

Multiagent 仿真与资源稀缺下的适应行为^①

赵纯均, 华立

(清华大学经济管理学院, 北京 100084)

摘要:以信息经济学和多 agent 系统理论为分析框架,借助美国 Santa Fe Institute 开发的仿真工具 Swarm,研究了短缺经济下个体的选择、竞争与适应行为,以及制度的进化过程。研究表明,简单的个体适应行为可以产生复杂的整体效果;个体基于历史以及相互间的信息交流导致了学习;满足一定的条件,新的制度安排将实现;个体之间信息的不对称可能使系统进化出类似道德风险和逆向选择的复杂行为。本研究有助于理解现实社会中普遍存在的占座现象;有助于解释财产权利以及制度的变迁。研究中所采用的多 agent 系统仿真方法为研究复杂的社会经济系统提供了一种相对新颖的观察视角。

关键词:多 agent 系统; 进化; Swarm; 短缺; 适应行为

分类号:N94

文献标识码:A

文章编号:1007-9807(1999)04-0001-07

0 引言

传统的新古典经济理论在分析中将个体视为同质的、追求效用最大化的完全理性的行为者,其基本方法建立在 19 世纪物理学理论(均衡、稳定、确定性动力学)之上,试图把研究对象看作结构简单的事物;而有关完全竞争、信息完全对等的基本假设与现实世界的情况有较大的出入。本世纪 60 年代以后,随着信息经济学的发展,开始引入不确定性、有限理性以及信息获取的成本等概念;同时制度经济学领域取得了可喜的进展,抛弃了原先的有关完全理性人的假设,认为个体受习惯和传统的驱动。90 年代以来,经济学研究中存在着这样一种看法,认为经济学应该建立在生物学理论之上(具有结构、特型、自组织和生命周期),永远处在发展的边缘;经济结构时常在组合、进化与发展。对经济系统的研究已经越来越不能回避复杂现象的存在,社会经济系统作为一类特殊的复杂系统,系统中的个体具有认知和学习的能力;作用者相互联系,平行发生作用,构成错综复杂的网络图景;系统表现出多层面、多结构,并能对未来具

有预期和展望。研究这类系统的目的一般不再局限于讨论系统的均衡与稳定,而是理解突现与适应之间相互关联发展的过程。

多 agent 系统(MAS 即 Multiagent System)理论是人工智能领域近些年来的研究热点。所谓 agent 指的是在复杂和动态变化环境中执行任务的自主或半自主的软件和硬件系统,是将知识和推理相结合的能动的实体。它们一般拥有自己的知识库、目标集和学习机制,能和环境协同作用、共同进化。一个多 agent 系统(MAS)由一组在系统结构中分别承担不同角色的 agent 组成,他们之间错综复杂的相互作用可能表现出单个 agent 所不具备的特性——突现行为(emergent behavior)——整体大于部分之和。

人工智能领域的进展与哲学和心理学等领域的突破是密不可分的。根据美国密西根大学遗传算法的创始人 Holland 教授与心理学家 Keith Holyoak, Richard Nisbett 和哲学家 Paul Thagard 多年的合作研究,认为人类的认知、学习过程建立在三项基本原则之上:知识能够以类似规则的心智结构来表达;这些规则始终处于竞争之中,经验

① 收稿日期:1999-02-01;修订日期:1999-07-10。

作者简介:赵纯均(1941-),男(汉族),四川荣昌人,清华大学经济管理学院教授,博士生导师。

使有用的规则越变越强,无用的规则越变越弱;具有说服力的新规则产生于旧规则的组合之中。他们的这些观点已经被多 agent 系统研究领域采用,并在航空港的管理、后勤运输、交通指挥系统、高级机器人系统以及分布式电力管理系统方面获得了成功的应用。

本文的基本思路即是借鉴上述心理学家和人工智能研究者的成果,采用多 agent 系统理论和计算机仿真相结合的方法,试图对现实社会中普遍存在的占座现象(采用标记或象征的方式表示资源的所有权)以及相应的正式或非正式的制度进化过程提供一种新颖的理解思路。在我们的模型中,行为个体是有限理性的,其行为基于相对简单的规则,但具有认知和适应的能力。个体对环境及其他个体的认识的信息是不完备的,个体特有的实践经历决定其具有不同的预测和判断能力。

仿真实验将借助 Santa Fe Institute 开发的 Swarm 软件进行。Swarm 最初是 Chris Langton

开发的一个用于研究人工生命的工具,后来经过 Nelson Minar, Roger Burkhart, Glen Ropella 等计算机专家的进一步完善;发展成为可用于多 agent 系统模型的仿真平台;它提供的一些分析和监测程序可使研究者将更多的精力放在描述 agent 的行为上,而不是计算机编程的技巧上。基于多 agent 理论的计算模型一般包括:创建的虚拟时空环境,显性的同时性控制,界定的实验对象以及各种监测和调整工具。在 Swarm 框架下,一个仿真程序主要由主程序(main)、监测模块(ObserverSwarm)以及模型表述模块(ModelSwarm)三部分组成。顾名思义,模型表述模块是整个计算模型的核心,agent 的系统环境、行为规则、学习机制以及目标集等都在这个模块中得到相应的表达。监测模块主要起辅助作用,通过图表等监测工具使实验者可以观测到系统的进化过程,或作进一步的数据分析。主程序的功能相对简单,仅提供模型的入口参数及激活功能。

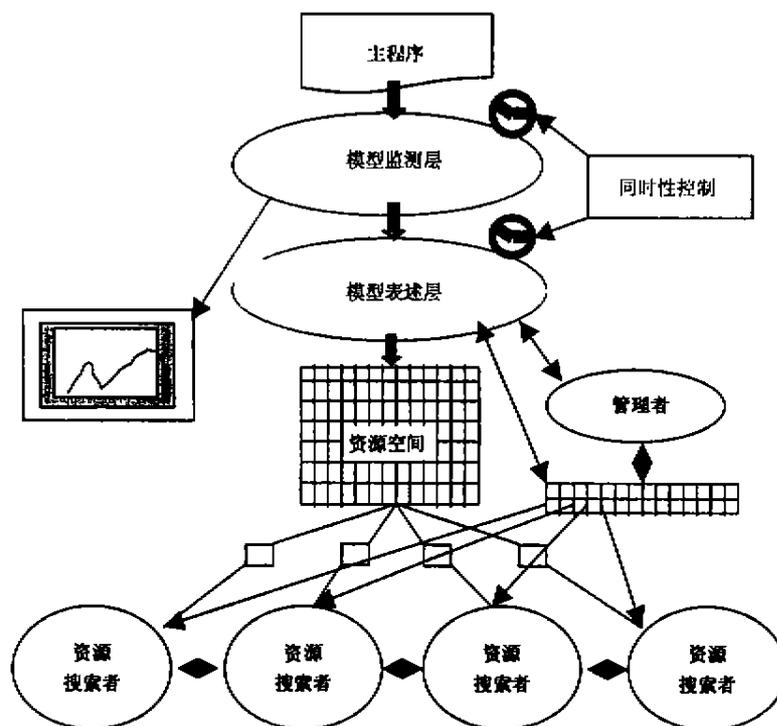


图1 基于 agent 的资源短缺模型框架

1 基于 agent 的资源短缺模型

资源短缺现象在现实社会中普遍存在,经济学本身即是一种研究稀缺资源的选择的科学。从

分析类似图书馆中的占座行为开始建立仿真模型,研究 agent 的学习、适应和相互竞争的进化过程;通过观察环境和作用者之间的交互作用,试图解释现实社会中在资源短缺的情况下的新的制度

安排的突现及发展过程. 这个模型的基本框架如图 1 所示.

模型的建立采用了自底向上的设计思想. 整个模型的核心在模型表述层, 它主要包括两类 agent. 一类是我们试图研究的资源搜索者, 即在资源短缺前提下寻找资源供给的人或公司; 另一类则是处于上层, 实施监控和调整的管理者, 对应于政府或现行的制度安排. 为了研究 agent 的学习和适应能力, 以及新制度的进化过程, 问题的抽象描述为: 在空间散布着一些资源可供开发, 作为资源寻求者的搜索者将按照某种行为规则去利用资源.

资源搜索者的基本策略集为: $\Phi = \{\mu_0, \mu_1, \mu_2, \mu_3, \mu_4\}$. 其中 μ_0 表示选择器械和维护时间的策略, 这和具体的维护效用曲线有关. μ_1 表示选择开始搜索时间的策略, 存在三条基本规则: (1) 固定的出发时间; (2) 将出发时间提前; (3) 将出发时间延后. μ_2 表示找到资源后的策略, 存在两种选择: (1) 马上开发利用; (2) 在资源上做好标记, 暂时离开. μ_3 表示在某个时段找不到闲余的资源, 但发现有些资源目前仅被作了标记而无人使用; 这时存在两种选择: (1) 使用该资源; (2) 放弃占用机会. μ_4 表示对资源作了标记的搜索者返回时, 双方对资源的所属会有一个讨价还价的过程. 双方都有放弃竞争和坚持己见的选择自由.

$$U_1(x, t) = 150 - 0.00025 \times (\text{MaintainTime} - 1730)^2 - \gamma \times \text{MaintainPeriod} \quad (1)$$

$$\text{MaintainPeriod} = 30 + [\text{UniformIntRand get IntegerWithMin}; -5 \text{ withMax}; 10] \quad (2)$$

维护过程所需的时间 MaintainPeriod 服从 25 分钟到 40 分钟之间的均匀分布. 一旦资源搜索者作出不维修或不做准备工作的决策, 就不能获得相应的附加效用值. 模型中还进一步假设, 管理者可以通过普测的方法, 了解资源搜索者对现行的制度调整的满意程度, 从而可以采取手段强化或弱化其调整的结果.

2 资源短缺下适应行为的仿真实验

在设定相关的参数之后, 开始仿真实验. 首先

当双方相互竞争的时候, 当前制度对以标记方式表示资源所属的方法所持的态度, 以及个体在以往竞争中的成败纪录将影响双方的谈判力度, 最终从概率意义上决定谁将在竞争中获胜.

每一个资源搜索者除了拥有以上的基本策略集外, 还有一个自我学习、相互交流, 不断适应的过程. 在这个模型中, 资源搜索者主要通过分析以往各种策略在使用中的表现情况, 以及与相邻的其他三个搜索者的信息交流来改善策略集中的行为规则组合的. 每一位资源搜索者拥有一个指标 $\text{SatisfyToInstitutions}$ 表示对当前的制度安排的满意程度; 它主要通过比较在调整期间的平均效用的变化情况进行参数的修正. 资源搜索者利用资源的效用曲线 $U(x, y, t)$ 由两部分组成. 第一部分效用函数记作 $U_1(x, t)$, 来自利用资源前所作的准备工作, 如机器的维修和保养. 它的大小与活动的时间安排有关, 即在某个时段安排某种活动获得的效用最大, 在其他时段则次之. 第二部分效用函数 $U_2(y, t)$ 则与实际利用资源程度有关, 即它是资源使用时间的函数 $U_2(x, t) = \gamma * \text{Use-Time}$. 在仿真中, 显性的同时性时段被选择在 $[1600, 2359]$; 并假设 $U_1(x, t)$ 在维护时间 $\text{MaintainTime} = 1730$ 时所获效用值最大. 不妨设 $U_1(x, t)$ 的计算公式为:

假设资源稀缺的程度为 1:2, 即每一份资源对应有两个搜索者希望占用. 假设空间中分布着 50 份资源, 100 位资源搜索者. 资源搜索者在初始时刻对策略集中的各种策略表现出同等偏好, 对当前的制度持较满意的态度 (初始化为 70 到 99 之间随机取值). 管理当局对以标记方式表示资源已被占用的初始态度置为 50, 即对采取这种方式既不赞成也不反对, 允许资源搜索者自由竞争. 系统运行 2000 轮后, 实验结果如图 2 所示.

由上图不难看出, 资源搜索者利用资源的平均效用稳步提高. 通过对历史情况的分析, 以及与

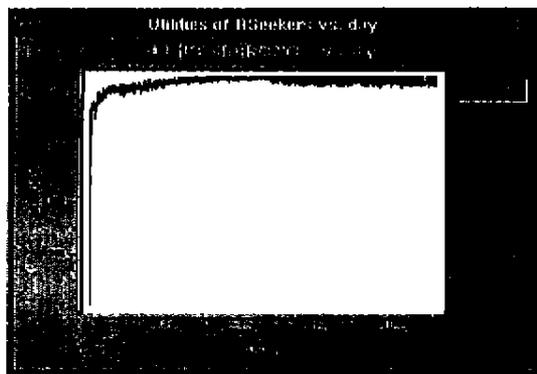


图 2a 资源搜索者的平均效用曲线

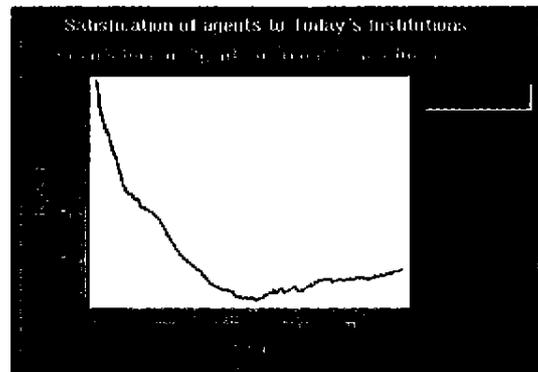


图 2b 资源搜索者的平均满意度曲线

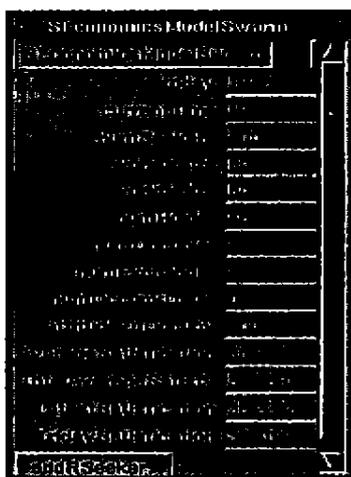


图 2c 参数监测工具

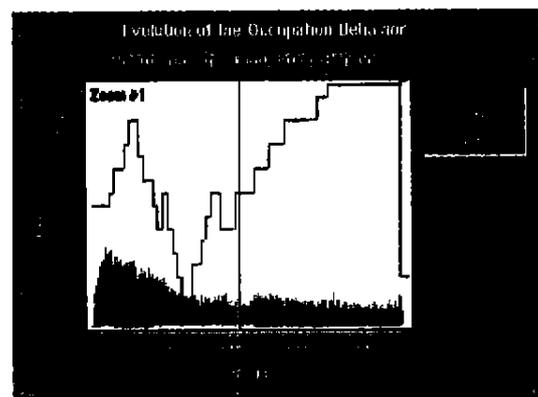


图 2d 新制度的演化变迁

相邻资源搜索者之间的信息交流,资源搜索者对策略集中的规则进行了优化组合,每个搜索者个体追求自身效用的最大化,其直接结果是空间中资源的整体利用效率提高了,管理者对以标记形式表示资源归属所持的态度也从最初的不置可否,发展到最后持默许的态度,这和现实社会中许多正式或非正式制度对财产所有权的承认方式具有相似之处。

实验进一步表明,资源搜索者选择的策略组合与管理者所持的态度密切相关。当管理者对占座行为持否定态度时,大多数资源搜索者采取的策略是:尽量提前出发的时间,找到资源后立即使用;晚到的搜索者不承认占座的合法性,试图夺取已被其他搜索者标记的资源。进化的直接结果是找到资源的搜索者不敢再离开物品,晚到的资源

搜索者则无从下手,当管理者对占座行为持不置可否的态度时,搜索者表现出对资源的激烈竞争,这时被标记的资源经常被其他搜索者占用;当标记者返回时,双方会有一个竞价和争夺的过程,当管理者对占座行为持赞成态度时,大多数搜索者选择早早去寻找资源的策略;找到后再逐步将维护时间调整到 1730 附近。在这种制度安排下,找到资源并作了标记的搜索者在以后的竞争过程中有较大的获胜机会,也迫使晚到的搜索者选择避免与物主之间的正面冲突,当然上面的策略调整还与具体资源搜索者的历史纪录,以及它周围的其他三个 agent 的表现有关。在过去的争夺过程中,有较多失败经历的 agent 表现出较大的惰性:它宁肯放弃开发前的准备工作也不愿意离开已占据的资源。同时还发现,表现好的资源搜索者经常

聚集在一起,表现差的也一样,这和他们之间不间断的信息交流是分不开的.从空间看,资源搜索者全体进化出错综复杂的网络式组织结构.

对照图 2a 和图 2d,资源搜索者的平均效用在管理者对占座行为持禁止态度时方差较少,而持赞同态度时方差较大.这是因为资源搜索者在制度对占座行为持肯定态度时,逐渐认识到:一旦自己已经找到资源并作了标记,便不必再匆忙返回.因为在这种制度安排下,资源搜索者在返回的时候有较大的可能收回被他人挪用的资源或者根本就不会有别的搜索者去使用已被标记的资源.资源平均利用效率的降低,一方面来自占据资源的搜索者选择了并非最优的行动策略;另一方面则由于其他的资源搜索者并不掌握占座的搜索者的个人信息,即对方会在什么时候返回及拥有的谈判实力;从而在现行的制度安排下选择了避免正面冲突的策略.仿真中所表现出来的这种现象和现实社会中的道德风险和逆向选择问题具有类似的内涵.在企业竞争和合作过程中,由于信息的不对称,产生类似旧车市场的逆向选择问题,不但造成资源利用效率的降低,有时甚至阻碍交易的发生;其根本原因和资源搜索者的行动选择颇为相似.交易双方都很难掌握对方的全部信息,从而只能根据大环境以及对方以往的经营业绩作一粗略的判断.所以从提高资源利用效率的角度来看,管理者所持的最佳态度似乎应该是禁止搜索者采用标记方式表明资源的所属;然而一旦管理者真地采取这种方式,资源搜索者为了使自己找到的资源不落入别人手中,逐渐会选择放弃前期准备工作的策略,最终导致资源整体利用效率的下滑和不满情绪的增加.这一点可以通过下面的仿真看到:在系统运行了 3 000 轮以后,突然改变管理者的态度,观察发生的现象.

正如在上面所分析的,资源搜索者对当前制度的强制调节表示了不满.也就是说在目前的短缺情况下,进化而来的制度安排是有效的 (efficient);在不改变既定短缺条件下强行禁止以标记方式表示资源所属的行为,只能导致系统内个体的不满和资源整体利用效率的降低.

在仿真中,还通过调整资源的短缺强度,以及当局一开始所采取的基本态度来观察系统的进化情况.实验表明,制度的进化过程虽有所不同,最

终的进化结果却大同小异.这一点和现实中的情况基本相符.进化路径的不同相应于每个资源搜索者具有一定的记忆能力,其策略的调整不是一蹴而就,而是需要一个适应的过程.

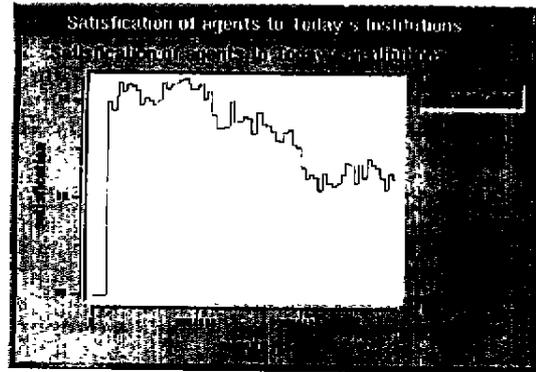


图 3a 资源搜索者的平均满意度曲线

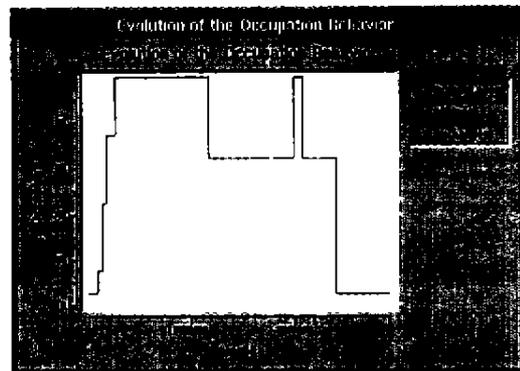


图 3b 新制度的演化变迁

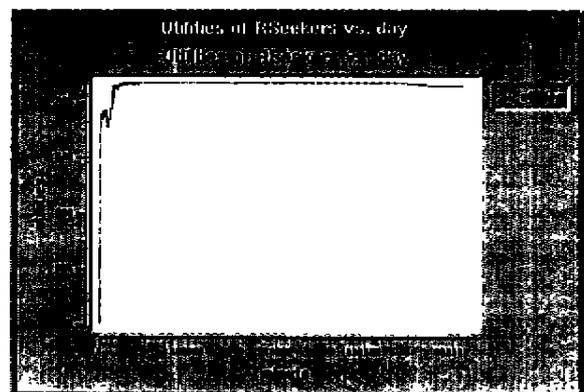


图 3c 资源搜索者的平均效用曲线

对比图 2a 和图 3c 的平均效用变化曲线,图 3c 中的曲线更平坦而少变化.这和管理者或当前制度所持态度的明朗情况是密切相关的.在图 2 的仿真实验中,管理者采取的是根据 agent 的反

应逐步调整现行政策的策略。在这种情况下,有些 agent 可能产生与现行调整方向相反的预测,把握不准制度的具体倾向,从而出现策略组合的多样性和意见的分散,表现为平均效用曲线的波动较大。相反,在图 3 中制度调整的幅度较大,制度的选择显示出对占座行为明显的好恶,这样资源搜索者就能够比较准确地把握当前的政策趋向,从而作出有利于自身的决策;表现为决策的单一和行动上的统一。

3 结论与意义

一种制度通常被定义为一套行为规则,他们被用于支配特定的行为模式与相互关系。一种组织则一般被看作是一个决策单位——一个家庭,一个企业,一个局——由它来实施对资源的控制。本文以信息经济学和多 agent 系统理论作为分析的框架,研究了资源稀缺下的制度进化以及组织

在环境中的学习和适应过程。研究表明,在资源稀缺的条件下,对资源采用标记方式表示其归属的新制度将进化出现;这种制度安排在不改变当前短缺条件的情况下是有效的,强行调整必将导致系统整体利益的下降和系统内组织的不满情绪的增加。在制度的进化与发展过程中,作为制度推行者的管理者和制度适应者的组织或其他决策单位相互作用,共同进化。管理者政令严明,态度明确,可使系统中的决策单位较快适应并作出准确的判断与预测,从而表现为上令下效,协调统一。

本文的意义在于为进行复杂的社会经济系统的研究提供了一种相对新颖的仿真方法,即多 agent 系统的仿真实验。它抛弃了传统经济学中许多未经证实的理论假设,通过定义个体的基本行为以及他们所处的局部环境,并激发系统的不断进化;从变化和突现行为中把握系统展现出来的更深层次的本质,从而为进一步的理论研究提供可资借鉴的素材和洞见。

参考文献:

- [1] Tomas B. Klos, Bart Nooteboom. Adaptive governance; the role of loyalty[R]. Santa Fe Institute working paper, May 26, 1998.
- [2] Arthur W B. Complexity in economic and financial markets[J]. Complexity, April, 1995, 1(1)
- [3] Arthur W B. Inductive reasoning and bounded rationality[J]. American Economic Review, 1994, 84, 406~411
- [4] Neil Gross. Introduction, into the wild frontier[J]. Business Week, June 23, 1997
- [5] Shareen Joshi. An explanation of generic behavior in an evolving financial market[C]. Santa Fe Institute Working Paper, 1998
- [6] 阎伟. 占座现象的实证观察和理论分析[M]. 经济研究, 1998, 74~80
- [7] Jacobson M J. Complexity and complex systems; a brief overview. SFI Web, 1998
- [8] Vidal J M. Computational agents that learn about agents; algorithms for their design and a predictive theory of their behavior. Doctor Dissertation of Philosophy, Computer Science and Engineering in the University of Michigan, 1998
- [9] von Martial F. Coordinating plans of autonomous agents[M]. Berlin, Springer-Verlag, c1992
- [10] Swarm Development Group. Documentation Set for Swarm 1. 2, June 25, 1998
- [11] Holland J H. Adaptation in natural and artificial systems[M]. MIT Press, 1992
- [12] L. E. 戴维斯, D. C. 诺斯. 制度变迁的理论概念与原因[M]. 上海: 上海三联出版社
- [13] 张维迎. 博弈论与信息经济学[M]. 上海: 上海人民出版社, 1996
- [14] Crutchfield J P, Mitchell M. The evolution of emergent computation[C]. Proceeding of the National Academy of Sciences, USA, 1995, 92, 10742~10746

Multiagent simulation and individual's adaptive behaviors in resource shortage

ZHAO Chun-jun, HUA Li

School of Economics and Management, Tsinghua University, Beijing 100084

Abstract: Theories of information economics and multiagent system are adopted as analysis frame for this paper, with Swarm as the simulation tool. Swarm is a software package for multiagent simulation of complex systems being developed at the Santa Fe Institute, which has been a useful tool for researchers in a variety of disciplines, especially artificial life. In this paper we use it to describe agents' selection, competition, and adaptive behaviors in resource shortage, and the evolution of institutions as a whole. Agents interact each other by local information and relatively simple reinforcement algorithms, and new data are collected from the environment and processed, which change the internal model of the agent in the end. Research evinces that simple adaptive behaviors of individuals can lead to complex system characteristics in higher level. Agent learns a lot from analysis of the operation history and by communications among them. New institution emerges in some conditions, which is often quite efficient in the context. Furthermore, we observe that complex system behaviors such as moral hazard and reverse selection may evolve out due to informational asymmetries. This research helps to explain the universal occupation phenomenon in real world, and the evolution of property right and institutions. Above all, multiagent simulation used in this paper offered a quite new perspective in research of complex social-economic phenomena.

Keywords: multiagent system(MAS); evolution; Swarm; shortage; adaptive behavior