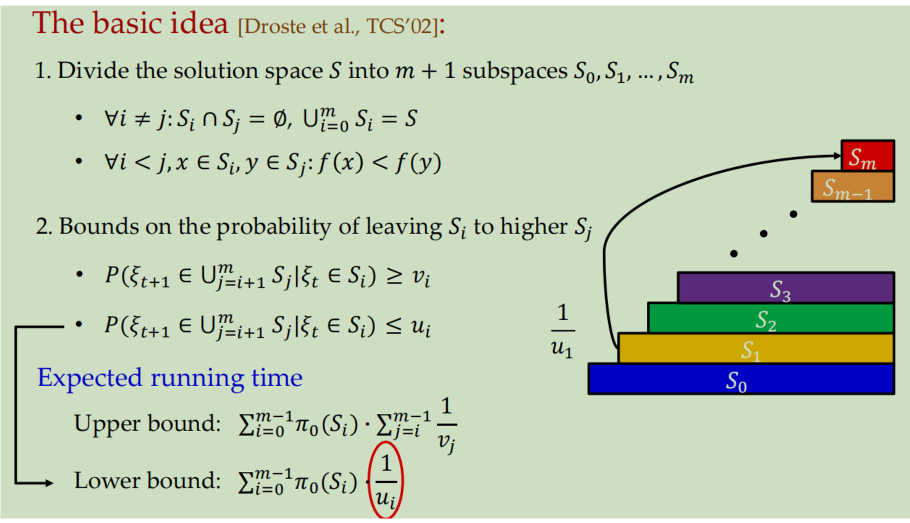
1 Problem 1: 基础概念 (5)

试理解适应层分析（fitness level method）、加性漂移分析（additive drift analysis）、乘性漂移分析（multiplicative drift analysis）三种理论分析工具的核心思想，并简述如何应用这三种分析方法。

答：适应层分析（fitness level method）的核心思想为：将解空间 S 划分为 m+1 个子空间 $S_0, S_i, ..., S_m$，并计算从 S_i 到更高层的 S_j 的概率的上界和下界，由此我们可以根据这些上下界来计算得到整个问题的期望运行时间的上下界。应用时先要按照问题特征对解空间进行划分，然后再根据问题的约束计算概率上下界，由此可以计算出期望运行时间的上下界。
加性漂移分析（additive drift analysis）的核心思想为：设计一个距离函数 $V(x)$ 来度量从状态空间中的一个状态 x 到目标状态空间的距离，然后计算期望单步漂移距离的上下界（与当前状态到目标状态空间的距离无关），由此可以计算出期望运行时间的上下界。应用时先要根据问题特征设计出合适的距离函数，然后再根据问题的约束计算单步距离上下界，由此可以计算出期望运行时间的上下界。
乘性漂移分析（multiplicative drift analysis）的核心思想在加性漂移分析的基础上有所改动：乘性漂移分析中期望单步漂移距离的上下界与当前状态到目标状态空间的距离有关。应用与加性漂移分析相似，除了在单步距离上下界和期望运行时间上下界的计算公式上有不同。

1、fitness level method适应层分析法



- 需要的时间，是概率的倒数。概率下届，能对应时间的上界

例题：(1+1)-EA for OneMax

(1+1)-EA:

Given a pseudo-Boolean function f:

1. $\boldsymbol{x} :=$ randomly selected from $\{0,1\}^n$.
2. Repeat until some termination criterion is met
3. $\quad \boldsymbol{x}' :=$ flip each bit of $\boldsymbol{x}$ with probability $1/n$.
4. $\quad$ if $f(\boldsymbol{x}') \geq f(\boldsymbol{x})$
5. $\quad\quad \boldsymbol{x} = \boldsymbol{x}'$.

Bit-wise mutation

OneMax:

$$max_{x\in\{0,1\}^n} \sum_{i=1}^{n} x_i$$

Count the number of 1-bits

Theorem. [Droste et al., TCS'02] The expected running time of the (1+1)-EA solving the OneMax problem is $O(n\log n)$.

Main idea:

the number of 1-bits

- Divide the solution space $\{0,1\}^n$ into $S_0, S_1, \dots, S_n$ with $S_i = \{x \in \{0,1\}^n \mid |x| = i\}$
- The probability of jumping to higher S_j from S_i is lower bounded by

$$P(\xi_{t+1} \in \cup_{j=i+1}^{m} S_j \mid \xi_t \in S_i) \geq \boxed{\frac{n-i}{n} \cdot (1-\frac{1}{n})^{n-1}}$$

keep the other bits unchanged

flip one of the $n-i$ 0-bits

$$v_j = \frac{n-j}{n} \cdot (1-\frac{1}{n})^{n-1}$$

Upper bound on the expected running time:

$$\sum_{i=0}^{n-1} \pi_0(S_i) \cdot \sum_{j=i}^{n-1} \frac{1}{v_j} \leq \sum_{j=0}^{n-1} \frac{1}{v_j}$$

$$= \sum_{j=0}^{n-1} \frac{n}{n-j} \frac{1}{(1-\frac{1}{n})^{n-1}} \leq en\sum_{j=1}^{n} \frac{1}{j} \in O(n\log n)$$

例题：(1+1)-EA for BinVal等价于LeadingOnes

定义 3 (BinVal). 一个规模为 n 的 *BinVal* 问题旨在找到一个 n 位的 *01* 串，以最大化

$$f(\boldsymbol{s}) = \sum_{i=1}^{n} 2^{n-i} s_i,$$

这里 s_i 指 $\boldsymbol{s} \in \{0,1\}^n$ 的第 i 位。

5 Problem 5: 求解 BinVal 问题 (30)

任选分析工具，求解 (1+1)-EA 算法找到 BinVal 问题最优解的期望运行时间的上界。

答：BinVal 问题等价于使二进制 01 串表示的十进制值最大的问题，而由二进制的性质可得，同等位数的情况下，高位为 1 的数大于所有高位为 0 的数，如 1000 大于最高位为 0 的所有四位二进制数。故和 LeadingOnes 问题一样，从左往右数连续的 1 的个数（一旦遇到 0 就不用考虑右边的 1 了）越多，适应度越高。

故同样使用适应层分析法，将解空间 $\{0,1\}^n$ 按从左往右数连续的 1 的个数（一旦遇到 0 就不用考虑右边的 1 了）划分为 n+1 个子空间 $S_0, S_i, ..., S_n$，其中 $S_i = \{x \in \{0,1\}^n \mid f(x) = i\}$，$f(x)$ 即 x 从左往右数连续的 1 的个数（一旦遇到 0 就不用考虑右边的 1 了）。

然后计算从 S_i 到更高层的 S_j 的概率的下界（从左往右数第一个为 0 的位翻转，前 i 个 1 不变）

$$P(\xi_{t+1} \in \cup_{j=i+1}^{m} S_j \mid \xi_t \in S_i) \geq \frac{1}{n}(1-\frac{1}{n})^i$$

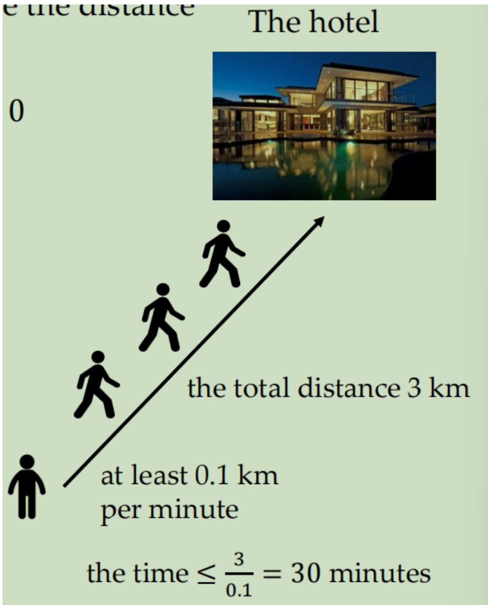
可以得到该问题的期望运行时间上界

$$\sum_{i=0}^{n-1} \pi_0(S_i) \sum_{j=i}^{n-1} \frac{1}{v_j} \leq \sum_{j=0}^{n-1} \frac{1}{v_j}$$

$$= \sum_{j=0}^{n-1} n \frac{1}{(1-\frac{1}{n})^i} \leq en^2 \in O(n^2)$$

故用适应层分析法来分析 (1+1)-EA 算法找到 BinVal 问题最优解的期望运行时间的上界为 $O(n^2)$

2、漂移分析法



用距离函数度量目标状态距离，到达目标时为0，否则大于0

- 目标函数记为LO（x）
- cl：最小步长

计算期望单步漂移距离的上下界

2. Bounds on the expected drift in one step

- $\mathrm{E}[V(\xi_t) - V(\xi_{t+1}) \mid \xi_t] \geq c_l$
- $\mathrm{E}[V(\xi_t) - V(\xi_{t+1}) \mid \xi_t] \leq c_u$

Expected running time

Upper bound: $\sum_{x \in \mathcal{X}} \pi_0(x) \cdot \frac{V(x)}{c_l}$

Lower bound: $\sum_{x \in \mathcal{X}} \pi_0(x) \cdot \frac{V(x)}{c_u}$

- 按与ζt是否有关分类

Additive [He & Yao, AIJ'01]:

$\mathrm{E}[V(\xi_t) - V(\xi_{t+1}) \mid \xi_t] \geq c_l$ → not depend on ξ_t

Upper bound: $\sum_{x \in \mathcal{X}} \pi_0(x) \cdot \frac{V(x)}{c_l}$

Multiplicative [Doerr et al., Algorithmica'12]:

$\mathrm{E}[V(\xi_t) - V(\xi_{t+1}) \mid \xi_t] \geq \delta \cdot V(\xi_t)$ → proportional to the current distance

Upper bound: $\sum_{x \in \mathcal{X}} \pi_0(x) \cdot \frac{1 + \ln (V(x)/V_{min})}{\delta}$

V_{min}: $\min\{V(x) \mid V(x) > 0\}$

3、additive drift analysis加性漂移分析法

例题：(1+1)-EA for LeadingOnes

> **最大化从最左端开始连续1的个数**

LeadingOnes: $max_{x\in\{0,1\}^n} \sum_{i=1}^{n} \prod_{j=1}^{i} x_j$

$LO(x)$

Count the number of consecutive 1-bits starting from the left

e.g., $f(11010) = 2$, $f(01111) = 0$

证明：上界

Theorem. [He & Yao, AIJ'01] The expected running time of the (1+1)-EA solving the LeadingOnes problem is $O(n^2)$.

想证上界，需要下界，因为下界cl在分母

Main idea:

- Design the distance function $V(x) = n - LO(x)$

the number of leading 1-bits

- The expected drift from a solution x with $LO(x) = i$ is lower bounded by

$$E[\,V(\xi_t) - V(\xi_{t+1}) \mid \xi_t = x] \ge \boxed{1 \cdot \frac{1}{n} \cdot (1-\frac{1}{n})^i}$$

keep the i leading 1-bits unchanged

$LO(x) = i \ \rightarrow \ge i+1$

flip the first 0-bit

$c_l = \frac{1}{en}$

Upper bound on the expected running time: $\sum_{x \in \mathcal{X}} \pi_0(x) \frac{V(x)}{c_l} \le \frac{n}{c_l}$

$= n/(\frac{1}{en}) = en^2 \in O(n^2)$

下界：把第一个0翻转，前面的1不动。则长度至少增益为1

> i次幂大于n-1次幂，从而放缩成1/e

4、multiplicative drift analysis乘性漂移分析法

例子：(1+1)-EA for MST

> 本身图是连通的

- 每一位代表一条边是否选
- 最小化问题，修改一下**(1+1)-EA**
- 适应度函数：联通子图的权最小

> 被c(x):连通分支数量支配，因为wub非常大

(1+1)-EA: Given a pseudo-Boolean function f:

1. $\boldsymbol{x} :=$ randomly selected from $\{0,1\}^n$.
2. Repeat until some termination criterion is met
3. $\quad \boldsymbol{x}' :=$ flip each bit of $\boldsymbol{x}$ with probability $1/n$.
4. $\quad$ if $f(\boldsymbol{x}') \ge f(\boldsymbol{x})$ $\longrightarrow$ $f(\boldsymbol{x}') \le f(\boldsymbol{x})$
5. $\quad\quad \boldsymbol{x} = \boldsymbol{x}'$.

Solution representation: $\boldsymbol{x} \in \{0,1\}^m \leftrightarrow$ a subgraph

e.g., $\{e_1, e_2, e_4\} \rightarrow 11010$ $\quad x_i = 1$ means that edge e_i is selected

Fitness function:

the number of connected components

the maximum edge weight

$$\min \ f(\boldsymbol{x}) = (c(\boldsymbol{x}) - 1) \cdot w_{ub} + \sum_{i: x_i = 1} w_i$$

the sum of edge weights

$w_{ub} = n^2 \cdot w_{max}$ to make a subgraph with less connected components better

Theorem. [Neumann & Wegener, TCS'07; Doerr et al., Algorithmica'12] The expected running time of the (1+1)-EA solving the MST problem is $O(m^2(\log n + \log w_{max}))$.

Using multiplicative drift analysis:

- design the distance function: $V(x) = f(x) - f_{opt}$
- analyze the expected drift:

$$E[V(\xi_t) - V(\xi_{t+1}) \mid \xi_t = x] \geq \frac{1}{em(m+1)} V(x)$$

proportional to the current distance

Upper bound on the expected running time:

$$\sum_{x \in \mathcal{X}} \pi_0(x) \cdot \frac{1 + \ln (V(x)/V_{min})}{\delta} \leq em(m+1)(1 + \ln (mw_{max}))$$

$V(x) \leq mw_{max}$　　$V_{min} \geq 1$　　$\in O(m^2(\log n + \log w_{max}))$

- 时间开销=找连通子图的时间+在连通子图中找MST的时间

> 为什么对？因为期望的线性可加性E[T]=E[T1]+E[T2]

第一步：找连通子图的时间：fitness level method

> 使c(x)=1
> 适应层分析法

- 要求：c(x)不会增加；且至少有c(x)-1条边，这些边插进去会使c(x)减少1
- 得出使c(x)减少1的概率：

the probability of decreasing $c(x)$ by 1 is at least $\frac{c(x)-1}{m}(1-\frac{1}{m})^{m-1}$

flip one of those $c(x)-1$ 0-bits　keep the other bits unchanged

> 放缩成1/e

fitness level method $\Longrightarrow$ The expected running time upper bound: $\sum_{c(x)=n}^{2} \frac{em}{c(x)-1} \in O(m \log n)$

第二步：在连通子图中找MST的时间

> 此时c(x)已达到1，且不会变，这个阶段是把权重减到最小
> 乘性飘移分析法

- V(x)设置为与最优fitness的差
- 步骤：删除多余边变成生成树+通过一对一的权重交换使权变小

> 定理：两颗生成树之间可以通过交换边来转移

- 先计算概率公式：

> 第一个等号，因为ζt是给定的

> 第二个等号，把V（ζτ+1）带进去

$$\mathrm{E}[V(\xi_t)-V(\xi_{t+1})\mid \xi_t=\boldsymbol{x}]=V(\boldsymbol{x})-\mathrm{E}[V(\xi_{t+1})\mid \xi_t=\boldsymbol{x}]=f(\boldsymbol{x})-\mathrm{E}[f(\xi_{t+1})\mid \xi_t=\boldsymbol{x}]$$

- m+1次行为平均权重减少量至少为:

$$\geq f(\boldsymbol{x})-\left(\sum_{i=1}^{m-(n-1)} f(\boldsymbol{y}^i)\cdot\frac{1}{m}(1-\frac{1}{m})^{m-1}+\sum_{i=1}^{n} f(\boldsymbol{z}^i)\cdot\frac{1}{m^2}(1-\frac{1}{m})^{m-2}+(1-\cdots)f(\boldsymbol{x})\right)$$

there exists a set of $m-(n-1)$ 1-bit flips and a set of n 2-bit flips such that the average weight decrease is at least $(f(\boldsymbol{x})-f_{opt})/(m+1)$

$$=\sum_{i=1}^{m-(n-1)}(f(\boldsymbol{x})-f(\boldsymbol{y}^i))\cdot\frac{1}{m}(1-\frac{1}{m})^{m-1}+\sum_{i=1}^{n}(f(\boldsymbol{x})-f(\boldsymbol{z}^i))\cdot\frac{1}{m^2}(1-\frac{1}{m})^{m-2}$$

$$\geq\boxed{\frac{1}{m}(1-\frac{1}{m})^{m-1}\cdot\frac{1}{m+1}}\cdot\left(\sum_{i=1}^{m-(n-1)}(f(\boldsymbol{x})-f(\boldsymbol{y}^i))\cdot+\sum_{i=1}^{n}(f(\boldsymbol{x})-f(\boldsymbol{z}^i))\right)$$

$$\geq\frac{1}{m}(1-\frac{1}{m})^{m-1}\cdot\frac{f(\boldsymbol{x})-f_{opt}}{m+1}\geq\frac{1}{em(m+1)}V(\boldsymbol{x})$$

用加性证乘性

2 Problem 2: 加深理解 (15)

课上已讲述“如何用加性漂移分析推出乘性漂移分析”，请将其证明过程整理完善，并说明乘性和加性漂移分析方法的区别。

答：首先定义距离函数

$$U(x)=\begin{cases}0 & \text{if } V(x)=0\\ 1+ln(V(x)/V_{min}) & \text{otherwise}\end{cases}$$

然后计算从某个不属于目标状态空间的状态 ξ_t（即 $V(\xi_t)>0$ 或 $U(\xi_t)>0$）的漂移:
如果 ξ_{t+1} 属于目标状态空间，则

$$U(\xi_t)-U(\xi_{t+1})=1+\ln(V(\xi_t)/V_{min})\geq 1=\frac{V(\xi_t)-V(\xi_{t+1})}{V(\xi_t)}$$

如果 ξ_{t+1} 不属于目标状态空间，则

$$U(\xi_t)-U(\xi_{t+1})=\ln(V(\xi_t)/V(\xi_{t+1}))$$

又有 $1+x\leq e^x$，可得

$$\frac{u}{w}=1+\frac{u-w}{w}\leq e^{\frac{u-w}{w}}$$

$$\ln\frac{u}{w} \le \frac{u-w}{w}$$
$$\ln\frac{w}{u} \ge \frac{w-u}{w}$$

故有

$$\ln(V(\xi_t)/V(\xi_{t+1}) \ge \frac{V(\xi_t)-V(\xi_{t+1})}{V(\xi_t)}$$
$$U(\xi_t)-U(\xi_{t+1}) = \ln(V(\xi_t)/V(\xi_{t+1}) \ge \frac{V(\xi_t)-V(\xi_{t+1})}{V(\xi_t)}$$

综上，无论 ξ_{t+1} 是否属于目标状态空间，都有

$$U(\xi_t)-U(\xi_{t+1}) \ge \frac{V(\xi_t)-V(\xi_{t+1})}{V(\xi_t)}$$

可继续计算从某个不属于目标状态空间的状态 ξ_t（即 $V(\xi_t)>0$ 或 $U(\xi_t)>0$）的期望漂移:

$$E[U(\xi_t)-U(\xi_{t+1}) \mid \xi_t] \ge \frac{E[V(\xi_t)-V(\xi_{t+1}) \mid \xi_t]}{V(\xi_t)} \ge \delta$$

由加性漂移分析的漂移下界和期望运行时间上界：

$$E[U(\xi_t)-U(\xi_{t+1}) \mid \xi_t] \ge c_l$$
$$\sum_{x\in\chi}\pi_0(x)\frac{U(x)}{c_l}$$

可推出乘性漂移分析的漂移下界和期望运行时间上界：

$$E[V(\xi_t)-V(\xi_{t+1}) \mid \xi_t] \ge \delta V(\xi_t)$$
$$\sum_{x\in\chi}\pi_0(x)\frac{1+ln(V(x)/V_{min})}{\delta}$$

二者的区别在于乘性漂移分析中期望单步漂移距离的上下界与当前状态到目标状态空间的距离有关，而加性漂移分析则无关。