.9 - 6.2) + x_2 \div 10$$



演化计算

代表性学术会议、期刊及资源

学术会议:

- ACM Genetic and Evolutionary Computation Conference (GECCO)
- ACM/SIGEVO Workshop on Foundations of Genetic Algorithms (FOGA)
- Parallel Problem Solving from Nature (PPSN)

学术期刊:

- Evolutionary Computation
- IEEE Transactions on Evolutionary Computation (TEC)

算法库:

- Nevergrad - A gradient-free optimisation platform

演化计算理论研究内容

- 收敛性
- 期望运行时间上界
- 期望运行时间下界
- 在合理时间内得到解的概率的下界
- 噪声、动态环境等不确定因素的影响
- ...

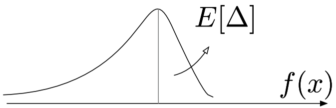
为什么要研究理论？

- 绝对正确
  - 任何人可重复、可检验
- 实验无法验证
  - 可得到问题规模 n 很大时的结果，如 TSP 问题中节点数 n 等
- 可以从原理上理解算法运行
- 可以指导设计和应用算法
- …

举个例子

$(1+1)$ EA：

<u>初始化：</u> 按位均匀随机生成一个个体 $x \in \{0,1\}^n$，设置变异率 $p = 1/n$
<u>优化：</u> 当 x 与 x^* 不相等时，进行
1. <u>变异：</u> 通过逐位按照概率 p 翻转 x 创建一个新个体 y
2. <u>选择：</u> 如果$f(y) \geq f(x)$，则用 y 替换 x



$$E[\Delta] \geq \frac{1}{n}\left(1-\frac{1}{n}\right)^i \geq \frac{1}{n}\left(1-\frac{1}{n}\right)^n > \frac{1}{en}$$

LeadingOnes:

$$\mathrm{LO}(x) := \sum_{i=1}^{n} \prod_{j=1}^{i} x_j$$

$$x : \{\overbrace{11\cdots1}^{i\text{个}}0**\cdots*\} \Rightarrow \mathrm{LO}(x) = i$$

所以，

$$E[T] = n/E[\Delta] = O(n^2)$$

目前面临的问题

- 通常只能分析简单的问题和简单的算法
- 结果不易得
- 缺乏更精确的结果，$O(n^2) = cn^2$，但经常不知道具体的 c
- $\dots$

博士论文题目：演化算法中的自适应机制

PhD Thesis Topic: Self-adaptation in Evolutionary Algorithms

一个典型的遗传算法

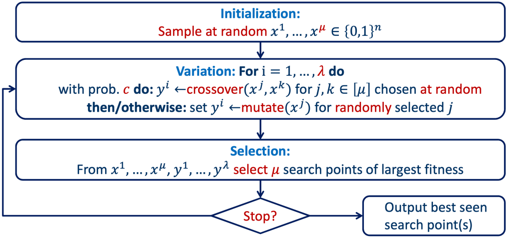


图: 一个典型的遗传算法[3]

- 有很多参数，如变异率、交叉率、种群数量、选择压力等
- 算法的性能很大程度上取决于参数的设置
- 随着优化的进行，最优参数也会发生变化

[3] Tutorial: Carla Doerr, Dynamic Parameter Choices in Evolutionary Computation, GECCO'19

参数的重要性（理论结果）

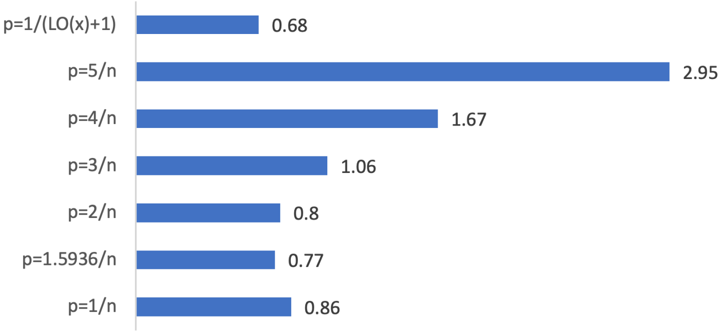


图: 不同变异率 p 下，$(1+1)$ EA 在 $\mathrm{LO}(x)$ 上的期望运行时间 $E[T]/n^2$ [4]

[4] Tutorial: Carla Doerr, Dynamic Parameter Choices in Evolutionary Computation, GECCO'19

优化参数的方法

- 尝试寻找全局最优参数
  - 多运行几次找出最优参数
  - 使用“超优化算法”优化参数，如 AutoML 算法等
- 在优化过程中动态地调整参数
  - <u>确定性算法：</u> 基于运行时间、运行状态等改变参数，如模拟退火算法等
  - <u>适应算法：</u> 制定一些规则改变参数，如 $1/5$-th 规则等
  - <u>自适应算法</u>：将参数也“编码”进个体，同个体一同“进化”

自适应演化算法

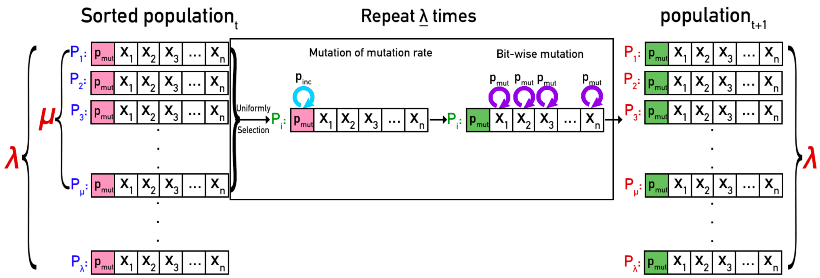


图: (μ, λ) 自适应演化算法

图: 自适应演化算法 [5]

$$\textsc{PeakedLO}(x) := \begin{cases} m & \text{if } x = 0^n \\ \sum_{i=1}^{n} \prod_{j=1}^{i} x_j & \text{otherwise} \end{cases}$$

- 当使用小变异率、大变异率或者均匀随机选择两中变异率时，期望运行时间为 $E[T] = e^{\Omega(n)}$
- 当使用自适应机制调整变异率时，期望运行时间为 $E[T] = O(n^2)$

[5] 来源：https://www.cs.bham.ac.uk/∼lehrepk/selfadapt/

有待研究的问题

- 是否能帮助算法跳出局部最优解
- 在噪音、动态环境等不确定环境下的性能
- 其他参数的自适应，如交叉率、种群数量、选择压力、交叉方法、选择方法等
- ...

谢谢大家！

欢迎交流