基于关系模型的进化算法收敛性分析与对比

黄 翰[1),2)] 林智勇[3)] 郝志峰[4)] 张宇山[5)] 李学强[1)]

[1)](华南理工大学软件学院 广州 510006)

[2)](南京大学计算机软件新技术国家重点实验室 南京 210093)

[3)](广东技术师范学院计算机科学学院 广州 510665)

[4)](广东工业大学计算机学院 广州 510006)

[5)](广东商学院数学与计算科学学院 广州 510320)

摘 要 研究建立了一种等态等价关系与强/弱态偏序关系模型,用于分析进化算法在收敛性上的等价性与可比性.基于吸收态Markov(马尔可夫)性,满足等态关系的进化算法具有等价的收敛性,从而在收敛性意义上实现了进化算法的等价类划分.在等态关系基础上,建立了弱态和强态的偏序关系,提出了一种对比进化算法收敛性的数学工具,在此基础上设计了更为强态的进化算法.文章运用所得理论分析了采用不同变异算子的(1+1)EA算法之间的关系,并用数值实验予以验证.文章提出的关系模型可以作为研究进化算法在收敛性上等价、对比和改进的一种理论基础.

关键词 进化算法;收敛性;等态关系;强/弱态关系

中图法分类号 TP301 **DOI号**:10.3724/SP.J.1016.2011.00801

Convergence Analysis and Comparison of Evolutionary Algorithms Based on Relation Model

HUANG Han[1),2)] LIN Zhi-Yong[3)] HAO Zhi-Feng[4)] ZHANG Yu-Shan[5)] LI Xue-Qiang[1)]

[1)](*School of Software Engineering, South China University of Technology, Guangzhou* 510006)

[2)](*The State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing* 210093)

[3)](*College of Computer Science, Guangdong Polytechnic Normal University, Guangzhou* 510665)

[4)](*School of Computer, Guangdong University of Technology, Guangzhou* 510006)

[5)](*School of Mathematics & Computational Science, Guangdong University of Business Studies, Guangzhou* 510320)

Abstract A relation-based model is proposed to analyze the equivalence and ordering of evolutionary algorithm (EA) in convergence. Based on the property of absorbing Markov chain, the EA of the same equivalence-in-status relation has the equal convergence, so that the partition of EAs can be implemented in convergence. According to equivalence-in-status relation, an ordering of stronger/weaker-in-status relation is proposed as a mathematical tool for comparing the convergence of evolutionary algorithm. Furthermore, a stronger EA is given to improve the convergence of EA. The relation of four (1+1)EAs with different mutation is analyzed by the proposed theorems as case studies with a proof of numerical results. The presented model and theorem can be a foundation to study the equivalence, ordering and improvement of evolutionary algorithm.

Keywords evolutionary algorithm; convergence; equivalence in status; stronger/weaker in status

---

收稿日期:2010-08-04;最终修改稿收到日期:2011-03-02.本课题得到国家自然科学基金(61003066,61070033,60873078)、教育部博士点基金(20090172120035)、中央高校基本科研业务费(2009ZM0052)和广东省高校优秀青年创新人才培育基金(LYM08074)、广东省自然科学基金项目(9251009001000005、10151601501000015)、广东省科技计划项目(2010B050400011,2008B080701005)资助.**黄 翰**,男,1980年生,博士,副教授,主要研究方向为进化计算方法的理论基础、进化计算方法的优化设计及其应用.E-mail:bssthh@163.com;hhan@scut.edu.cn.**林智勇**,男,1977年生,博士,副教授,主要研究兴趣包括算法设计与分析、机器学习和计算智能等.**郝志峰**,男,1968年生,博士,教授,博士生导师,主要研究领域为算法设计与分析、代数学及其应用等.**张宇山**,男,1975年生,博士研究生,讲师,主要研究方向为进化算法理论基础等.**李学强**,男,1983年生,博士研究生,主要研究方向为进化多目标优化算法等.

1 引 言

进化算法(Evolutionary Algorithms, EAs)是一类随机优化算法,其中包括遗传算法(GA)、进化策略(ES)和进化规划(EP)等.它们利用某些启发式规则进行优化,其思想根源来自于进化论.EAs研究近期被重新划分为4个分支:进化优化、进化学习、进化设计和进化算法的理论基础[1].收敛性分析是理论基础研究中的一项重要内容,它是评估进化算法正确性的一种重要手段.目前,关于EAs收敛性方面的研究结果已有不少,但多数是基于不可归约马氏链而得到的,并且是针对单种算法作收敛性分析,多种算法收敛性对比的理论研究工作则非常缺乏.

20世纪90年代,进化算法的理论基础研究开始兴起.由于GA的应用越来越广泛,并表现出原有经典方法所没有的优势,许多学者开始思考其性能优劣的理论解释,希望能得出有理论依据的结论,用于指导算法设计.Holland[2]的模式定理无疑是最早的理论基础.但是,模式定理仅适用于二进制编码,其结论难以运用到实际分析中,且存在着一些反例,因此,模式定理并不能充分地揭示GA的理论意义.Frantz[3]首先发现了GA欺骗问题(GA-Deceptive),它能使算法偏离最优解而早熟,进而让模式定理无效.Goldberg[4]和Bethke[5]运用Walsh函数和模式对GA进行了分析,并据此设计了最小骗问题的例子.遗憾的是,Walsh方法需要计算解空间中所有的点才能得到分析结果,因而无法用于指导实际的算法设计.基于二进制编码和单点交叉的积木块假设[6-7]是GA性能分析中的另一种重要观点:低阶、短距、高平均适应度的模式(积木块)在遗传算子作用下能生成高阶、长距、高平均适应度的模式,最终生成全局最优解.但由于只是一个假设、并没有理论证明,而且存在着反例,因此,积木块假设并没有得到普遍认可.

马尔可夫链(Markov)数学工具的引入让GA的理论基础研究取得了实质性的突破.Goldberg和Segrest[8]首先提出用有限齐次马氏链分析GA.此后,Eiben等人[9]用马氏链证明了保留最优个体(Elitist)的GA在概率意义上的全局收敛性.Rudolph[10]则用有限齐次马氏链证明了带有复制、交换和突变操作的标准GA不能收敛到全局最优解;据此,他认为标准GA不适应于静态函数优化问题[11],并建议改变复制策略以获得全局收敛.在此基础上,Bäck[12]和Muhlenbein[13]研究了全局收敛遗传算法的时间复杂性.

国内学者也在进化算法的收敛性方面做了不少关键性工作.徐宗本等人[14]证明了变异概率为0的多种情形下GA会过早收敛,并给出了增大变异概率等3个措施以保证新措施下的GA是全局收敛的.张讲社等人[15]给出了一种整体退火遗传算法,并证明了其全局收敛性.梁艳春等人[16]从理论上探讨了GA的收敛条件与收敛速度等问题,给出了等价遗传算法平均收敛速度的解析表达式.徐宗本等人[17]尝试运用鞅论研究GA的几乎必然强收敛性.周育人和闵华清等人[18]在变异算子为Gauss变异以及目标函数为连续函数的假设下,利用概率论的基本理论证明了多目标进化算法MOEAs的强收敛性.他们[19]还证明了球函数上的(μ,λ)ES的线性收敛性,并导出了相应的收敛因子和效率公式;依据效率公式,他们从理论上分析了(μ,λ)进化策略中参数μ和λ的最佳比值.王宇平等人分析了基于平滑技术全局优化进化算法[20-21]、求解TSP的量子遗传算法[21]和动态多目标优化进化算法[22]的收敛性.

多数研究结果[8-22]的局限性在于需要基于不可归约Markov性,即要求算法在无穷次迭代中能遍历问题的解空间.另外,已有研究结论多数是只分析单种进化算法的收敛性,较少考虑多种进化算法之间收敛性的对比.通过在收敛性方面建立等价关系和偏序关系数学模型,本文对生成与测试(generate-and-test)[1]方法框架的进化算法进行等价划分和对比研究,获得了无需基于不可归约Markov性的收敛性分析结论.

2 进化算法的等态关系模型

2.1 进化算法的随机过程模型

进化算法的种类很多,包括了遗传算法、进化规划与进化策略等.其中,遗传算法所采用的染色体编码方式有多种,包括二进制、实数、字符甚至是混合结构;进化规划和进化策略则采用实数编码.种群是一组染色体的集合,通常各种进化算法都是利用某种随机化的方法而获得初始种群.染色体的评估是根据给定的适值函数计算每个染色体的适应值.染色体的筛选主要是根据它们的适应值从当前种群中选择出若干染色体组成新的种群,供算法的下一轮迭代使用.在筛选方式上,遗传算法采用赌轮选择,

进化策略采用排序选择，进化规划则是采用竞争选择等. 为了生成新的染色体，遗传算法采取交叉和变异等算子，而进化规划和进化策略则仅采用变异算子. 综上所述，本文仅研究符合如下生成与测试(generate-and-test)[1]方法框架的进化算法.

> 1. 初始化，产生由 n 条染色体组成的初始种群 $\mathbb{P}_0$，$t=0$.
> 2. 对种群 $\mathbb{P}_t$ 中的染色体逐一进行评估.
> 3. 根据(父代)种群 $\mathbb{P}_t$ 产生一组染色体，形成(子代)种群 $\widetilde{\mathbb{P}}_t$.
> 4. 对 $\widetilde{\mathbb{P}}_t$ 中的染色体逐一进行评估.
> 5. 如果满足停机条件则输出最优解并终止算法.

图 1　本文研究的进化算法的基本流程

在进化算法中，染色体实质上是对所求问题的解的某种编码. 给定问题 $\mathcal{P}$ 和求解 $\mathcal{P}$ 的进化算法 $\mathcal{A}$，我们用 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$ 表示 $\mathcal{A}$ 在进化过程中第 t 代的种群. 对种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$ 中的第 i 条合法染色体 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(i)$，通过解码(形式上记为 $\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(i))$)可以得到 $\mathcal{P}$ 的一个可行解(为简单起见，我们总是假定染色体是合法的. 事实上，当染色体不合法时，可以通过某些方法进行修补而令其合法[23-24]). 进而，通过解码，种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$ 将与问题 $\mathcal{P}$ 的一组解相对应，形式上我们可用向量的形式表示如下

$$\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}})=(\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(1)),\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(2)),\cdots,\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(n))).$$

显然，这样的解向量和算法 $\mathcal{A}$ 的执行策略密切相关，并随着种群的进化而变化. 若将这样的向量视为一个取值于某个状态空间的随机向量，则我们可以获得一个随机过程，该过程刻画了算法 $\mathcal{A}$ 求解 $\mathcal{P}$ 的进化过程. 具体地，我们作如下定义.

定义 1(n 阶状态空间).　给定问题 $\mathcal{P}$，设其可行解集为 $\mathcal{FS}^{\mathcal{P}}$. 定义 $\Omega_n^{\mathcal{P}} \triangleq \underbrace{\mathcal{FS}^{\mathcal{P}}\times\mathcal{FS}^{\mathcal{P}}\cdots\times\mathcal{FS}^{\mathcal{P}}}_{n}$ 为由 $\mathcal{P}$ 的可行解集生成的 n 阶状态空间，其中×表示集合的笛卡尔积.

定义 2(算法对应的随机过程).　设 $\mathcal{A}$ 是符合图 1 生成测试框架的用于求解问题 $\mathcal{P}$ 的进化算法，令 $\xi_t^{\mathcal{A},\mathcal{P}}=\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}})$，即与 $\mathcal{A}$ 的第 t 代种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$(通过某种解码)相对应的问题 $\mathcal{P}$ 的解向量，称 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 为算法 $\mathcal{A}$ 求解 $\mathcal{P}$ 的随机过程.

显然，当 $\mathcal{A}$ 的种群规模为 n 时，我们总有 $\xi_t^{\mathcal{A},\mathcal{P}}\in\Omega_n^{\mathcal{P}}(t=0,1,2,\cdots)$，因此，$\Omega_n^{\mathcal{P}}$ 可以作为随机过程 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 的状态空间. 进一步地，我们首先指出随机过程 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 的 Markov 性，亦即有以下引理.

引理 1.　设算法 $\mathcal{A}$ 的种群规模为 n，则在状态空间 $\Omega_n^{\mathcal{P}}$ 中 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 具有 Markov 性.

证明.　根据图 1 的算法流程可知，$\mathcal{A}$ 的第 $t+1$ 代种群 $\mathbb{P}_{t+1}^{\mathcal{A},\mathcal{P}}$ 是从第 t 代种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$ 及其生成的子代种群 $\widetilde{\mathbb{P}}_t^{\mathcal{A},\mathcal{P}}$ 中筛选出来的，所以，对于 $\forall\widetilde{\Omega}\subseteq\Omega_n^{\mathcal{P}}$ 和 $t=0,1,2,\cdots$，有

$$P\{\xi_{t+1}^{\mathcal{A},\mathcal{P}}\in\widetilde{\Omega}\,|\,\xi_0^{\mathcal{A},\mathcal{P}},\cdots,\xi_t^{\mathcal{A},\mathcal{P}}\}=P\{\xi_{t+1}^{\mathcal{A},\mathcal{P}}\in\widetilde{\Omega}\,|\,\xi_t^{\mathcal{A},\mathcal{P}}\}.$$

这表明，$\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 具有 Markov 性.　　证毕.

引理 1 的结论早在学者们[10]研究遗传算法的收敛性时就提出，这里套用来描述进化算法(图 1)具有 Markov 性. 进化算法的主要目的是为了搜索问题的最优解，这意味着，我们需要如下关于最优状态空间的概念.

定义 3(最优状态空间).　设 $\Omega_n^{\mathcal{P}}$ 是 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 的状态空间，称 $\Omega_n^{\mathcal{P},opt}\subset\Omega_n^{\mathcal{P}}$ 为相应的最优状态空间，若对于 $\forall V\in\Omega_n^{\mathcal{P},opt}$ 其中包含一个分量是问题 P 的最优解.

显然，当 $P\{\xi_t^{\mathcal{A},\mathcal{P}}\in\Omega_n^{\mathcal{P},opt}\}=1$ 时，进化算法 $\mathcal{A}$ 能在有限时间内(第 t 次迭代)找到问题的最优解；而当 $\lim\limits_{t\to+\infty}P\{\xi_t^{\mathcal{A},\mathcal{P}}\in\Omega_n^{\mathcal{P},opt}\}=1$ 时，进化算法可以在迭代时间趋于无穷时找到问题的最优解. 在此基础上，我们给出进化算法收敛的概念.

定义 4(算法收敛).　设 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 为算法 $\mathcal{A}$ 求解 $\mathcal{P}$ 的随机过程，相应的最优状态空间为 $\Omega_n^{\mathcal{P},opt}$. 称进化算法 $\mathcal{A}$ 是收敛的，若 $\lim\limits_{t\to+\infty}P\{\xi_t^{\mathcal{A},\mathcal{P}}\in\Omega_n^{\mathcal{P},opt}\}=1$ 成立.

不难发现，在前面的描述中，我们通过使用上标来强调随机过程 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 不仅与问题 $\mathcal{P}$ 有关，还与算法 $\mathcal{A}$ 有关. 为了便于研究，我们作如下进一步的假设. 首先，假定所要求解的问题 $\mathcal{P}$ 是一个离散优化问题，它的可行解集为 $\mathcal{FS}^{\mathcal{P}}=\{s_i\,|\,i=1,2,\cdots\}$. 其次，假定 $\mathcal{EAS}=\{\mathcal{A}_1,\mathcal{A}_2,\cdots,\mathcal{A}_k\}$ 是用于求解问题 $\mathcal{P}$ 的 k 个不同进化算法的集合，其中每个进化算法在实现时对应的种群规模不超过 N. 如前所述，随机过程 $\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 的状态空间 $\Omega_n^{\mathcal{P}}$ 和算法种群规模有关，而不同的算法可能选择不同的种群规模. 因此，为了能对多个算法进行统一研究，我们需要一个扩展的状态空间，它能适用于所有种群规模不超过 N 的进化算法. 具体地，给定问题 $\mathcal{P}$，我们定义扩展的状态空间如下：

$$\Omega^{\mathcal{P}}\triangleq\Omega_1^{\mathcal{P}}\cup\Omega_2^{\mathcal{P}}\cup\cdots\cup\Omega_N^{\mathcal{P}}.$$

显然，经过这样的处理，各个算法对应的随机过程不仅具有相同的状态空间 $\Omega_n^{\mathcal{P}}$，也具有相同的最优状态空间，形式上可记为 $\Omega_n^{\mathcal{P},opt}$. 此外，由于状态空间 $\Omega^{\mathcal{P}}$ 是离散的，根据引理 1 可知，$\mathcal{EAS}$ 中的每个进化算法对应一个 Markov 链. 本文接下来的内容就是基于这些具有相同状态空间的 Markov 链对相应的进化算法进行比较研究. 给定问题 $\mathcal{P}$，为了便于描述，

在后文中，我们将去掉符号上标处的 $\mathcal{P}$（譬如，$\Omega^{\mathcal{P}}$ 和 $\Omega^{\mathcal{P},opt}$ 将分别简记为 Ω 和 Ω^{opt}）.

2.2 进化算法的等态关系模型

本节将在笛卡尔积 $\mathcal{EAS}\times\mathcal{EAS}$ 上构造一个等态关系，为此，先引入如下两个辅助定义.

定义 5（非衰退序列）. 实序列 $\{\sigma_t\}_{t=0}^{+\infty}$ 是非衰退的，如果 $\sigma_t>0$ 且 $\prod\limits_{t=0}^{+\infty}(1-\sigma_t)=0$.

定义 6（α 可达状态集）. 设进化算法 $\mathcal{A}$ 对应的 Markov 链为 $\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$，对给定的非衰退序列 $\alpha=\{\alpha_t\}_{t=0}^{+\infty}$，称

$$\Xi_t^{\mathcal{A}}(\alpha)=\{\omega\,|\,P(\xi_t^{\mathcal{A}}=\omega\,|\,\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt})\geq\alpha_t,\omega\in\Omega\}$$

为 $\mathcal{A}$ 在第 t 次迭代的 α 可达状态集（$t=1,2,\cdots$）.

$\Xi_t^{\mathcal{A}}(\alpha)$ 包含的状态是进化算法 $\mathcal{A}$ 可以在第 t 次迭代以非 0 概率达到的，而且这个概率不衰退为 0.

定义 7（等态关系）. 给定进化算法集合 $\mathcal{EAS}$，对 $\mathcal{EAS}$ 中任意两个算法 $\mathcal{A}$ 和 $\mathcal{B}$（对应的随机过程分别为 $\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$ 和 $\{\xi_t^{\mathcal{B}}\}_{t=1}^{+\infty}$），若对于任意的非衰退序列 α 都存在另一个非衰退序列 β，使得 $\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}=\Xi_t^{\mathcal{B}}(\beta)\cap\Omega^{opt}$（$t=1,2,\cdots$）成立，反之亦然，则称进化算法 $\mathcal{A}$ 等态于 $\mathcal{B}$，记为 $\mathcal{A}\cong\mathcal{B}$，而 $\cong$ 称为 $\mathcal{EAS}\times\mathcal{EAS}$ 上的一个等态关系.

如果两个进化算法满足等态关系，则两个算法在每次迭代时，以非衰退概率产生的种群集合与最优状态集合的交集是一样的. 直观解释：算法 $\mathcal{A}$ 能达到的最优状态与算法 $\mathcal{B}$ 达到的是一致的. 虽然如此，算法 $\mathcal{A}$ 和 $\mathcal{B}$ 达到最优状态所花的时间可能不一样，即等态关系的性质与收敛速度无关. 下面证明等态关系是 $\mathcal{EAS}\times\mathcal{EAS}$ 上的等价关系.

定理 1. 等态关系 $\cong$ 是 $\mathcal{EAS}\times\mathcal{EAS}$ 的一个等价关系.

证明. （1）自反性. 显然，对任意 $\mathcal{A}\in\mathcal{EAS}$ 有：$\mathcal{A}\cong\mathcal{A}$.（2）对称性. 对任意 $\mathcal{A},\mathcal{B}\in\mathcal{EAS}$，根据定义 7 有 $\mathcal{A}\cong\mathcal{B}\Leftrightarrow\mathcal{B}\cong\mathcal{A}$.（3）传递性. 对任意 $\mathcal{A},\mathcal{B},\mathcal{C}\in\mathcal{EAS}$，当 $\mathcal{A}\cong\mathcal{B}$ 时，亦即对于任意非衰退序列 α 都存在一个非衰退序列 β 使得 $\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}=\Xi_t^{\mathcal{B}}(\beta)\cap\Omega^{opt}$（$t=1,2,\cdots$）成立. 同理，$\mathcal{B}\cong\mathcal{C}$ 意味着存在一个非衰退序列 γ 使得 $\Xi_t^{\mathcal{B}}(\beta)\cap\Omega^{opt}=\Xi_t^{\mathcal{C}}(\gamma)\cap\Omega^{opt}$（$t=1,2,\cdots$）成立. 由此可知，$\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}=\Xi_t^{\mathcal{C}}(\gamma)\cap\Omega^{opt}$（$t=1,2,\cdots$）也成立，即 $\mathcal{A}\cong\mathcal{B},\mathcal{B}\cong\mathcal{C}\Rightarrow\mathcal{A}\cong\mathcal{C}$. 证毕.

既然等态关系 $\cong$ 也是 $\mathcal{EAS}\times\mathcal{EAS}$ 上的一个等价关系，那么，我们可以据此对 $\mathcal{EAS}$ 中的算法进行分类，亦即得到如下等态类的概念.

定义 8（等态类）. 给定进化算法集合 $\mathcal{EAS}$，$\cong$ 是 $\mathcal{EAS}\times\mathcal{EAS}$ 上的等态关系，则称集合 $[\mathcal{A}]_{\cong}=\{\tilde{\mathcal{A}}\,|\,\tilde{\mathcal{A}}\in\mathcal{EAS}\wedge\tilde{\mathcal{A}}\cong\mathcal{A}\}$ 为进化算法 $\mathcal{A}\in\mathcal{EAS}$ 在 $\mathcal{EAS}$ 上诱导的等态类.

第 3 节将给出等态类和进化算法收敛性的关系.

3 等态关系与进化算法收敛性

3.1 基于等态关系的进化算法收敛性等价分析

等态关系实现了对进化算法的一个等价类划分，属于同类的进化算法有可能搜索到最优状态集（包含最优解的种群状态集）是一致的. 如果允许足够多次实验且迭代次数无限，同一等态类中的进化算法都能达到一致的最优状态. 下面证明等态类中进化算法在收敛性上的这一特点.

首先，需要引入吸收态 Markov 链的定义[8].

定义 9（吸收态 Markov 链）. 给定 Markov 链 $\{\xi_t\}_{t=0}^{+\infty}$，相应的最优状态空间为 Ω^{opt}，称 $\{\xi_t\}_{t=0}^{+\infty}$ 是吸收态的，若有 $P\{\xi_{t+1}\notin\Omega^{opt}\mid\xi_t\in\Omega^{opt}\}=0$（$t=1,2,\cdots$）.

吸收态 Markov 链一旦到达最优状态空间，就会吸附在其中永远不会出来. Eiben 等人[9] 用 Markov 链证明了保留最优个体（Elitist）GA 的概率性全局收敛之后，许多进化算法的设计都采用 Elitist 保留策略；这些算法都可以建模为吸收态 Markov 链. 定义 9 的模型适合绝大多数的进化算法.

结合非衰退实序列，满足吸收态 Markov 性的进化算法 $\mathcal{A}$ 有引理 2 所描述的性质.

引理 2. 给定算法 $\mathcal{A}$ 对应的吸收态 Markov 链 $\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$，若存在 $t'\geq0$ 使得当 $t>t'$ 时有 $P\{\xi_t^{\mathcal{A}}\in\Omega^{opt}\,|\,\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}\geq\alpha_t$ 且 $\{\alpha_t\}_{t=1}^{+\infty}$ 为非衰退的，则算法 $\mathcal{A}$ 收敛.

证明. 若存在 $t'\geq0$ 使得当 $t>t'$ 时有

$$P\{\xi_t^{\mathcal{A}}\in\Omega^{opt}\,|\,\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}\geq\alpha_t,$$

则也有：$P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\,|\,\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}\leq1-\alpha_t$.

若记 $\bar{P}_t^{\mathcal{A}}=\prod\limits_{i=0}^{t-1}P\{\xi_{i+1}^{\mathcal{A}}\notin\Omega^{opt}\,|\,\xi_i^{\mathcal{A}}\notin\Omega^{opt}\}$，则有

$$\lim_{t\to\infty}\bar{P}_t^{\mathcal{A}}\leq\prod_{t=0}^{+\infty}(1-\alpha_t).$$

又 $\prod\limits_{t=0}^{+\infty}(1-\alpha_t)=0$，所以 $\lim\limits_{t\to\infty}\bar{P}_t^{\mathcal{A}}=0$. 这表明，$\mathcal{A}$ 在迭代时间趋于无穷时必能到达最优状态空间.

因为 $\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$ 是吸收态 Markov 链，根据定义 9 有 $P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\,|\,\xi_{t-1}^{\mathcal{A}}\in\Omega^{opt}\}=0$（$t=1,2,\cdots$），亦即 $P\{\xi_t^{\mathcal{A}}\in\Omega^{opt}\,|\,\xi_{t-1}^{\mathcal{A}}\in\Omega^{opt}\}=1$（$t=1,2,\cdots$）. 据此，利用全概率公式，可得

$$
\begin{aligned}
&P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\}=\\
&\quad P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\mid\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}P\{\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}+\\
&\quad P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\mid\xi_{t-1}^{\mathcal{A}}\in\Omega^{opt}\}P\{\xi_{t-1}^{\mathcal{A}}\in\Omega^{opt}\}=\\
&\quad P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\mid\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}P\{\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt}\}=\\
&\quad P\{\xi_0^{\mathcal{A}}\notin\Omega^{opt}\}\prod_{i=0}^{t-1}P\{\xi_{i+1}^{\mathcal{A}}\notin\Omega^{opt}\mid\xi_i^{\mathcal{A}}\notin\Omega^{opt}\}=\\
&\quad P\{\xi_0^{\mathcal{A}}\notin\Omega^{opt}\}\bar{P}_t^{\mathcal{A}}
\end{aligned}
$$

注意到$\lim_{t\to\infty}\bar{P}_t^{\mathcal{A}}=0$，利用上式，有

$$\lim_{t\to\infty}P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\}=0.$$

进而，有

$$\lim_{t\to\infty}P\{\xi_t^{\mathcal{A}}\in\Omega^{opt}\}=1-\lim_{t\to\infty}P\{\xi_t^{\mathcal{A}}\notin\Omega^{opt}\}=1.$$

证毕.

引理2源自早期学者研究进化规划算法收敛性时的类似结论[7,23-24]，这里重新阐述主要为了说明进化算法收敛的一个充分条件.利用引理2，我们可得到下面定理2，它揭示了等态关系模型与进化算法收敛性之间的联系.

定理2.　给定进化算法集合$\mathcal{EAS}$，$[\mathcal{A}]_{\cong}$是由$\mathcal{A}\in\mathcal{EAS}$在$\mathcal{EAS}$上诱导的等态类.假设有某一$\bar{\mathcal{A}}\in[\mathcal{A}]_{\cong}$是收敛的，对任意$\tilde{\mathcal{A}}\in[\mathcal{A}]_{\cong}$，若$\tilde{\mathcal{A}}$对应的随机过程$\{\xi_t^{\tilde{\mathcal{A}}}\}_{t=0}^{+\infty}$为吸收态Markov链，则$\tilde{\mathcal{A}}$也收敛.

证明.　(1)首先指出：对于$\bar{\mathcal{A}}$，存在$t'>0$和非衰退序列α使得当$t>t'$时有$\Xi_t^{\bar{\mathcal{A}}}(\alpha)\cap\Omega^{opt}\neq\varnothing$.若不然，则对于任意$t'>0$和非衰退序列$\alpha$，都存在一个$t_0>t'$使得$\Xi_{t_0}^{\bar{\mathcal{A}}}(\alpha)\cap\Omega^{opt}=\varnothing$，即有$P\{\xi_{t_0}^{\bar{\mathcal{A}}}\in\Omega^{opt}\}=0$，据此，$\lim_{t\to\infty}P\{\xi_t^{\bar{\mathcal{A}}}(P)\in\Omega^{opt}\}=1$不能成立，并导致与$\bar{\mathcal{A}}\in[\mathcal{A}]_{\cong}$的收敛事实矛盾.

(2)其次，对任意$\tilde{\mathcal{A}}\in[\mathcal{A}]_{\cong}$，显然有$\tilde{\mathcal{A}}\cong\bar{\mathcal{A}}$.根据定义7，对于任意的非衰退序列$\alpha$都存在一个非衰退序列$\beta$使得：

$$\Xi_t^{\bar{\mathcal{A}}}(\alpha)\cap\Omega^{opt}=\Xi_t^{\tilde{\mathcal{A}}}(\beta)\cap\Omega^{opt}\neq\varnothing\,(t>t').$$

再根据定义6有$P\{\Xi_t^{\tilde{\mathcal{A}}}(\alpha)\cap\Omega^{opt}\neq\varnothing\}\geq\alpha_t$，所以，$P\{\xi_t^{\tilde{\mathcal{A}}}\in\Omega^{opt}|\xi_{t-1}^{\tilde{\mathcal{A}}}\notin\Omega^{opt}\}\geq\beta_t\,(t>t')$而且$\prod_{t=0}^{+\infty}(1-\beta_t)=0$.由于$\tilde{\mathcal{A}}$对应的随机过程$\{\xi_t^{\tilde{\mathcal{A}}}\}_{t=0}^{+\infty}$为吸收态Markov链，根据引理2可知$\tilde{\mathcal{A}}$收敛.　证毕.

定理2说明了在等态类中，具有吸收态Markov链性质的进化算法具有共同收敛的性质.这也说明了：只要证明了等态类中一个进化算法的收敛性，便得到同类其它进化算法(具有吸收态Markov链性质)的收敛性.进化算法一般都具有吸收态Markov链的性质，定理2反映了一个事实：如果要分析一个进化算法(满足吸收态Markov链性质)是否收敛，可以通过研究其所在等态类中的其它进化算法的收敛性达到目的.推论1从另一角度反映了这一事实.

推论1.　给定进化算法集合$\mathcal{EAS}$，$[\mathcal{A}]_{\cong}$是由$\mathcal{A}\in\mathcal{EAS}$在$\mathcal{EAS}$上诱导的等态类，若存在$\tilde{\mathcal{A}}\in[\mathcal{A}]_{\cong}$对应的$\{\xi_t^{\tilde{\mathcal{A}}}\}_{t=0}^{+\infty}$为吸收态Markov链且$\tilde{\mathcal{A}}$不收敛，则任意$\bar{\mathcal{A}}\in[\mathcal{A}]_{\cong}$也不收敛.

证明.　根据定理2，用反证法容易得证.

推论1体现了具有吸收态Markov链性质的进化算法与等态类中其它算法的另一个有趣关系：若存在一个进化算法(满足吸收态Markov性)不收敛，则该算法所属等态类中的所有进化算法均不收敛.也就是说，要证明等态类中的进化算法都不收敛，只需要证明类中一个具有吸收态Markov性的进化算法不收敛即可.在实际中，只要采用精英保留等改进，便可以使进化算法具有吸收态Markov链的性质，所以研究对象不难找到.

3.2　基于强/弱态关系的进化算法收敛性对比

基于等态关系，可以给出相应的偏序关系，用于进化算法收敛的对比.

定义10(强/弱态关系).　给定进化算法集合$\mathcal{EAS}$，对$\mathcal{EAS}$中任意两个算法$\mathcal{A}$和$\mathcal{B}$(对应的随机过程分别为$\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$和$\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$)，若对于任意的非衰退序列$\alpha$都存在另一个非衰退序列$\beta$，使得，$\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}\subseteq\Xi_t^{\mathcal{B}}(\beta)\cap\Omega^{opt}\,(t=1,2,\cdots)$成立，则称$\mathcal{A}$弱态于$\mathcal{B}$，记为$\mathcal{A}\prec\mathcal{B}$，或$\mathcal{B}$强态于$\mathcal{A}$，记为$\mathcal{B}\succ\mathcal{A}$.

强/弱态关系的直观解释：若$\mathcal{A}\prec\mathcal{B}$，则算法$\mathcal{A}$能找到的最优解，算法$\mathcal{B}$也能找到；反之，算法$\mathcal{B}$能找到的最优解，算法$\mathcal{A}$不一定能找到.由定义10，$\prec$满足反自反、非对称且传递，所以，强/弱态是一种偏序关系.下面的定理和推论刻画了满足强/弱态关系的两个进化算法在收敛性方面的关系.

定理3.　给定进化算法集合$\mathcal{EAS}$，对$\mathcal{EAS}$中任意两个算法$\mathcal{A}$和$\mathcal{B}$，若$\mathcal{A}\prec\mathcal{B}$且$\mathcal{A}$收敛，而$\mathcal{B}$对应的$\{\xi_t^{\mathcal{B}}\}_{t=1}^{+\infty}$为吸收态Markov链，则$\mathcal{B}$也收敛.

证明.　(1)首先，类似定理2证明的第(1)部分，可证得，对于$\mathcal{A}$存在$t'>0$和非衰退序列α，当$t>t'$时，有$\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}\neq\varnothing$.

(2)其次，根据定义10由$\mathcal{A}\prec\mathcal{B}$可知，当$t>t'$时$\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}\subseteq\Xi_t^{\mathcal{B}}(\beta)\cap\Omega^{opt}$，其中$\alpha$和$\beta$均为非衰退序列.再根据定义7，可知$P\{\Xi_t^{\mathcal{B}}(\beta)\cap\Omega^{opt}\neq\varnothing\}\geq\beta_t$.因此，当$t>t'$时，有

$$P\{\xi_t^{\mathcal{B}}\in\Omega^{opt}\mid\xi_{t-1}^{\mathcal{B}}\notin\Omega^{opt}\}\geq\beta_t.$$

注意到$\{\xi_t^{\mathcal{B}}\}_{t=1}^{+\infty}$为吸收态 Markov 链，根据引理 2，可知 $\mathcal{B}$ 是收敛的. 证毕.

定理 3 的结论也表明了：只要证明了一个进化算法的收敛性，便得到强态于该算法的其它具有吸收态 Markov 链性质的进化算法收敛性. 与定理 2 类似，定理 3 的一个直接作用在于：如果难以分析一个吸收态过程进化算法是否收敛，可以转而研究弱态于该算法其它算法的收敛性，以达到研究目的. 下面的推论 2 则给出了另一个方向的结论.

推论 2. 给定进化算法集合 $\mathcal{EAS}$，若 $\mathcal{A}\in\mathcal{EAS}$ 对应的$\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$为吸收态 Markov 链且 $\mathcal{A}$ 不收敛，则 $\mathcal{EAS}$ 中弱态于 $\mathcal{A}$ 的算法皆不收敛.

证明. 根据定理 3，用反证法容易得证.

推论 2 表明：具有吸收态 Markov 链性质的进化算法不收敛，则导致弱态于该算法的其它算法不可能收敛. 也就是说，要证明某个进化算法不收敛，只需要证明链中强态于该算法的一个具有吸收态 Markov 性的进化算法不收敛即可.

3.3 基于等态关系的进化算法收敛判别定理

这一小节基于等态关系模型给出进化算法的收敛判别定理，并相应提出进化算法收敛性改进的一个基本思想.

推论 3. 给定算法 $\mathcal{A}$ 对应的吸收态 Markov 链$\{\xi_t^{\mathcal{A}}\}_{t=1}^{+\infty}$，若存在 $t'\geq 0$ 和非衰退序列 α 使得当 $t>t'$ 时 $\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}\neq\varnothing$ 成立，则 $\mathcal{A}$ 收敛.

证明. 根据定义 5，$\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}\neq\varnothing\,(t>t')$ 成立意味着当 $t>t'$ 时必然存在某一状态 $\omega_t\in\Omega^{opt}$ 使得 $P(\xi_t^{\mathcal{A}}=\omega_t\,|\,\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt})\geq\alpha_t$，亦即，$P(\xi_t^{\mathcal{A}}\in\Omega_t^{opt}\,|\,\xi_{t-1}^{\mathcal{A}}\notin\Omega^{opt})\geq\alpha_t$. 注意到 α 为非衰退序列，即 $\prod_{t=0}^{+\infty}(1-\alpha_t)=0$. 因此，由引理 2 可知 $\mathcal{A}$ 是收敛的. 证毕.

推论 3 给出了一个进化算法收敛的判定原则. 注意，这里的证明并不要求 Markov 链（过程）的状态转移矩阵是不可归约的. 推论 3 条件比以往学者提出的收敛性证明条件[8-22,25]要弱；以往收敛分析条件可以简单表示为：存在 $t'\geq 0$ 和非衰退序列 α，当 $t>t'$ 时，$\Xi_t^{\mathcal{A}}(\alpha)=\Omega^{opt}$. 这个条件要求进化算法对应的 Markov 链在迭代时间趋于无穷时至少一次遍历整个状态空间（不可归约性），显然比推论 3 条件要苛刻. 因此，本节得到了一种新型的收敛性判别条件：满足吸收态 Markov 链性的进化算法能否收敛，关键取决于算法产生的可达状态集与最优状态集是否从某次迭代开始都能有交集，亦即，存在 $t'\geq 0$ 和非衰退序列 α 使得当 $t>t'$ 时有 $\Xi_t^{\mathcal{A}}(\alpha)\cap\Omega^{opt}\neq\varnothing$.

然而，在实际算法设计中，最优状态集 Ω^{opt} 一般是未知的. 所以，对于满足吸收态 Markov 链（过程）的进化算法，直接的改进方案就是设计一个强态于原算法的进化算法. 具体分析如下：

给定两个进化算法 $\mathcal{A}$ 和 $\mathcal{B}$，$\mathcal{B}$ 是 $\mathcal{A}$ 的修改版本，满足：$\mathcal{B}\succ\mathcal{A}$ 而且对应的随机过程具有吸收态 Markov 性. 当存在序列 α 满足定义 5，存在 $V\in\Omega^{opt}\wedge V\in\Xi_t^{\mathcal{A}}(\alpha)\,(t=1,2,\cdots)$ 时，由 $\mathcal{B}\succ\mathcal{A}$ 的性质有 $V\in\Xi_t^{\mathcal{B}}(\beta)$，即此时即可以认为 $\mathcal{B}$ 在收敛方面不亚于 $\mathcal{A}$.

不仅如此，当 $\forall V\in\Omega^{opt}\wedge V\notin\Xi_t^{\mathcal{A}}(\alpha)$ 且 $\exists\tilde{V}\in\Omega^{opt}\wedge\tilde{V}\in\Xi_t^{\mathcal{B}}(\beta)\,(t=1,2,\cdots)$ 时，$\mathcal{B}$ 的收敛性能优于 $\mathcal{A}$.

当然，因为这种基于强态关系的改进可能增加计算代价，所以，在设计更强态进化算法进行收敛性改进时，需要考虑所需增加的计算量，如此才能得到最有效的改进效果. 计算量的分析属于进化算法的另一理论主题——计算时间复杂性研究，我们在其它研究[26-28]中进行了讨论，本文分析主要针对收敛性.

4 满足等态和强/弱态关系的进化算法举例

4.1 采用不同变异算子的 4 个(1+1)EA 算法

本节根据经典的进化算法设计，给出满足等态和强/弱态关系的算法例子.

(1+1)EA 算法符合图 1 框架，算法有两个特点：种群规模为 1，变异为 1 位取反，无交叉算子. 因为其算法结构简单，(1+1)EA 算法的分析曾经出现在多位学者的理论研究文献中[23-24,29-30]，以极大化问题为例，算法框架如图 2 所示.

1. 初始化. 随机产生一个 0-1 串染色体. 每个染色体记为$(b_1,b_2,\cdots,b_m)$，$b_i\in\{0,1\}$，$i=1,2,\cdots,m$.
2. 选择. 如果子代染色体在适应值上优于父代染色体，则取代之成为当前染色体；否则保留原父代染色体. 如果停机条件满足则输出最优解.
3. 变异. 通过变异算子产生一个子代染色体. 跳转至步 2 选择.

图 2 (1+1)EA 算法流程

本小节研究 4 个具有不同变异算子的(1+1)EA：

(1) EA-I 算法在变异上采用逐位取反，即每一个 b_i 都有 1/2 的概率取反.

(2) EA-II 算法同样采用逐位取反，对每一个 b_i 有 $1/n$ 的概率取反.

(3) EA-III 算法则仅随机选择一个 b_i 取反.

(4) EA-IV 的变异采用插入变异，即随机选择一个 b_i 将其插入另外一个随机选择的位置 b_j 之前：

$$(b_1,\cdots,b_j,\cdots,b_i,\cdots,b_n)\rightarrow(b_1,\cdots,b_i,b_j,\cdots,b_n).$$

4 个算法在其它方面的设计一致，4 种变异算子的设计可以参考文献[23-24]. 以 4 种(1+1)EA 算法为例是因为其是进化算法理论研究的经典对象，并不代表研究提出的等态关系模型仅适用于(1+1)EA 算法. 理论上，只要满足图 1 生成与测试框架的进化算法皆可以用等态关系模型进行分析. 限于篇幅，文章不再列举其它例子.

4.2　满足等态关系的进化算法举例

例 1.　EA-I 与 EA-II 的关系.

可以证明 EA-I 等态于 EA-II. 由于两种算法采用相同的初始化设置，所以，基于同样的初始状态时，两种算法可达的状态集也是相等的. 从选择策略上，EA-I 和 EA-II 对应的随机过程都是吸收态 Markov 链. 命题 1 给出了这一结论.

命题 1.　EA-I 等态于 EA-II，即 EA-I≌EA-II.

证明.　假设进化算法 EA-I 和 EA-II 对应的随机过程分别为 $\{\xi_t^{\mathcal{A}_1}\}_{t=0}^{+\infty}$ 和 $\{\xi_t^{\mathcal{A}_2}\}_{t=0}^{+\infty}$. 因为 EA-I 和 EA-II 都是采用 0-1 编码染色体 $(b_1,b_2,\cdots,b_m)$，所以 $\xi_t^{\mathcal{A}_1}$ 和 $\xi_t^{\mathcal{A}_2}$ 都属于离散状态空间 $\widetilde{\Omega}$，$\widetilde{\Omega}$ 中每一个元素都是 m 维 0−1 串. 记 $\widetilde{\Omega}^{opt}$ 是最优状态空间，对于 $\forall V\in\widetilde{\Omega}^{opt}$ 和 $t=0,1,\cdots$，假设 V 与 $\xi_t^{\mathcal{A}_1}$ 有 m_1 个位置不同，V 与 $\xi_t^{\mathcal{A}_2}$ 有 m_2 个位置不同. 所以，根据两个算法的变异算子，可以计算：

$$P\{\xi_{t+1}^{\mathcal{A}_1}=V\mid\xi_t^{\mathcal{A}_1}\notin\widetilde{\Omega}^{opt}\}=\begin{pmatrix}m\\m_1\end{pmatrix}^{-1}\cdot\left(\frac{1}{2}\right)^{m_1}\cdot\left(\frac{1}{2}\right)^{m-m_1}=\alpha_t^{(1)}\tag{1}$$

$$P\{\xi_{t+1}^{\mathcal{A}_2}=V\mid\xi_t^{\mathcal{A}_2}\notin\widetilde{\Omega}^{opt}\}=\begin{pmatrix}m\\m_2\end{pmatrix}^{-1}\cdot\left(\frac{1}{n}\right)^{m_2}\cdot\left(1-\frac{1}{n}\right)^{m-m_2}=\alpha_t^{(2)}\tag{2}$$

易证 $\{\alpha_t^{(1)}\}_{t=0}^{+\infty}$，$\{\alpha_t^{(2)}\}_{t=0}^{+\infty}$ 满足非衰退性，下面证明两个算法的等态关系.

根据定义 6，由于 V 的任意性，$\Xi_t^{\mathcal{A}_1}(\alpha^{(1)})\cap\widetilde{\Omega}^{opt}\neq\varnothing$ 和 $\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})\cap\widetilde{\Omega}^{opt}\neq\varnothing$ 成立. 对于任意非衰退实序列 $\beta^{(1)}=\{\beta_t^{(1)}\}$ 满足 $\beta_t^{(1)}>\alpha_t^{(1)}$，存在非衰退实序列 $\beta^{(2)}=\{\beta_t^{(2)}\}$ 满足 $\beta_t^{(2)}>\alpha_t^{(2)}$，使得 $\Xi_t^{\mathcal{A}_1}(\beta^{(1)})\cap\widetilde{\Omega}^{opt}=\Xi_t^{\mathcal{A}_2}(\beta^{(2)})\cap\widetilde{\Omega}^{opt}=\varnothing(t=0,1,\cdots)$ 成立；类似地，对任意 $\gamma^{(1)}=\{\gamma_t^{(1)}\}$ 满足 $\gamma_t^{(1)}\le\alpha_t^{(1)}$，存在 $\gamma^{(2)}=\{\gamma_t^{(2)}\}$ 满足 $\gamma_t^{(2)}\le\alpha_t^{(2)}$，使得 $\Xi_t^{\mathcal{A}_1}(\gamma^{(1)})\cap\widetilde{\Omega}^{opt}=\Xi_t^{\mathcal{A}_2}(\gamma^{(2)})\cap\widetilde{\Omega}^{opt}=\varnothing(t=0,1,\cdots)$；反之亦然. 由定义 6，EA-I 等态于 EA-II，即 EA-I≌EA-II.　证毕.

命题 1 说明了 EA-I 和 EA-II 虽然有着不同的变异算子，但是在等态意义下是相同的. 根据引理 2，$\{\xi_t^{\mathcal{A}_1}\}_{t=0}^{+\infty}$ 和 $\{\xi_t^{\mathcal{A}_2}\}_{t=0}^{+\infty}$ 都是 Markov 链. 由(1+1)EA 算法的选择策略，两种进化算法一旦找到最优解(适值最大)就会保持到永远，所以，根据定义 8，$\{\xi_t^{\mathcal{A}_1}\}_{t=0}^{+\infty}$ 和 $\{\xi_t^{\mathcal{A}_2}\}_{t=0}^{+\infty}$ 都是吸收态 Markov 链. 根据定理 2，若 EA-I 收敛，EA-II 也收敛；反之亦然(根据推论 1).

4.3　满足强/弱态的关系进化算法举例

例 2.　EA-II 与 EA-III 的关系.

命题 2. EA-II 强态于 EA-III，即 EA-II≻EA-III.

证明.　假设进化算法 EA-II 和 EA-III 对应的随机过程分别为 $\{\xi_t^{\mathcal{A}_2}\}_{t=0}^{+\infty}$ 和 $\{\xi_t^{\mathcal{A}_3}\}_{t=0}^{+\infty}$. 类似定理 2 证明过程，记 $\widetilde{\Omega}^{opt}\subseteq\widetilde{\Omega}$ 是最优状态空间，对于 $\forall V\in\widetilde{\Omega}^{opt}$ 和 $t=0,1,\cdots$，假设 V 与 $\xi_t^{\mathcal{A}_2}$ 有 m_2 个位置不同，V 与 $\xi_t^{\mathcal{A}_3}$ 有 m_3 个位置不同. $P\{\xi_{t+1}^{\mathcal{A}_2}=V\mid\xi_t^{\mathcal{A}_2}\notin\widetilde{\Omega}^{opt}\}$ 计算如式(2)，而 $P\{\xi_{t+1}^{\mathcal{A}_3}=V\mid\xi_t^{\mathcal{A}_3}\notin\widetilde{\Omega}^{opt}\}$ 如式(3)：

$$P\{\xi_{t+1}^{\mathcal{A}_3}=V\mid\xi_t^{\mathcal{A}_3}\notin\widetilde{\Omega}^{opt}\}=\begin{cases}\left(\dfrac{1}{m}\right)^{m_3}, & m_3=0,1\\0, & m_3>1\end{cases}=\alpha_{t+1}^{(3)}\tag{3}$$

其中 $0<\alpha_t^{(2)},\alpha_t^{(3)}\le 1(t=1,2,\cdots)$ 而且 $\prod_{t=0}^{+\infty}(1-\alpha_t^{(2)})=0$，但是 $\prod_{t=0}^{+\infty}(1-\alpha_t^{(3)})=0$ 只有在 $m_3=0$，1 时成立. 据定义 6，$V\in\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})$ 成立，而 $V\in\Xi_t^{\mathcal{A}_3}(\alpha^{(3)})$ 也只有在 $m_3=0$，1 时成立. 由于 V 的任意性，对 $\forall V\in\widetilde{\Omega}^{opt}$，若 $V\in\Xi_t^{\mathcal{A}_3}(\alpha^{(3)})$ 则 $V\in\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})$，若 $V\in\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})$ 则 $V\in\Xi_t^{\mathcal{A}_3}(\alpha^{(3)})$ 只有在 $m_3=0,1$ 时成立，即不一定成立. 所以，对于任意非衰退实序列 $\beta^{(3)}=\{\beta_t^{(3)}\}$ 满足 $\beta_t^{(3)}>\alpha_t^{(3)}$，存在非衰退实序列 $\beta^{(2)}=\{\beta_t^{(2)}\}$ 满足 $\beta_t^{(2)}>\alpha_t^{(2)}$ 使得 $\Xi_t^{\mathcal{A}_2}(\beta^{(2)})\cap\widetilde{\Omega}^{opt}\supseteq\Xi_t^{\mathcal{A}_3}(\beta^{(3)})\cap\widetilde{\Omega}^{opt}=\varnothing$ 对 $t=1,2,\cdots$ 成立. 类似地，对任意 $\gamma^{(3)}=\{\gamma_t^{(3)}\}$ 满足 $\gamma_t^{(3)}\le\alpha_t^{(3)}(m_3=0,1)$ 和 $0<\gamma_t^{(3)}<1(m_3>1)$，存在 $\gamma^{(2)}=\{\gamma_t^{(2)}\}$ 满足 $\gamma_t^{(2)}\le\alpha_t^{(2)}$，使得 $\Xi_t^{\mathcal{A}_2}(\gamma^{(2)})\cap\widetilde{\Omega}^{opt}\supseteq\Xi_t^{\mathcal{A}_3}(\gamma^{(3)})\cap\widetilde{\Omega}^{opt}=\varnothing$. 反之亦然. 因此，EA-II 强态于 EA-III，即 EA-II≻EA-III.　证毕.

命题 2 说明采用随机一位变异方式的 EA-III 是弱态于采用逐位变异方式的 EA-II. 根据引理 1，$\{\xi_t^{\mathcal{A}_3}\}_{t=0}^{+\infty}$ 也是 Markov 链. 根据(1+1)EA 算法的选择策略，进化算法 EA-III 一旦找到最优解(适值最

大)就会保持到永远，所以，由定义 9，$\{\xi_t^{\mathcal{A}_3}\}_{t=0}^{+\infty}$ 也是吸收态 Markov 链. 根据命题 2，若 EA-III 收敛，EA-II 也收敛；根据推论 2，若 EA-II 不收敛，则 EA-III 也不收敛.

例 3.　EA-I、EA-II 与 EA-IV 的关系.

因为 EA-I、EA-II 和 EA-IV 三个算法的初始化设置一样，则它们的初始化可达状态集相等. 当基于相同状态时，EA-IV 能达到的最优状态，EA-I 和 EA-II 也都能达到；但是 EA-I 和 EA-II 达到的最优状态，EA-IV 则不一定能达到. 如基于所有基因位取值为 1 的初始状态(1，…，1)，EA-I 和 EA-II 可以达到含有“0”的状态，但是 EA-IV 则不能. 因此，可以得出命题 3 结论.

命题 3.　EA-IV 弱态于 EA-II 和 EA-I，即 EA-IV$\prec$EA-II$\simeq$EA-I.

证明.　假设进化算法 EA-II 和 EA-IV 对应的随机过程分别为 $\{\xi_t^{\mathcal{A}_2}\}_{t=0}^{+\infty}$ 和 $\{\xi_t^{\mathcal{A}_4}\}_{t=0}^{+\infty}$. 类似定理 4 证明过程，记 $\widetilde{\Omega}^{opt}\subseteq\widetilde{\Omega}$ 是最优状态空间，对于 $\forall V\in\widetilde{\Omega}^{opt}$ 和 $t=0,1,\cdots$，假设 V 与 $\xi_t^{\mathcal{A}_2}$ 有 m_2 个位置不同，V 与 $\xi_t^{\mathcal{A}_4}$ 有 m_4 个位置不同. $P\{\xi_{t+1}^{\mathcal{A}_2}=V\,|\,\xi_t^{\mathcal{A}_2}\notin\widetilde{\Omega}^{opt}\}$ 计算如式(2)，而当 $\xi_t^{\mathcal{A}_4}$ 可以通过一次插入变异达到 V 时，

$$P\{\xi_{t+1}^{\mathcal{A}_4}=V\,|\,\xi_t^{\mathcal{A}_4}\notin\widetilde{\Omega}^{opt}\}=\frac{1}{m}\cdot\frac{1}{m-1}=\alpha_{t+1}^{(4)}\quad(4)$$

而当 $\xi_t^{\mathcal{A}_4}$ 无法可以通过一次插入变异达到 V 时，$P\{\xi_{t+1}^{\mathcal{A}_4}=V\,|\,\xi_t^{\mathcal{A}_4}\notin\widetilde{\Omega}^{opt}\}=0$. 其中 $\alpha_t^{(2)},\alpha_t^{(4)}>0(t=1,2,\cdots)$ 而且 $\prod_{t=0}^{+\infty}(1-\alpha_t^{(2)})=0$，但是 $\prod_{t=0}^{+\infty}(1-\alpha_t^{(4)})=0$ 只有在 $\xi_t^{\mathcal{A}_4}$ 可以通过一次插入变异达到 V 时才能保证成立. 根据定义 6，$V\in\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})$ 成立，而 $V\in\Xi_t^{\mathcal{A}_4}(\alpha^{(4)})$ 也只有 $\xi_t^{\mathcal{A}_4}$ 可以通过一次插入变异达到 V 时成立. 由于 V 的任意性，对 $\forall V\in\widetilde{\Omega}^{opt}$，若 $V\in\Xi_t^{\mathcal{A}_4}(\alpha^{(4)})$ 则 $V\in\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})$，若 $V\in\Xi_t^{\mathcal{A}_2}(\alpha^{(2)})$ 则 $V\in\Xi_t^{\mathcal{A}_4}(\alpha^{(4)})$ 不一定成立. 所以，类似命题 1 和命题 2 对应的证明，对任意非衰退序列 $\beta^{(4)}$ 存在一个非退序列 $\beta^{(2)}$，使得对 $t=1,2,\cdots$，都有 $\Xi_t^{\mathcal{A}_2}(\beta^{(2)})\cap\widetilde{\Omega}^{opt}\supseteq\Xi_t^{\mathcal{A}_4}(\beta^{(4)})\cap\widetilde{\Omega}^{opt}$ 成立. 因此，EA-II 强态于 EA-IV，即 EA-IV$\prec$EA-II$\simeq$EA-I.　　证毕.

命题 3 说明了采用插入变异方式的 EA-IV 也是弱态于采用逐位变异的 EA-II. 考虑 EA-IV 对应的随机过程也属于吸收态 Markov 链，根据推论 2，如果 EA-IV 收敛，EA-I 和 EA-II 也必然收敛. 根据推论 2，若 EA-I 和 EA-II 不收敛，EA-IV 也肯定不收敛.

5　数值实验分析

为了验证研究理论分析结论的正确性，我们测试了 EA-I、EA-II、EA-III 和 EA-IV 求解线性、欺骗性优化问题的收敛性. 测试问题函数如下，求解结果见表 1 与表 2.

F1：$f=\sum_{i=1}^{M}x_i$，最优解(0，…，0)，最优值 0；

F2：$f=(x_1-1)^2+\sum_{i=2}^{M}(x_i-x_1)^2$，最优解(1，…，1)，最优值 0；

F3：$f=\sum_{i\in J_1}(x_i-1)^2+\sum_{i\in J_2}x_i^2$，最优解奇数位为 1，偶数位为 0，最优值 0；

F4：$f=100(x_1+x_2-1)^2+(x_1-1)^2+x_2^2+\sum_{i=3}^{M}x_i^2$，最优解(1，0，…，0)，最优值 0；

F5：$f=\sum_{i=1}^{M-1}(x_i-x_{i+1})^2$，最优解(1，…，1)和(0，…，0)，最优值 0.

F1 和 F3 是线性优化函数，F2、F4 和 F5 是欺骗性优化问题. 特别地，F5 欺骗性最大，除了全 0 和全 1 的最优解之外，次优解的跳跃性很大.

表 1　20 次计算的对比结果(M=10)

算法	指标	F1	F2	F3	F4	F5
EA-I	命中数	20	20	20	20	20
	迭代数	1108	932	527	668	680
EA-II	命中数	20	20	20	20	20
	迭代数	44	54	47	77	433
EA-III	命中数	20	14	20	14	4
	迭代数	27	21	20	19	19
EA-IV	命中数	0	0	4	0	0
	迭代数	$>10^4$	$>10^4$	13.5	$>10^4$	$>10^4$

表 2　20 次计算的对比结果(M=30)

算法	指标	F1	F2	F3	F4	F5
EA-I	命中数	20	14	20	20	1
	迭代数	$>10^5$	$>10^5$	$>10^5$	$>10^5$	$>10^5$
EA-II	命中数	20	15	20	20	2
	迭代数	259	200	247	610	73685
EA-III	命中数	20	15	20	12	1
	迭代数	98	94	93	101	74
EA-IV	命中数	0	0	0	0	0
	迭代数	$>10^5$	$>10^5$	$>10^5$	$>10^5$	$>10^5$

4 个(1＋1)EA 算法对每个问题都计算了 20 次. 表 1 和表 2 中的“命中数”是 20 次计算中达到最优解的次数，“迭代数”是达到最优解迭代次数的 20 次平均. 实验结果表明：

(1) EA-I 与 EA-II 在 5 个问题的求解"命中数"指标上非常相近，可以认为两种算法的收敛性是相当的. 实验结果验证了命题 1 结论的正确性，即 EA-I 等态于 EA-II. 若 EA-I 收敛，EA-II 也收敛；反之亦然.

(2) EA-II 总体求解结果比 EA-III、EA-IV 优，可以认为 EA-II 具有较好的收敛性. 实验结果验证了命题 2 和 3 结论的正确性，即 EA-II 强态于 EA-III、EA-IV. 如果 EA-III 或 EA-IV 收敛，EA-II 也必然收敛. 反之，EA-II 不收敛，EA-III 和 EA-IV 也肯定不收敛.

(3) 满足等态或强态关系的算法在"迭代数"方面有显著差异. 例如，EA-I 的"迭代数"远远超过了 EA-II，这与研究的理论分析并没有矛盾. 等态关系仅仅说明了两个算法在收敛性上存在等价性，而收敛性是基于迭代时间趋于无穷的假设，所以，允许算法在收敛速度和计算时间上存在差异. EA-II 强态于 EA-III 和 EA-IV 也有类似的情况出现.

本节得到的实验结果与之前理论分析的结论一致，验证了基于等态关系模型的进化算法收敛性分析与对比的相关结论. 进一步的研究将基于收敛速度和计算时间分析对研究模型进行改进.

6　结　论

基于进化算法的可达状态集与吸收态 Markov 性，建立了可用进化算法收敛性对比分析的两种关系模型，即"等态关系"与"强/弱态关系"，并且证明了前者是进化算法收敛性之间的一种等价关系，而后者则是一种偏序关系. 满足等态关系的若干个进化算法具有同收敛的性质，而强态/弱态关系则可以用于对比进化算法收敛性的强弱. 基于提出的关系模型，得到了一种新型的收敛性判别条件：满足吸收态 Markov 链性质的进化算法能否收敛，关键取决于算法产生的可达状态集与最优状态集是否从某次迭代开始都能有交集. 同时，提出了一种新的用于改进进化算法的思想：设计一个强态于原算法的进化算法. 未来研究将完善等态与强/弱态关系模型，并用于指导具体的算法设计和改进. 此外，还将针对进化算法的计算时间，建立相应的关系模型，从而对进化算法的计算时间复杂性进行比较研究.

致　谢　衷心感谢伯明翰大学姚新教授、北京大学苏开乐教授、华南理工大学杨晓伟教授、周育人教授与匿名审稿专家，他们对本文修改提出了宝贵建议！

参　考　文　献

[1] Yao X, Xu Y. Recent advances in evolutionary computation. Journal of Computer Science and Technology, 2006, 21(1): 1-18

[2] Holland J H. Adaptation in Natural and Artificial Systems. 1st ed. 1975, 2nd ed, Cambridge, MA: MIT Press, 1992

[3] Frantz D R. Non-linearities in genetic adaptive search[Ph. D. dissertation]. University of Michigen, Michigan, USA, 1972

[4] Goldberg D E. Simple genetic algorithms and the minimal deceptive problem//Davis L ed. Genetic Algorithms and Simulated Annealing. London Pitman, 1987: 74-78

[5] Bethke A D. Genetic algorithms as function optimizers [Ph. D. dissertation]. University of Michigan, Michigan, USA, 1980

[6] Grefenstette J J. Deception considered harmful//Whitley L D. Foundations of Genetic Algorithms. San Mateo, CA: Morgan Kaufmann, 1993, 2: 75-91

[7] Rudolph G. Convergence Properties of Evolutionary Algorithms. Hamburg: Verlag Dr. KovaYc, 1997

[8] Goldberg D E, Segrest P. Finite Markov chain analysis of genetic algorithm//Proceeding of the 2nd International Conference on Genetic Algorithms. Cambridge, MA, 1987: 7-8

[9] Eiben A E, Aarts E H, Van Hee K M. Global convergence of genetic algorithms: An infinite Markov chain analysis//Schwefel H P. Parallel Problem Solving from Nature. Heidelberg, Berlin: Springer-Verlag, 1991: 4-12

[10] Rudolph G. Convergence properties of canonical genetic algorithms. IEEE Transaction on Neural Networks, 1994, 5(1): 96-101

[11] De Jong K A. Are genetic algorithms function optimizers?//Proceedings of the Second Parallel Problems Solving from Nature Conference. Brussels, Belgium, 1992: 3-14

[12] Bäck T. The interaction of mutation rate, selection and self-adaptation within a genetic algorithm//Proceedings of the Parallel Problem Solving from Nature. Brussels, Belgium, 1992, 2: 84-94

[13] Muhlenbein H. How genetic algorithms really work. I: Mutation and hillclimbing//Proceedings of the Parallel Problem Solving from Nature. Brussels, Belgium, 1992, 2: 15-25

[14] Xu Zong-Ben, Gao Yong. Characteristic analysis and prevention on premature convergence in genetic algorithms. Science in China, Series E: Technological Sciences, 1996, 26(4): 364-375(in Chinese)
(徐宗本，高勇. 遗传算法过早收敛现象的特征分析及其预防. 中国科学(E辑), 1996, 26(4): 364-375)

[15] Zhang Jiang-She, Xu Zong-Ben, Leung Yee. Global annealing genetic algorithm and its convergence analysis. Science in China, Series E: Technological Sciences, 1997, 27(2): 154-164(in Chinese)
(张讲社，徐宗本，梁怡. 整体退火遗传算法及其收敛充要条件. 中国科学(E辑), 1997, 27(2): 154-164)

[16] Liang Yan-Chun, Zhou Chun-Guang, Wang Zai-Shen. Convergence analysis of an equivalent genetic algorithm based on extended strings. Chinese Journal of Computers, 1997, 20(8): 686-694(in Chinese)
(梁艳春,周春光,王在申. 基于扩展串的等价遗传算法的收敛性. 计算机学报,1997,20(8):686-694)

[17] Xu Zong-Ben, Nie Zan-Kan, Zhang Wen-Xiu. Almost sure convergence of genetic algorithms: A martingale approach. Chinese Journal of Computers, 2002, 25(8): 785-793(in Chinese)
(徐宗本,聂赞坎,张文修. 遗传算法的几乎必然强收敛性——鞅方法. 计算机学报,2002,25(8):785-793)

[18] Zhou Yu-Ren, Min Hua-Qing, Xu Xiao-Yuan, Li Yuan-Xiang. Multi-objective evolutionary algorithm and its convergence. Chinese Journal of Computers, 2004, 27(10): 1415-1421(in Chinese)
(周育人,闵华清,许孝元,李元香. 多目标演化算法的收敛性研究. 计算机学报,2004,27(10):1415-1421)

[19] Zhou Yu-Ren, Yue Xi-Shun, Zhou Ji-Xiang. Convergence rate and efficiency of evolutionary algorithms. Chinese Journal of Computers, 2004, 27(11): 1485-1491(in Chinese)
(周育人,岳喜顺,周继香. 演化算法的收敛速率与效率分析. 计算机学报,2004,27(11):1485-1491)

[20] Wang Yu-Ping, Liu Da-Lian. A global optimization evolutionary algorithm and its convergence based on a smooth scheme and line search. Chinese Journal of Computers, 2006, 29(4): 670-675(in Chinese)
(王宇平,刘大莲. 基于平滑技术和一维搜索的全局优化进化算法及其收敛性. 计算机学报,2006,29(4):670-675)

[21] Wang Yu-Ping, Li Ying-Hua. A novel quantum genetic algorithm for TSP. Chinese Journal of Computers, 2007, 35(5): 747-755(in Chinese)
(王宇平,李英华. 求解 TSP 的量子遗传算法. 计算机学报,2007,35(5):747-755)

[22] Liu Chun-An, Wang Yu-Ping. Evolutionary algorithm for dynamic multi-objective optimization problems and its convergence. Acta Electronica Sinica, 2007, 35(6): 1118-1121(in Chinese)
(刘淳安,王宇平. 动态多目标优化的进化算法及其收敛性分析. 电子学报,2007,35(6):1118-1121)

[23] Chen Guo-Liang, Wang Xu-Fa, Zhuang Zhen-Quan, Wang Dong-Sheng. Genetic algorithm and its applications. Beijing: Posts & Telecom Press, 1996(in Chinese)
(陈国良,王煦法,庄镇泉,王东生. 遗传算法及其应用. 北京:人民邮电出版社,1996)

[24] Li Min-Qiang, Kou Ji-Song, Lin Dan, Li Shu-Quan. The principles and applications of genetic algorithm. Beijing: Science Press, 2002(in Chinese)
(李敏强,寇纪淞,林丹,李书全. 遗传算法的基本理论与应用. 北京:科学出版社,2002)

[25] Zhang Wen-Xiu, Liang Jie. Mathematical foundation of genetic algorithms. Xi'an: Xi'an Jiaotong University Press, 2004(in Chinese)
(张文修,梁怡. 遗传算法的数学基础. 西安:西安交通大学出版社,2004)

[26] Huang H, Wu C G, Hao Z F. A pheromone-rate-based analysis on the convergence time of ACO algorithm. IEEE Transactions on Systems, Man, and Cybernetics-Part B, 2009, 39(4): 910-923

[27] Huang Han, Hao Zhi-Feng, Qin Yong. Time complexity of evolutionary programming. Journal of Computer Research and Development, 2008, 45(11): 1850-1857(in Chinese)
(黄翰,郝志峰,秦勇. 进化规划算法的时间复杂性分析. 计算机研究与发展,2008,45(11):1850-1857)

[28] Huang Han, Hao Zhi-Feng, Wu Chun-Guo, Qin Yong. The convergence speed of ant colony optimization. Chinese Journal of Computers, 2007, 30(8): 1344-1353(in Chinese)
(黄翰,郝志峰,吴春国,秦勇. 蚁群算法的收敛速度分析. 计算机学报,2007,30(8):1343-1353)

[29] He J, Reeves C, Witt C, Yao X. A note on problem difficulty measures in black-box optimization: Classification, realizations and predictability. Evolutionary Computation, 2007, 15(4): 435-443

[30] Droste S, Jansen T, Wegener I. On the analysis of the (1+1) evolutionary algorithms. Theoretical Computer Science, 2002, 276(1-2): 51-81

附 录.

符号说明表

符 号	说 明
$\mathcal{P}$	待求解的优化问题,本文主要研究组合优化问题
s_i	问题 $\mathcal{P}$的第 i 个可行解
$\mathcal{FS}^{\mathcal{P}}$	问题 $\mathcal{P}$的可行解集,$\mathcal{FS}^{\mathcal{P}}=\{s_1,s_2,\cdots\}$
$\mathcal{A},\mathcal{B}$	求解问题 $\mathcal{P}$的进化算法
$\mathcal{EAS}$	一组进化算法,$\mathcal{EAS}=\{\mathcal{A}_1,\mathcal{A}_2,\cdots,\mathcal{A}_k\}$
$\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$	求解 $\mathcal{P}$的进化算法 $\mathcal{A}$ 在第 t 代的种群
n	种群规模
$\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(i)$	种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$中的第 i 条染色体($i=1,2,\cdots,n$)
$\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$	由种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$衍生得到的子代种群
$\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(i))$	对染色体 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}(i)$进行解码而得的 $\mathcal{P}$的一个可行解
$\chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}})$	由种群 $\mathbb{P}_t^{\mathcal{A},\mathcal{P}}$经解码而得的 $\mathcal{P}$的一组可行解(解向量),即有,$\chi()=(\chi((1)),\chi((2)),\cdots,\chi((n)))$
$\underbrace{X_1\times X_2\times\cdots\times X_n}_{n}$	笛卡尔积,$\underbrace{X_1\times X_2\times\cdots\times X_n}_{n}\triangleq\{\{x_1,x_2,\cdots,x_n\}\mid x_i\in X_i,i=1,\cdots,n\}$

（续　表）

符号	说　明
$\Omega_n^{\mathcal{P}}$	$\Omega_n^{\mathcal{P}} \triangleq \underbrace{\mathcal{FS}^{\mathcal{P}} \times \mathcal{FS}^{\mathcal{P}} \cdots \times \mathcal{FS}^{\mathcal{P}}}_{n}$，由问题 $\mathcal{P}$ 的所有 n 维解向量组成的一个空间
$\xi_t^{\mathcal{A},\mathcal{P}}$	$\xi_t^{\mathcal{A},\mathcal{P}} \triangleq \chi(\mathbb{P}_t^{\mathcal{A},\mathcal{P}})$ 被视为一个随机向量，相应地，$\{\xi_t^{\mathcal{A},\mathcal{P}}\}_{t=1}^{+\infty}$ 被视为一个在状态空间 $\Omega_n^{\mathcal{P}}$ 上的随机过程
$\Omega^{\mathcal{P}}$	$\Omega^{\mathcal{P}} \triangleq \Omega_1^{\mathcal{P}} \cup \Omega_2^{\mathcal{P}} \cup \cdots \cup \Omega_N^{\mathcal{P}}$，一个扩展的状态空间，适于所有种群规模不超过 N 的求解 $\mathcal{P}$ 的进化算法
$\cong$	进化算法之间的(二元)等态关系

HUANG Han, born in 1980, Ph.D., associate professor. His research interests include the fields of theoretical foundation of evolutionary computation methods, evolutionary optimization, and application of evolutionary algorithms, etc.

LIN Zhi-Yong, born in 1977, Ph. D., associate professor. His research interests include algorithm design & analysis, machine learning and computation intelligence, etc.

HAO Zhi-Feng, born in 1968, Ph. D., professor, Ph. D. supervisor. His research interests include the fields of algebra group, analysis and design of algorithms, SAT, combinational optimization and data mining.

ZHANG Yu-Shan, born in 1975, Ph. D. candidate, lecturer. His mainly interests focus on the mathematical foundation of bio-inspired algorithms, including analysis of convergence and time complexity.

LI Xue-Qiang, born in 1983, Ph. D. candidate. His mainly interests focus on the evolutionary algorithm.

Background

Theoretical foundation research which includes algorithm modeling, convergence analysis and runtime analysis is one of the most important topics of evolutionary computation. Convergence has been studied for several years since it was a hot theoretical topic. As a result, researchers paid more attention to the study of runtime analysis. The authors' recent work also focuses on runtime analysis of ant colony optimization, evolutionary programming and particle swarm optimization. However, it doesn't mean the study of convergence analysis is not significant. As we know, most of the conclusion of convergence analysis is based on an irreducible Markov chain model which is similar to a conclusion of an exhaustion algorithm or a stochastic algorithm. Furthermore, the result is only for convergence analysis of single algorithm, but not for comparison with different algorithms in convergence performance. It stagnated researches into the convergence analysis of evolutionary algorithms.

This paper presents a theoretical work focusing on the practical application of evolutionary algorithm. How powerful the algorithm is, how to compare an algorithm with others and how to improve the performance of algorithms are interesting questions in the application. Therefore, a relation-based model is proposed to analyze the equivalence and ordering of evolutionary algorithm (EA) in convergence. According to equivalence-in-status relation, an ordering of stronger/weaker-in-status relation is proposed as a mathematical tool for comparing the convergence of evolutionary algorithm.

This work was supported by National Natural Science Foundation of China (61003066, 60873078, 61070033), Doctoral Program of the Ministry of Education (20090172120035), Natural Science Foundation of Guangdong Province (9151008901000165), Key Technology Research and Development Programs of Guangdong Province (2009B010800026, 10151601501000015) and the Fundamental Research Funds for the Central Universities, SCUT (2009ZM0052).

The presented model and theorem can be a foundation to study the equivalence, ordering and improvement of evolutionary algorithm. In the future study, the improved relation model will be studied to discuss the convergence and computational time in order to make more fundamental improvement of EAs.