

热力学遗传算法计算效率的改进^{*}

应伟勤^{1,2+}, 李元香¹, SHEU Phillip C-Y^{1,2}

¹(武汉大学 软件工程国家重点实验室,湖北 武汉 430072)

²(Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697, USA)

Improving the Computational Efficiency of Thermodynamical Genetic Algorithms

YING Wei-Qin^{1,2+}, LI Yuan-Xiang¹, SHEU Phillip C-Y^{1,2}

¹(State Key Laboratory of Software Engineering, Wuhan University, Wuhan 430072, China)

²(Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697, USA)

+ Corresponding author: E-mail: wyqing@uci.edu

Ying WQ, Li YX, Sheu PCY. Improving the computational efficiency of thermodynamical genetic algorithms. *Journal of Software*, 2008,19(7):1613–1622. <http://www.jos.org.cn/1000-9825/19/1613.htm>

Abstract: Thermodynamical genetic algorithms (TDGA) simulate the competitive model between energy and entropy in annealing to harmonize the conflicts between selective pressure and population diversity in GA. But high computational cost restricts the applications of TDGA. In order to improve the computational efficiency, a measurement method of rating-based entropy (RE) is proposed. The RE method can measure the fitness dispersal with low computational cost. Then a component thermodynamical replacement (CTR) rule is introduced to reduce the complexity of the replacement, and it is proved that the CTR rule has the approximate steepest descent ability of the population free energy. Experimental results on 0-1 knapsack problems show that the RE method and the CTR rule not only maintain the excellent performance and stability of TDGA, but also remarkably improve the computational efficiency of TDGA.

Key words: genetic algorithm; thermodynamics; computational efficiency; diversity measurement; replacement rule

摘要: 热力学遗传算法(thermodynamical genetic algorithms,简称 TDGA)借鉴固体退火过程中能量与熵的竞争模式来协调 GA 中“选择压力”和“种群多样性”之间的冲突.然而 TDGA 目前极高的计算代价限制了其应用.为了提高 TDGA 的计算效率,首先定义一种等级熵(rating-based entropy,简称 RE)度量方法,它能以较小的计算成本度量种群中个体适应值的分散程度.然后引入分量热力学替换规则(component thermodynamical replacement,简称 CTR),有效地降低了替换规则的复杂度.同时也证明了 CTR 规则具有驱动种群自由能近似最速下降的能力.在 0-1 背包问题上的实验结果表明,RE 方法和 CTR 规则在保持 TDGA 良好的性能与稳定性的同时,极大地提高了其计算效率.

* Supported by the National Natural Science Foundation of China under Grant Nos.60473014, 60773009 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z290 (国家高技术研究发展计划(863)); the China Scholarship Council under Grant No.2007101731 (国家留学基金); the Natural Science Foundation of Hubei Province of China under Grant No.2007ABA009 (湖北省自然科学基金)

Received 2007-12-18; Accepted 2008-03-14

关键词: 遗传算法;热力学;计算效率;多样性度量;替换规则

中图法分类号: TP18 文献标识码: A

传统遗传算法(genetic algorithm,简称GA)作为一种基于“优胜劣汰”自然选择机制的优化方法,同时面临着收敛速度慢和早熟收敛两方面的困扰^[1].其中早熟收敛是指由于种群过早地遗失多样性而导致算法经常在搜索过程中陷入局部最优的现象.早熟收敛也极大地影响了GA的稳定性.Whitley^[2]认为,GA中最重要的两个因素就是“选择压力”和“种群多样性”.很多学者从这两方面提出了改进GA的方法.例如,可以通过精英选择策略、稳态替换策略^[3]等加大GA的选择压力.在维持种群多样性方面,通过适应度共享^[4]来创造出小生境进化环境是最常见的一种方法;Li等人^[5]则将GA中的种群视为一个动力学系统,个体视为粒子,遗传操作视为粒子的碰撞或移动,通过驱动所有粒子尽可能地运动来保持种群多样性.但是Whitley也指出,“选择压力”和“种群多样性”是相互制约的^[2].增大选择压力虽然可以加快算法的收敛速度,但会导致种群多样性的更快流失,增加算法早熟收敛的风险;降低选择压力虽然可以增大算法搜索到全局最优值的概率,但会降低搜索效率.因此,为了使GA具有良好的性能,必须在提高“选择压力”和保持“种群多样性”之间维持适当的平衡.

为了定量协调“选择压力”和“种群多样性”的冲突,Mori等人^[6-8]从固体退火过程^[9]得到启发提出了热力学遗传算法(thermodynamical genetic algorithm,简称TDGA).在退火过程中,系统在恒温下的任何状态改变都遵循自由能极小定律^[9]:对于与外界交换热量而温度保持不变的封闭系统,系统状态的自发变化总是朝着自由能减少的方向进行,当自由能达到极小值时系统达到平衡态.在热力学中,系统的自由能 $F=E-HT$,其中 E 是系统的能量, H 是系统的热力学熵.根据该定律,能量减少和熵增加均有利于减少自由能,因此系统的任何改变都可被视为 E 和 H 的竞争结果,温度 T 决定着两者的相对权重.退火中的这种竞争机制有效地协调了两者的冲突,使金属固体最终稳定达到低能低熵的基态.在退火过程与种群收敛过程之间存在着诸多相似性.例如,种群可视为热力学系统,个体对应粒子,种群的负平均适应值、种群多样性和一个权重控制参数 T 则分别可担当退火中能量、熵和温度的角色.TDGA正是通过模拟退火中能量与熵之间的竞争模式来定量协调“选择压力”与“种群多样性”的冲突.

Mori^[6]最初在 0-1 背包问题上验证了TDGA的有效性,之后Maekawa^[10]巧妙地设计了熵的定义,将TDGA成功应用到TSP问题.其后,Ding^[11]曾将TDGA用于求解多目标优化问题,Zheng^[12]也将TDGA中的热力学选择机制推广到粒子群优化算法.但是TDGA极高的计算代价限制了它的进一步实际应用,这也是TDGA目前存在的主要问题,还没有得到很好的解决.导致TDGA计算效率极低的两个主要因素在于:1) TDGA使用基因熵来度量种群多样性^[6],编码较长时基因熵的计算成本很大;2) TDGA使用贪婪热力学替换规则驱动种群自由能快速下降^[6],而该规则具有较高的时间复杂度.为了提高TDGA的计算效率,本文引入了一种等级熵度量方法和一种分量热力学替换规则.其中等级熵方法在适应值空间中动态度量种群多样性,不再依赖于个体编码长度;而分量热力学替换规则具有较低的时间复杂度,同时,本文亦运用极限夹逼准则完整地证明了该规则具有驱动种群自由能近似最速下降的能力.在0-1背包问题上的实验结果显示,等级熵度量方法和分量热力学替换规则在保持TDGA良好的性能和稳定性的同时,极大地提高了其计算效率.

1 种群多样性度量

在GA中引入退火过程的竞争机制时首先要解决的两个关键问题是:1) 如何定义能够度量种群多样性的熵;2) 如何设计能够驱动种群自由能快速下降的热力学替换规则.这也是本文主要讨论的两个问题.

1.1 基因熵度量方法

在文献[6]中应用TDGA求解二进制编码的0-1背包问题时,Mori采用种群在每一基因位上基因的信息熵的和来度量种群多样性.本文称其为基因熵(gene-based entropy,简称GE),它度量了种群的基因多样性.然而,当编码较长时,所有基因位上信息熵的重复性计算过程花费了很高的计算成本.

1.2 等级熵度量方法

为了更高效地度量种群多样性,在文献[13]中我们先获取适应值上下界,然后在上下界之间定义熵.本文进一步通过动态地在适应值空间的活跃窗口中划分等级而引入了等级熵(rating-based entropy,简称 RE)这一概念,从适应值分散程度的角度来度量种群多样性.

定义 1(个体的绝对能量). 设 S 为搜索空间, $f: S \rightarrow R$ 为目标函数,则对于任意个体 $X_r \in S$,在 f 为最大优化问题时个体 X_r 的绝对能量 $e(X_r) = f(X_r)$,在 f 为最小优化问题时 $e(X_r) = -f(X_r)$.

定义 2(活跃窗口和等级). 设 $P_t = (X_1, X_2, \dots, X_N) \in S^N$ 为 TDGA 的第 t 代种群,则 TDGA 在第 t 代的活跃窗口 $w_t = [l_t, u_t]$ 由如下两条规则生成:1) $w_0 = [l_0, u_0]$, 这里, $l_0 = \min\{e(X_r) | X_r \in P_0\}$, $u_0 = \max\{e(X_r) | X_r \in P_0\}$; 2) 若 $w_t = [l_t, u_t]$ 且 TDGA 第 t 代的子种群为 $O_t = (X_{N+1}, X_{N+2}, \dots, X_{N+M}) \in S^M$, 那么 $w_{t+1} = [l_{t+1}, u_{t+1}]$, 这里, $l_{t+1} = \min(l_t, \min\{e(X_r) | X_r \in O_t\})$, $u_{t+1} = \max(u_t, \max\{e(X_r) | X_r \in O_t\})$. 我们称 $\beta_i^{(t)}$ 为活跃窗口 w_t 的第 i 个等级, 这里,

$$\beta_i^{(t)} = \left(\frac{a^{i-1} - 1}{a^{K-1} - 1} (u_t - l_t) + l_t, \frac{a^i - 1}{a^{K-1} - 1} (u_t - l_t) + l_t \right) \cap [l_t, u_t], \quad a > 1, 0 \leq i \leq K - 1, K \geq 2 \quad (1)$$

其中, a 为比例因子常量, K 为等级数. 对于任意 $X_r \in P_t$, 若 $e(X_r) \in \beta_i^{(t)}$, 则称个体 X_r 落入到了等级 $\beta_i^{(t)}$ 中.

定义 3(等级熵). 设 $P_t = (X_1, X_2, \dots, X_N) \in S^N$ 和 $w_t = [l_t, u_t]$ 分别为 TDGA 在第 t 代的种群和活跃窗口, 且种群 P_t 中落入 w_t 的第 i 个等级 $\beta_i^{(t)}$ 中的个体数为 n_i , 那么称 $H(w_t, P_t)$ 为种群 P_t 在窗口 w_t 中的等级熵, 这里,

$$H(w_t, P_t) = - \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \frac{n_i}{N} \quad (2)$$

绝对能量度量了个体的优劣程度, 绝对能量越低的个体越优秀. 活跃窗口 w_t 记录了到第 t 代为止搜索过程已这个体的能量范围, 它将随搜索过程的进行不断扩张. 在对活跃窗口进行分级时, 遵循了适应值区域越优分级越精细的原则. 可以证明活跃窗口中的所有等级正好构成了该窗口的一个划分. 等级熵 $H(w_t, P_t)$ 以较小计算成本度量了种群中个体适应值的分散程度, 它不依赖于个体的编码长度. 可以证明, 当 P_t 中所有个体全部落入同一等级时等级熵 $H(w_t, P_t)$ 取最小值 0, 当 P_t 中落入各等级的个体数都相等, 即 $n_0 = n_1 = \dots = n_{K-1}$ 时, $H(w_t, P_t)$ 达到最大值 1.

2 热力学替换规则

热力学替换规则的设计是引入退火过程竞争机制的另一个关键问题. 首先定义种群的自由能.

定义 4(相对能量). 设 $P_t = (X_1, X_2, \dots, X_N) \in S^N$ 和 $w_t = [l_t, u_t]$ 分别为 TDGA 在第 t 代的种群和活跃窗口, 对于任意 $X_r \in P_t$, 称 $e'(w_t, X_r)$ 是个体 X_r 在窗口 w_t 中的相对能量, 这里,

$$e'(w_t, X_r) = \frac{l_t - e(X_r)}{l_t - u_t} \quad (3)$$

定义 5(自由能). 设 $P_t = (X_1, X_2, \dots, X_N) \in S^N$ 为 TDGA 在第 t 代的种群, 则称 $E(w_t, P_t)$ 为在窗口 w_t 中种群 P_t 的能量, 称 $F(w_t, T, P_t)$ 为在窗口 w_t 和温度 $T(T \geq 0)$ 下种群 P_t 的自由能, 这里,

$$E(w_t, P_t) = \frac{1}{N} \sum_{X_r \in P_t} e'(w_t, X_r) \quad (4)$$

$$F(w_t, T, P_t) = E(w_t, P_t) - TH(w_t, P_t) \quad (5)$$

自由能极小化是退火过程中驱动恒温下的热力学系统从能量与熵的竞争不平衡态趋向竞争平衡态的动力. 与此类似, 在进化过程中的每一温度 T_k 下算法应驱动种群快速最小化其自由能. 为此需要设计一种热力学替换规则, 使下一代种群的自由能最速下降. 这种热力学替换规则的任务是从父代种群 $P_t = (X_1, X_2, \dots, X_N)$ 的 N 个个体及其子种群 $O_t = (X_{N+1}, X_{N+2}, \dots, X_{N+M})$ 中的 M 个个体中挑选出 N 个个体组成下一代种群 $P_{t+1}^{(E)}$, 使其具有的自由能 $F(w_{t+1}, T, P_{t+1}^{(E)})$ 最小. 然而, 精确地最小化每一代的自由能本身是一个很难的组合优化问题, 其复杂度高达 $O((N+K)C_{N+M}^N)$, 实际中是不可行的.

2.1 贪婪热力学替换规则

为了降低替换规则的复杂度, 文献[6]Mori 采用了贪婪热力学替换(greedy thermodynamical replacement, 简

称 GTR)规则.GTR 规则采用贪婪的策略逐个往下一代种群中填充个体.GTR 规则的计算开销较之精确方法有很大降低,其复杂度为 $O(N(N+M)K)$,但在实际应用中,计算成本仍相当高,而且其驱动自由能快速下降的能力还没有得到有效证明.

2.2 分量热力学替换规则

本文通过自由能分量的概念将种群的自由能分派到其各个个体上,进而引入了一种分量热力学替换(component thermodynamical replacement,简称 CTR)规则,极大地简化了热力学替换规则的设计.

定义 6(自由能分量). 设 $P_t=(X_1, X_2, \dots, X_N) \in \mathcal{S}^N$ 和 $w_t=[l_t, u_t]$ 分别为 TDGA 在第 t 代的种群和活跃窗口,对于个体 $X_r \in P_t$,若 X_r 落入在 w_t 的第 b 个等级 $\beta_t^{(b)}$ 中,且种群 P_t 中落入等级 $\beta_t^{(b)}$ 中的个体数为 n_b ,那么称 $F_c(w_t, T, P_t, X_r)$ 为在窗口 w_t 和温度 T 下个体 X_r 在种群 P_t 中的自由能分量,这里,

$$F_c(w_t, T, P_t, X_r) = e'(w_t, P_t) + T \log_K \left(\frac{n_b}{N} \right) \quad (6)$$

定理 1. 种群 P_t 的自由能等于该种群中各个个体自由能分量的均值,即

$$F(w_t, T, P_t) = \frac{1}{N} \sum_{X_r \in P_t} F_c(w_t, T, P_t, X_r) \quad (7)$$

定理 1 揭示了自由能与自由能分量之间的关系,本文省略了定理 1 的证明,其证明过程可参考第 2.3 节中引理 2.下面算法 1 给出了分量热力学替换规则 CTR 的描述.

算法 1. 分量热力学替换规则 CTR.

输入:第 $t+1$ 代的活跃窗口 w_{t+1} ,当前温度 T ,第 t 代种群 P_t 和第 t 代的子种群 O_t .

输出:第 $t+1$ 代种群 P_{t+1} .

- (1) 将子种群 O_t 的 M 个个体添加到父代种群 P_t 中,从而生成一个规模为 $N+M$ 的中间种群 P'_{t+1} ;
- (2) 计算中间种群 P'_{t+1} 中所有个体的自由能分量 $F_c(w_{t+1}, T, P'_{t+1}, X_r), X_r \in P'_{t+1}$;
- (3) 从 P'_{t+1} 中挑选出 M 个具有最大自由能分量的个体;
- (4) 删除 P'_{t+1} 中这 M 个个体即可得到下一代种群 P_{t+1} .

2.3 CTR 规则的近似最速下降能力

CTR 规则具有比 GTR 规则更低的复杂度 $O((N+M)M)$,明显降低了 GTR 规则的计算代价.同时可以证明,CTR 规则具有驱动种群的自由能近似最速下降的能力,这保证了 CTR 规则的有效性.本节在引入两个引理的基础上,运用极限夹逼定理,从理论上完整地证明了 $M \ll N$ 时,由 CTR 规则生成的种群 $P_{t+1}^{(C)}$ 的自由能 $F(w_{t+1}, T, P_{t+1}^{(C)})$ 与下一代种群的可能最小自由能 $F(w_{t+1}, T, P_{t+1}^{(E)})$ 的差值趋于 0.

引理 1. 对于父代种群 P_t 和子种群 O_t ,设 $P_{t+1}^{(E)}$ 为自由能最小的下一代种群, $P_{t+1}^{(C)}$ 为由 CTR 规则生成的下一代种群,即 $P_{t+1}^{(C)} = \text{CTR}(w_{t+1}, T, P_t, O_t)$,那么,

$$\frac{1}{N} \sum_{X_r \in P_{t+1}^{(C)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) - \frac{1}{N} \sum_{X_r \in P_{t+1}^{(E)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) \leq 0 \quad (8)$$

证明:设 P'_{t+1} 中个体已按自由能分量升序排列,即 $P'_{t+1} = \{X_1, X_2, \dots, X_N, \dots, X_{N+M}\}$,当 $1 \leq u \leq v \leq M+N$ 时, $F_c(w_{t+1}, T, P'_{t+1}, X_u) \leq F_c(w_{t+1}, T, P'_{t+1}, X_v)$.此时 $P_{t+1}^{(C)} = \{X_1, X_2, \dots, X_N\}$,那么,

$$\sum_{X_r \in P_{t+1}^{(C)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) = \sum_{u=1}^N F_c(w_{t+1}, T, P'_{t+1}, X_u) \quad (9)$$

由于 $P_{t+1}^{(E)} \subseteq P'_{t+1}$,可设 $P_{t+1}^{(E)}$ 中个体已按在 P'_{t+1} 中出现顺序排列,即 $P_{t+1}^{(E)} = \{X_{q(1)}, X_{q(2)}, \dots, X_{q(N)}\}, 1 \leq q(1) < q(2) < \dots < q(N) \leq M+N$,那么,

$$\sum_{X_r \in P_{t+1}^{(E)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) = \sum_{u=1}^N F_c(w_{t+1}, T, P'_{t+1}, X_{q(u)}) \quad (10)$$

由数学归纳法容易证得 $1 \leq u \leq N$ 时 $u \leq q(u)$, 于是有 $(w_{t+1}, T, P'_{t+1}, X_u) \leq F_c(w_{t+1}, T, P'_{t+1}, X_{q(u)})$, 进而,

$$\frac{1}{N} \sum_{u=1}^N F_c(w_{t+1}, T, P'_{t+1}, X_u) \leq \frac{1}{N} \sum_{u=1}^N F_c(w_{t+1}, T, P'_{t+1}, X_{q(u)}) \quad (11)$$

联合式(9)~式(11), 得到式(8). □

引理 2. 设 $P_{t+1}^{(E)}$ 为自由能最小的下一代种群, $P_{t+1}^{(C)} = CTR(w_{t+1}, T, P_t, O)$, 那么,

$$0 \leq F(w_{t+1}, T, P_{t+1}^{(C)}) - \frac{1}{N} \sum_{X_r \in P_{t+1}^{(C)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) \leq T \log_K((M+N)/N) \quad (12)$$

$$0 \leq F(w_{t+1}, T, P_{t+1}^{(E)}) - \frac{1}{N} \sum_{X_r \in P_{t+1}^{(E)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) \leq T \log_K((M+N)/N) \quad (13)$$

证明: 首先可为活跃窗口 w_{t+1} 引入一个等级指示函数:

$$g(X_r, \beta_{t+1}^{(i)}) = \begin{cases} 1, & \text{if } e(X_r) \in \beta_{t+1}^{(i)} \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

使用该指示函数, $P_{t+1}^{(C)}$ 中第 i 等级的个体数 n_i 以及 P'_{t+1} 中第 i 等级的个体数 n'_i 可描述为

$$n_i = \sum_{u=1}^N g(X_u, \beta_{t+1}^{(i)}), n'_i = \sum_{u=1}^{N+M} g(X_u, \beta_{t+1}^{(i)}) \quad (15)$$

且它们具有如下性质:

$$\sum_{i=0}^{K-1} n_i = N, \sum_{i=0}^{K-1} n'_i = N+M, n_i \leq n'_i \quad (16)$$

根据定义 6, 个体 X_u 在种群 P'_{t+1} 中的自由能分量亦可等价表示为

$$F_c(w_{t+1}, T, P'_{t+1}, X_u) = e'(w_{t+1}, X_u) + T \sum_{i=0}^{K-1} g(X_u, \beta_{t+1}^{(i)}) \log_K \left(\frac{n'_i}{N+M} \right) \quad (17)$$

由式(9)、式(15)、式(17)和定义 5 可得:

$$\begin{aligned} F(w_{t+1}, T, P_{t+1}^{(C)}) - \frac{1}{N} \sum_{X_r \in P_{t+1}^{(C)}} F_c(w_{t+1}, T, P'_{t+1}, X_r) &= F(w_{t+1}, T, P_{t+1}^{(C)}) - \frac{1}{N} \sum_{u=1}^N F_c(w_{t+1}, T, P'_{t+1}, X_u) \\ &= \left(\frac{1}{N} \sum_{u=1}^N e'(w_{t+1}, X_u) + T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n_i}{N} \right) \right) - \frac{1}{N} \sum_{u=1}^N \left(e'(w_{t+1}, X_u) + T \cdot \sum_{i=0}^{K-1} g(X_u, \beta_{t+1}^{(i)}) \log_K \left(\frac{n'_i}{N+M} \right) \right) \\ &= T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n_i}{N} \right) - T \cdot \frac{1}{N} \cdot \sum_{i=0}^{K-1} \left(\log_K \left(\frac{n'_i}{N+M} \right) \sum_{u=1}^N g(X_u, \beta_{t+1}^{(i)}) \right) \\ &= T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n_i}{N} \right) - T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n'_i}{N+M} \right) = -T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n'_i/(N+M)}{n_i/N} \right) \end{aligned} \quad (18)$$

一方面, 由式(16)可得:

$$-T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n'_i/(N+M)}{n_i/N} \right) \leq -T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n_i/(N+M)}{n_i/N} \right) = -T \cdot \log_K \left(\frac{N}{N+M} \right) \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} = T \cdot \log_K \left(\frac{N+M}{N} \right) \quad (19)$$

另一方面, 由自然对数的性质 $\ln x \leq x-1$ 可得:

$$\frac{n_i}{N} \log_K \left(\frac{n'_i/(N+M)}{n_i/N} \right) = \frac{n_i}{N} \cdot \frac{1}{\ln K} \cdot \ln \left(\frac{n'_i/(N+M)}{n_i/N} \right) \leq \frac{1}{\ln K} \cdot \frac{n_i}{N} \cdot \left(\frac{n'_i/(N+M)}{n_i/N} - 1 \right) = \frac{1}{\ln K} \cdot \left(\frac{n'_i}{N+M} - \frac{n_i}{N} \right) \quad (20)$$

由式(16)和式(20)可得:

$$-T \cdot \sum_{i=0}^{K-1} \frac{n_i}{N} \log_K \left(\frac{n'_i/(N+M)}{n_i/N} \right) \geq -T \cdot \frac{1}{\ln K} \cdot \left(\sum_{i=0}^{K-1} \frac{n'_i}{N+M} - \sum_{i=0}^{K-1} \frac{n_i}{N} \right) = -T \cdot \frac{1}{\ln K} \cdot \left(\frac{N+M}{N+M} - \frac{N}{N} \right) = 0 \quad (21)$$

联合式(18)、式(19)和式(21)可得式(12). 同理亦可证明式(13). □

定理 2. 设 $y=M/(N+M)$ 和 $D(w_{t+1}, T, P_t, O) = F(w_{t+1}, T, P_{t+1}^{(C)}) - F(w_{t+1}, T, P_{t+1}^{(E)})$, 则

$$\lim_{y \rightarrow 0^+} (D(w_{t+1}, T, P_t, O_t)) = 0 \quad (22)$$

证明:由引理 2 中式(12)和式(13)相减可得如下不等式:

$$D(w_{t+1}, T, P_t, O_t) \leq T \log_K ((M+N)/N) + \left(\frac{1}{N} \sum_{X_r \in P_{t+1}^{(C)}} F_c(w_{t+1}, T, P_{t+1}, X_r) - \frac{1}{N} \sum_{X_r \in P_{t+1}^{(E)}} F_c(w_{t+1}, T, P_{t+1}, X_r) \right) \quad (23)$$

然后由式(23)和引理 1 可得:

$$D(w_{t+1}, T, P_t, O_t) \leq T \log_K ((M+N)/N) \quad (24)$$

由于 $P_{t+1}^{(E)}$ 是自由能最小的下一代种群,因此,

$$D(w_{t+1}, T, P_t, O_t) \geq 0 \quad (25)$$

同时存在如下极限:

$$\lim_{y \rightarrow 0^+} (T \log_K ((M+N)/N)) = 0 \quad (26)$$

运用极限的夹逼准则,由(24)~式(26)可得式(22). \square

定理 2 告诉我们,当 $M \ll N$ 时,CTR 规则具有驱动种群自由能至少以接近最快的进化速度下降的能力.定理 2 保证了在 $M \ll N$ 时 CTR 规则的有效性.

3 算法描述

采用等级熵度量方法和分量热力学替换规则的 TDGA 算法的一般流程见算法 2.

算法 2. 使用 RE 方法和 CTR 规则的热力学遗传算法 TDGA(RE+CTR).

输入:搜索空间 S 和目标函数 f .

输出:最优解 ζ .

- (01) 设置比例因子 a , 等级数 K , 以及温度 T_k 下的 Markov 链长 L_k ;
- (02) 随机生成 N 个个体作为初始种群 P_0 , 并评估 P_0 中个体;
- (03) 按定义 2 中规则 1) 生成第 0 代的活跃窗口 w_0 ;
- (04) $T_0=10; t=0; k=0$;
- (05) **while**(Termination_test(P_t))==False){
- (06) **for**($i=0; i < L_k; i++$){
- (07) 通过均匀选择以及交叉和变异操作生成 M 个子个体;
- (08) 将这 M 个个体组织成子种群 O_t , 并评估 O_t 中个体;
- (09) 按定义 2 中规则 2) 生成第 $t+1$ 代的活跃窗口 w_{t+1} ;
- (10) $P_{t+1}=CTR(w_{t+1}, T_k, P_t, O_t)$;
- (11) $t=t+1$;
- (12) }
- (13) $k=k+1$;
- (14) $T_k=T_0/(1+k)$;
- (15) }
- (16) 将第 t 代的最好个体作为最优解 ζ 输出;

上述 TDGA(RE+CTR) 算法在每一温度 T_k 下进行 L_k 代的竞争, 在温度 T_k 下的这些竞争形成了一条长为 L_k 的 Markov 链, 这里, 长 L_k 应随问题的难度而增大. 初始温度 T_0 、温度衰减函数 $T_k = \alpha(T_0, k)$ 以及温度 T_k 下 Markov 链的长度 L_k 共同构成了冷却进度表. 根据固体退火过程, 初始温度应足够高, 并且温度衰减应足够缓慢. 考虑到本文中种群的能量和等级熵均已归一化, 温度值的设置不必再依赖于具体问题的适应值范围, 故本文采用了统一的初始温度 $T_0=10$ 和温度衰减函数 $T_k=T_0/(1+k)$.

4 实验结果及分析

为了检验上述采用等级熵方法和分量替换规则的TDGA(RE+CTR)算法的性能及计算效率,我们用GALib^[14]和C++实现了相关算法,并在 0-1 背包问题上对这些算法进行了一系列对比实验.所有实验均在AMD 3600+ CPU/1G内存的计算机上进行.

4.1 测试例

0-1 背包问题是典型的NP完全难题,Martello^[15]指出 0-1 背包问题的难度在很大程度上受价值和重量的相关性影响,并且越高相关性的问题具有越高的预期求解难度.因此,他提出了 3 种类型的背包问题:无相关(un-correlated,简称UC)、弱相关(weakly correlated,简称WC)和强相关(strongly correlated,简称SC).本文按Pisinger^[16]提出的背包问题测试例构造算法为每种类型的背包问题生成一个测试例,它们对应的Pisinger构造算法的输入参数以及背包容量、理论上界以及目前已知最优解见表 1.

Table 1 Test instances
表 1 测试例

Instance	Parameters					Capacity	Upper bound	Best known solution
	Size	Range	Sum	Seed	Type			
KP1	100	1 000	1 000	750	uc	36 202	49 940	49 910
KP2	100	1 000	1 000	750	wc	36 202	38 249	38 245
KP3	100	1 000	1 000	750	sc	38 217	39 080	39 077

4.2 对比实验

本文选择 4 种遗传算法来对比分析CTR规则的性能和计算效率.4 种算法分别为简单遗传算法(simple genetic algorithm,简称SGA)、稳态遗传算法(steady state genetic algorithm,简称SSGA)^[3]、采用等级熵方法与贪婪规则的热力学遗传算法TDGA(RE+GTR)以及采用等级熵方法与分量规则的热力学遗传算法TDGA(RE+CTR).所有算法都使用相当的终止条件,均采用二进制串编码、均匀交叉算子和翻转变异算子,并对不满足约束的个体都按贪婪准则进行修补.算法的共同参数如下设置:种群规模 $N=80$,交叉概率 $P_c=0.6$,变异概率 $P_m=0.02$,终止条件 *Termination_test* 在评估次数 $s \geq 3.2 \times 10^6$ 时满足.SSGA和TDGA中子种群 O_i 的规模 $M=8$,TDGA中温度 T_k 下的Markov链长度 $L_k=1000$,等级熵方法RE的比例因子常量 $a=2$,等级数 $K=1/5 \times N=16$.

应用 4 种算法各求解 3 个测试例 20 次.表 2 列出了在 3 个测试例上 4 种算法的性能统计信息,包括成功收敛到已知最优解的比率、20 次最好个体适应值的平均数和标准差、平均运行时间(s)和平均首达时间(s),其中首达时间表示算法首次达到最优解的时间.我们根据首次达到最优解时的个体评估次数 s 对算法的每次运行情况进行如下性能评价:(1) 优秀(excellent): $s \leq 1.6 \times 10^5$;(2) 良好(good): $1.6 \times 10^5 < s \leq 3.2 \times 10^6$;(3) 糟糕(poor): $s > 3.2 \times 10^6$.图 1 显示了在KP2和KP3上 4 种算法在 20 次运行中的性能评价分布直方图,其中纵坐标为 20 次运行中算法获得各性能评价的比率.图 2 给出了在KP2和KP3测试例上 4 种算法种群等级熵的平均变化曲线.

表 2 清晰地展示了 TDGA 的稳定性.对于求解难度较低的 KP1 问题,4 种算法每次都能成功达到最优解.对于求解难度较高的 KP2 和 KP3 问题,SGA 和 SSGA 的成功收敛率急剧下降;而 TDGA(RE+GTR)和 TDGA(RE+CTR)每次都成功收敛,表现出了更稳健的搜索能力.正是由于 TDGA(RE+GTR)和 TDGA(RE+CTR)具有极高的稳定性,其在 KP2 和 KP3 上解的平均值与方差明显优于 SGA 和 SSGA.在图 1(a)和图 1(b)中选择压力较大的 SSGA 在 KP2 和 KP3 上均呈现了较严重的“两极分化”现象,即算法要么在进化初期达到最优解,要么就达不到最优解.这表明 SSGA 经常出现早熟收敛现象.从图 2(a)和图 2(b)可以看出,SSGA 在进化早期等级熵快速下降至 0,种群多样性迅速流失,这是导致早熟收敛的主要原因.而 TDGA(RE+GTR)和 TDGA(RE+CTR)的等级熵在进化初期先迅速增大,而后随温度降低缓缓下降.TDGA(RE+GTR)和 TDGA(RE+CTR)由于成功地维持了“选择压力”和“种群多样性”之间的平衡,因而表现出了极强的持续优化能力,有效避免了这种“两极分化”现象的发生.

表 2 中数据也充分说明了 CTR 规则具有较高的计算效率.TDGA(RE+CTR)的运行时间几乎与 SSGA 相当,略高于 SGA.这与 GTR 规则极高的计算代价形成了鲜明的对比,TDGA(RE+GTR)的运行时间约 25 倍于

TDGA(RE+CTR),36 倍于 SGA.首达时间也是衡量 GA 性能的一个重要指标.表 2 显示了 TDGA(RE+CTR)在 3 个测试例上都具有比 TDGA(RE+GTR)快得多的首达时间,这应得益于 CTR 规则同时具有良好的稳定性和较高的计算效率.

Table 2 Statistic results of algorithms on three instances
表 2 算法在 3 个测试例上的统计结果

Test instance	Algorithm	Success ratio	Mean	Standard deviation	Time	First hitting time
KP1	SGA	20/20	49 910.00	0	101.28	=2.283
KP1	SSGA	20/20	49 910.00	0	127.52	=0.338
KP1	TDGA(RE+GTR)	20/20	49 910.00	0	3 643.76	=15.122
KP1	TDGA(RE+CTR)	20/20	49 910.00	0	140.34	=0.742
KP2	SGA	2/20	38 226.5	21.197 9	98.12	>91.555
KP2	SSGA	3/20	38 237.35	8.253 9	128.12	>108.947
KP2	TDGA(RE+GTR)	20/20	38 245	0	3 629.36	=498.856
KP2	TDGA(RE+CTR)	20/20	38 245	0	144.40	=10.520
KP3	SGA	0/20	39 039.15	7.073	101.24	>101.24
KP3	SSGA	8/20	39 071.7	5.788 8	148.72	>91.649
KP3	TDGA(RE+GTR)	20/20	39 077	0	3 661.84	=251.385
KP3	TDGA(RE+CTR)	20/20	39 077	0	143.76	=8.992

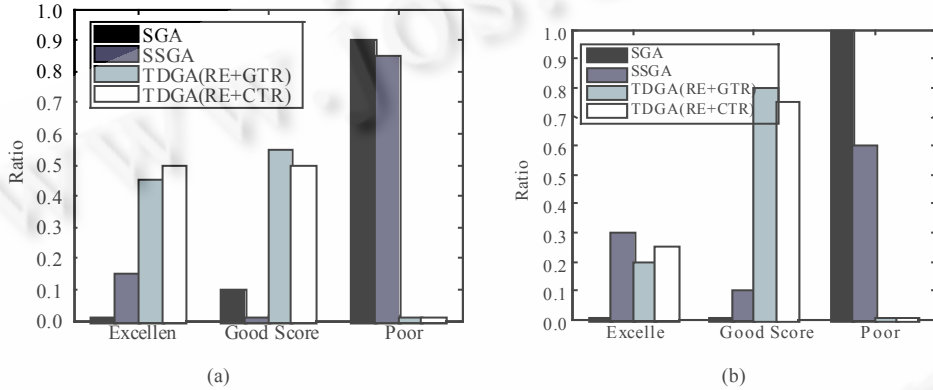


Fig.1 Performance distributions of algorithms on KP2 (a) and KP3 (b)
图 1 算法在 KP2(a)和 KP3(b)上的性能分布直方图

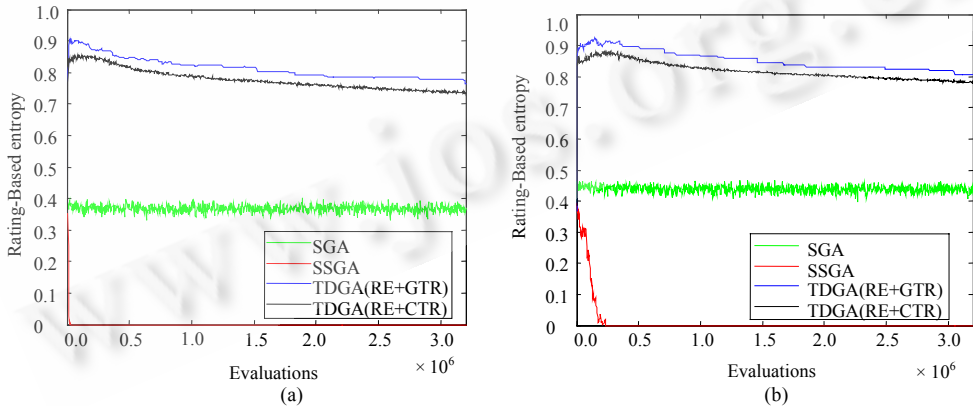


Fig.2 RE curves of algorithms on KP2 (a) and on KP3 (b)
图 2 算法在 KP2(a)和 KP3(b)上的等级熵变化曲线

为了进一步对比分析等级熵方法的计算效率,我们也在 3 个测试例上各运行采用基因熵方法与贪婪规则的 TDGA(GE+GTR)热力学遗传算法 10 次,图 3(a)呈现了采用基因熵的 TDGA(GE+GTR)和采用等级熵的 TDGA(RE+GTR)之间的平均运行时间对比直方图,图 3(b)进一步呈现了采用 GTR 规则的 TDGA(RE+GTR)和采用 CTR 规则的 TDGA(RE+CTR)之间的平均运行时间对比直方图.由图 3 看出,对于采用 GE 方法和 GTR 规则

的 TDGA 算法,若采用 RE 方法代替 GE 方法,计算成本稳定下降了约 16 倍;若进一步采用 CTR 规则代替 GTR 规则,计算成本继续稳定下降约 25 倍.RE 方法和 CTR 规则两者的使用将 TDGA 的计算效率共提高了约 400 倍.

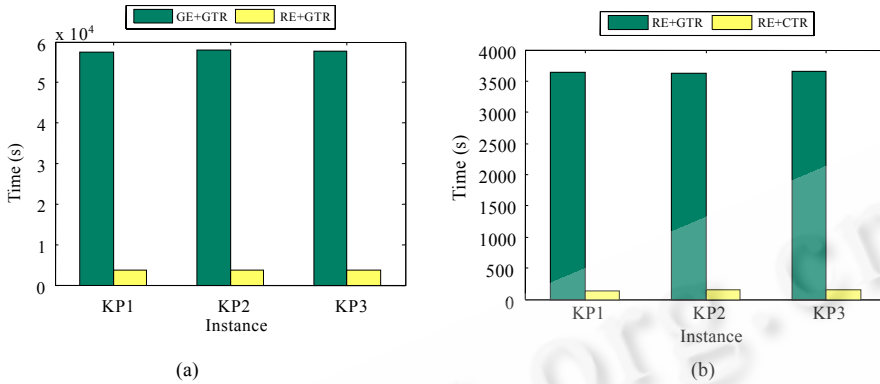


Fig.3 Computational cost comparisons between GE and RE (a) and between GTR and CTR (b)

图 3 GE 与 RE 之间(a)以及 GTR 与 CTR 之间(b)的计算代价对比直方图

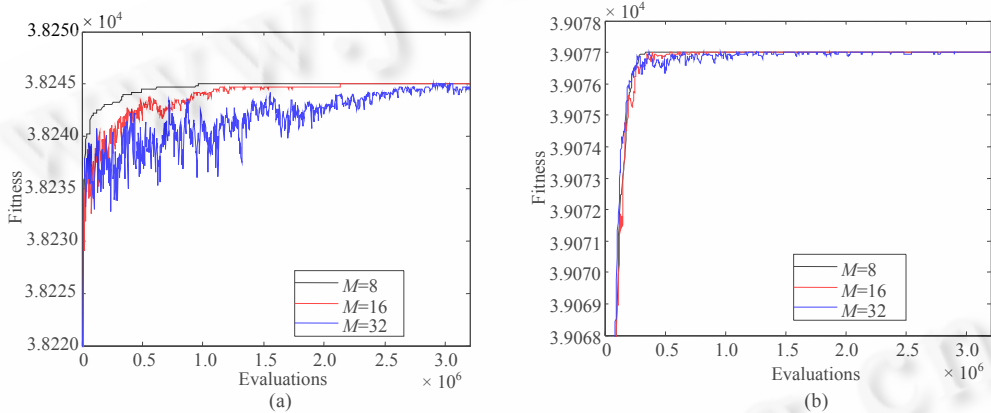


Fig.4 Convergence curves of TDGA(RE+CTR) on KP2 (a) and KP3 (b)

图 4 在 KP2(a)和 KP3(b)上 TDGA(RE+CTR)的收敛曲线

根据定理 2 我们也测试了子种群规模 M 对 CTR 规则的性能影响,图 4 给出了 $M=8,16,32$ 时 TDGA(RE+CTR) 在 KP2 和 KP3 上的平均收敛曲线.实验结果显示了在 $M=16$ 和 $M=32$ (此时 $y>0.1$)时 TDGA(RE+CTR)在 KP2 和 KP3 上均出现了较明显的振荡现象,而且 M 值越大振荡越剧烈.这表明, M 增大时,TDGA(RE+CTR)的稳定性有所降低.这些结果也再次验证了定理 2 的论述.因此,为了保证 TDGA(RE+CTR)具有较好的性能, M 应取较小的值,实验结果表明,当 $y<0.1$ 时,一般即可满足要求.

5 结束语

针对 TDGA 目前计算代价过高的问题,本文引入了等级熵多样性度量方法 RE 和分量热力学替换规则 CTR 以改善 TDGA 的计算效率.RE 方法在适应值空间中划分个体等级,能够以较小的计算成本度量种群中个体适应值的分散程度.CTR 规则将种群的自由能分派到其各个个体,有效地降低了热力学替换规则的复杂度.在 0-1 背包问题上的实验结果表明,RE 方法和 CTR 规则使 TDGA 一方面成功地借鉴了固体退火中能量与熵的竞争模式,系统地协调了 GA 中“选择压力”和“种群多样性”之间的冲突,另一方面极大地提高了其计算效率.

然而使用等级熵度量方法时需要确定比例因子常量 a 和等级数 K 等参数,因此,如何设计计算成本低且依赖参数少的多样性度量方法值得进一步研究.另外,将 TDGA(RE+CTR)应用于函数优化和多目标优化领域也是

我们下一步要做的工作.

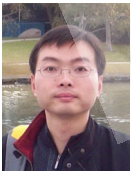
致谢 感谢武汉大学软件工程国家重点实验室演化计算讨论组的老师和同学对本文工作给予的建议和帮助.

References:

- [1] Su XH, Yang B, Wang YD. A genetic algorithm based on evolutionary stable strategy. *Journal of Software*, 2003,14(11): 1863–1868 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/14/1863.html>
- [2] Whitley D. The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. In: Schaffer J, ed. *Proc. of the Int'l Conf. on Genetic Algorithms*. San Mateo: Morgan Kaufmann Publishers, 1989. 116–121.
- [3] Vavak F, Forgarty TC. Comparison of steady state and generational genetic algorithms for use in nonstationary environments. In: Fukuda T, Furuhashi T, eds. *Proc. of the IEEE Conf. on Evolutionary Computation*. New York: IEEE Press, 1996. 192–195.
- [4] Sareni B, Krähenbühl L. Fitness sharing and niching methods revisited. *IEEE Trans. on Evolutionary Computation*, 1998,2(3): 97–106.
- [5] Li YX, Zou XF, Kang LS, Michalewicz Z. A new dynamical evolutionary algorithm based on statistical mechanics. *Journal of Computer Science and Technology*, 2003,18(3):361–368.
- [6] Mori N, Yoshida J, Tamaki H, Kita H, Nishikawa Y. A thermodynamical selection rule for the genetic algorithm. In: Fogel DB, ed. *Proc. of the IEEE Conf. on Evolutionary Computation*. New York: IEEE Press, 1995. 188–192.
- [7] Mori N, Kita H, Nishikawa Y. Adaptation to a changing environment by means of the feedback thermodynamical genetic algorithm. In: Eiben AE, *et al.*, eds. *Proc. of the IEEE Conf. on Parallel Problem Solving from Nature*. Berlin: Springer-Verlag, 1998. 149–158.
- [8] Mori N, Takeda M, Matsumoto K. A comparison study between genetic algorithms and bayesian optimize algorithms by novel indices. In: Rothlauf F, ed. *Proc. of the Genetic and Evolutionary Computation Conf.* New York: ACM Press, 2005. 1485–1492.
- [9] Kang LS. *Non-Numerical Parallel Algorithms (1)—Simulation Annealing Algorithms*. Beijing: Science Press, 1997 (in Chinese).
- [10] Maekawa K, Mori N, Tamaki H, Kita H, Nishikawa Y. A genetic solution for the traveling salesman problem by means of a thermodynamical selection rule. In: Fukuda T, Furuhashi T, eds. *Proc. of the IEEE Conf. on Evolutionary Computation*. New York: IEEE Press, 1996. 529–534.
- [11] Ding W, Hu T, Zhang HG. Multi-Objective optimization by a new dynamical evolutionary algorithm based on the information entropy. In: Zhao MS, Shi ZZ, eds. *Proc. of the Int'l. Conf. on Neural Networks and Brain*. New York: IEEE Press, 2005. 50–53.
- [12] Zheng BB, Li YX, Shen XJ, Zheng BJ. A new dynamic particle swarm optimizer. In: Wang TD, *et al.*, eds. *Proc. of the Int'l Conf. on Simulated Evolution and Learning*. Berlin: Springer-Verlag, 2006. 481–488.
- [13] Ying WQ, Li YX, Peng SJ, Wang WW. A steep thermodynamical selection rule for evolutionary algorithms. In: Shi Y, *et al.*, eds. *Proc. of the Int'l Conf. on Computational Science*. Berlin: Springer-Verlag, 2007. 997–1004.
- [14] Massachusetts Institute of Technology. GALib: A C++ library of genetic algorithm components. 1996. <http://lancet.mit.edu/ga/>
- [15] Martello S, Toth P. *Knapsack Problems: Algorithms and Computer Implementations*. New York: John Wiley & Sons, 1990.
- [16] Pisinger D. *Algorithms for knapsack problems* [Ph.D. Thesis]. Copenhagen: University of Copenhagen, 1995.

附中文参考文献:

- [1] 苏小红,杨博,王亚东.基于进化稳定策略的遗传算法.软件学报,2003,14(11):1863–1868. <http://www.jos.org.cn/1000-9825/14/1863.html>
- [9] 康立山.非数值并行算法(第1册)——模拟退火算法.北京:科学出版社,1997.



应伟勤(1979—),男,湖北仙桃人,博士生,主要研究领域为智能计算,语义计算.



SHEU Phillip C-Y(1956—),男,教授,IEEE Fellow,主要研究领域为语义计算,软件工程,复杂生物医学系统.



李元香(1962—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为智能计算,并行计算.