

基于 Agent 仿真模型的校核与验证研究

伊文英 李 波

(西安交通大学电子信息与工程学院 西安 710049)

摘 要 首先对国内外基于 Agent 仿真模型的校核与验证方面的发展情况进行了述评;然后,提出了一个完整的基于 Agent 仿真模型的校核与验证框架,该方案中包括表面验证、参数灵敏度分析、模型校准与运行时验证;最后以环境经济政策仿真模型中的校核与验证为例对该框架进行了简单说明。

关键词 基于 Agent 仿真模型,校核与验证,经济仿真

中图分类号 TP312 **文献标识码** A

Research on Verification and Validation of Agent-based Simulation Models

YI Wen-ying LI Bo

(Department of Electronic Information and Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract We firstly reviewed the development of verification and validation simulation models using agent-based paradigm. Then a process for verification and validation of agent-based simulation models was proposed, which combines face validation, sensitivity analysis, calibration and operational validation. Finally, the verification and validation of EE-PSS (economic and environmental policy simulation system) was taken as an example to simply illustrate the suggested process.

Keywords Agent-based simulation models, Verification and validation, Economic simulation

1 国内外发展情况述评

仿真模型的有效性是仿真模型最重要的属性。仿真模型的有效性评估是当今仿真技术研究中一个十分重要的关键课题,也是当今国际仿真界研究的一个热点问题和仿真学术会议的重要议题。

仿真模型的有效性评估主要包括模型校核(Verification)和验证(Validation)两部分内容,2000 年之前的发展国内已有详实的介绍^[1]。仿真模型有效性评估是 VV&A 的重要组成部分,关于 VV&A 的发展情况在文献[2]中也有详细介绍^[2]。

基于 Agent 的建模与仿真(ABMS)是研究复杂系统的有效途径和建模仿真方法学,是当前最具活力、有所突破的仿真方法学,已经成为系统仿真领域的一个新的研究方向,主要应用于经济领域、社会科学领域和军事领域等^[3]。有效性评估是 Agent 仿真模型开发中的一个重要的环节。但是在基于 Agent 的建模与仿真中,宏观与微观被有机地联系起来,主体具有主动、适应性,随机因素被引入,同时验证所需的实际数据的不易得、不可得,使得基于 Agent 的有效性评估比较特殊,因此,Agent 仿真模型的有效性评估需要使用一些特别的校核与验证方法。

国外在基于 Agent 仿真模型的有效性评估方面的研究开始得较早,虽然迄今为止仍未形成完整的有效性评估标准和

规范,特别是在经济模型中只使用特征事实来验证模型,如 ASPEN 和 EURACE,但是总体来讲已经取得了显著的成果:Robert G. Sargent 重点研究了仿真模型的 V&V,描述了 4 种仿真模型有效性的评估方法,给出了模型开发过程的 V&V 示例,重点讨论了概念模型、仿真模型、模型功能和数据有效性的验证过程,给出了模型验证的基本步骤和 V&V 的文档管理办法,对基于 Agent 的仿真模型的校核与验证有一定的借鉴意义^[4]。2006 年 Macro 和 Elinor 指出绝大部分的 Agent 仿真模型的有效性评估仅停留在“概念验证”阶段,他们把 Agent 仿真模型和理论验证方法相结合,提出了 4 种适用于 Agent 仿真模型的理论验证方法:案例研究(case studies)、特征事实(stylized facts)、角色扮演游戏(role-playing games)和室内实验(laboratory experiments)^[5]。2009 年 Ngo 通过研究热带雨林中迁移农业的 Agent 仿真模型提出了一个通用的基于 Agent 仿真模型的校核与验证方案,但是部分细节与他的项目更相关^[6]。德国的学者 Franziska 在 Ngo 的基础上提出了一个基于 Agent 仿真模型的验证过程,包括表面验证(face validation)、敏感性分析(sensitivity analysis)、校准(calibration)和统计验证(statistical validation)^[7]。

我国 20 世纪 70 年代末才引入仿真技术,一直以来都是重点关注仿真模型的开发和应用,模型校核与验证发展缓慢。但是进入 21 世纪以来,随着仿真技术的发展,VV&A 逐渐得到重视,基本上对国外的 VV&A 理论研究进行了较好的跟

本文受国家环境保护公益性行业科研专项项目:我国环境与经济政策总体示范研究(201109076)资助。

伊文英(1985—),女,硕士生,主要研究方向为高性能仿真;李 波(1968—),男,教授,主要研究方向为分布式计算及软件工程、高性能仿真、模型数值模拟及可视化、地理信息系统、决策支持系统等,E-mail: boblee@xjtu.edu.cn.

踪并且 VV&A 取得了较好的发展。但在基于 Agent 仿真模型的校核与验证方面的研究比较少,少有文献可以参考。廖守亿和戴金海在复杂系统和基于 Agent 的建模与仿真方面的研究比较深入,提出了基于 Agent 的建模与仿真框架,并对基于 Agent 的建模与仿真模型的校核与验证进行了探讨,提出了两种基于 Agent 的模型有效性验证方法——预测有效和描述有效,但是研究集中在对方法本身的逻辑分析和描述上,没有给出具体的实现方案^[9]。随着基于 CAS 理论的多主体仿真的发展,特别是 ACE 的出现和发展,我国基于 Agent 仿真模型的校核与验证方面的研究工作亟需得到加强。

2 仿真模型校核与验证的定义及在 M&S 中的应用模式

如同 M&S(Modeling and Simulation)本身的概念,V&V 的基本概念也是一个逐渐发展和完善的过程,虽然 V&V 是整个 M&S 领域的概念,但是对于不同类型的 M&S,V&V 工作所适用的技术方法、工作流程等存在差别,但是 V&V 的概念是没有差别的,因此在 V&V 的概念上没有必要强调 M&S 的类型(由于确认是权威机构验收模型、确定模型是否可用于具体应用的过程,这里不予考虑)。在此本文使用 V&V 进行如下定义^[6]:

校核(Verification):在整个 M&S 生命周期中,促进并决定模型、仿真系统及相关属于精确描述开发者的概念描述及相关技术规范的活动和过程。

验证(Validation):在整个 M&S 生命周期中,根据 M&S 开发的预期目的,促进并确定模型、仿真系统及相关数据精确描述被仿真对象的活动和过程。

V&V 是贯穿于 M&S 开发的全生命周期的活动。V&V 工作的对象在 M&S 开发的初期主要是所有的模型及相关数据,而在 M&S 开发后期,主要是整个仿真系统及其相关数据。简单的描述模型开发过程中它们之间的关系可以用图 1 表示,但是这种方式过于简单,即没有准确地描述出 V&V 与 M&S 全生命周期各阶段的关系;而且图中部分内容还需要有进一步的推敲,比如概念模型不仅仅需要验证,同样需要进行校核工作^[4]。

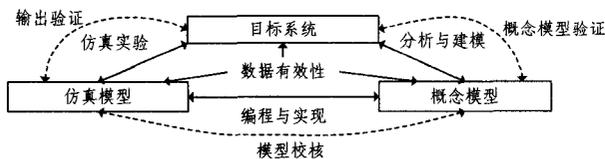


图 1 M&S 与 V&V 的关系

3 基于 Agent 仿真中存在的问题

本节简要地描述一些 Agent 仿真中存在的问题,这些问题可能会给仿真模型校核与验证带来困难。

3.1 模型输出的描述

如果在一个仿真模型中可以找到能够恰当表示系统输出的变量,则该模型可以使用理论验证或统计验证技术。在宏观层面上很容易找到这样的一个量,比如一个经济仿真系统中的变量 GDP,但是由于很难合理地捕捉到个体行为的特点,比如社会系统中个体的决策行为,因此很难在主体微观层面上找到这样的变量。

3.2 关注瞬间动态特性

基于 Agent 的仿真系统非常适合用来研究主体的瞬时动态特性,主体的交互与演化最终形成了一般均衡或稳定的状态。仿真的最终状态可以使用传统的验证技术进行验证,但是仿真系统焦点的瞬时动态特性也需要被验证。因此,统计方法这种间接验证则不如基于时间序列的验证有用,当前基于时间序列的趋势分析和数据挖掘或许可以用来验证 Agent 仿真系统的瞬时动态。然而,可能由于数据的不易获取,当前这方面的研究还是一片空白。

3.3 非线性与脆弱性

非线性与脆弱性指的是涌现现象,即基于 Agent 仿真模型中由于主体与主体、主体与环境之间相互影响和作用而导致的复杂系统整体的宏观行为。从复杂系统的个体行为无法直接推导出系统的宏观行为,所以涌现现象具有不可预测性。在仿真模型中仿真结果与参数变化之间具有非线性关系,微小的参数调整可能导致迥异的输出结果,这样就很难分析出参数与输出之间的关系。ABMS 区别于传统的建模与仿真方法的关键原因是其可以捕捉复杂系统的涌现现象,也正是这一点使得对基于 Agent 仿真模型的校核与验证比较困难。

3.4 验证数据的不可得、不易得

基于 Agent 仿真模型主要应用于复杂系统的研究,例如经济系统、社会系统、军事系统等,我们无法在目标系统上做实验,所以无法得到目标系统的数据,即使目标系统宏观层面上的一些数据可以直接找到或者收集整理而得,但与仿真实验的输出本身存在一定偏差。

模型行为的理论验证中微观层面所需要的数据通常也是不可得的,因为微观层面个体行为级别的数据很难获取;另外,调研的成本昂贵而且具有潜在的片面性。但是这些数据又是 Agent 仿真模型主体层面验证所必需的,这样就给基于 Agent 仿真模型验证带来困难。

4 基于 Agent 仿真模型的校核与验证框架

在基于 Agent 仿真模型的开发中,概念模型的验证非常重要,如果一个仿真模型的概念模型有效性没有得到充分验证,那么这个模型的开发就没有意义。基于 Agent 仿真模型中概念模型的验证与一般的 M&S 概念模型验证方法类似,这里不再赘述。本节重点研究基于 Agent 仿真模型的校核与验证,初步提出了一个系统的 ABMS 校核与验证方案。

基于 Agent 仿真模型的有效性验证主要包括两个方面:1)内部有效性验证,主要指仿真模型与概念模型的一致性验证,即仿真模型的逻辑合理性验证,称为模型校核;2)外部有效性验证,主要指输出验证,旨在验证仿真模型的输出结果与目标系统数据的一致性。

假设概念模型已经得到验证,ABMS 中概念模型必须转换成等价的可计算模型,校核内容包括阐明设计细节,以及判断根据概念模型是否正确地对概念模型进行计算机编码实现。

校核与验证从一个可运行的模型开始,图 2 展示了一个完整的 M&S 生命周期中的 V&V 过程,接下来将对这些步骤进行详细介绍,其中表面验证与灵敏度分析、模型校准通常被结合在一起称为校核,输出验证通常被称为模型验证。

图 2 中各个部分是逐步迭代进行的。要得到一个合理的

仿真系统模型,需要反复地进行校核与验证工作。一个仿真模型只有在其实验环境下完整地执行校核与验证过程才能输出可靠、有效的结果。但是在基于 Agent 的仿真系统中,很少有仿真模型会应用完整的校核与验证系统对模型进行校验,多数仿真模型仅仅对运行有效性进行验证,EEPSS 系统也仅使用了其中的一部分校验方法。

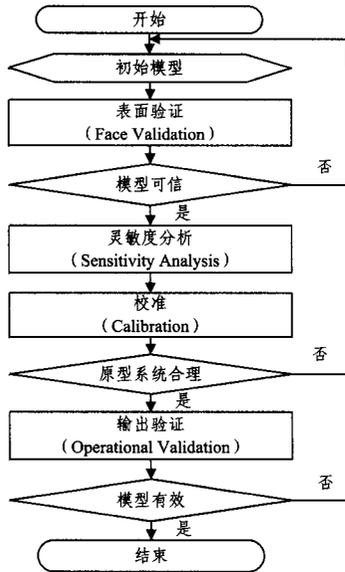


图2 基于 Agent 仿真模型的校核与验证框架

下面是对校核与验证过程中的每个阶段进行详细介绍,第5节将介绍该方案在基于 Agent 的环境经济仿真系统(Economic and Environmental Policy Simulation System, EEPSS)中的应用。

4.1 模型校核

模型校核旨在确信模型的逻辑合理性,确定、优化模型参数使得模型有效,包括表面验证、灵敏度分析与模型校准,主要通过修改代码和算法来确定影响模型输出的输入参数以及使用优化方法确定一组最优的模型参数值,使得模型输出结果与实际数据能够更好地相符。校核包括表面验证和检验模型的稳定性两个过程。

4.1.1 表面验证

表面验证在仿真模型开发中起着很重要的作用,目的是使验证仿真模型与概念模型之间保持一致性,主要应用于仿真模型开发的早期。表面验证主要是对系统进行描述和检查,例如形成规范化的文档、检查程序代码等,包含大量模型细节设计规范的检查。表面验证中与输出无关的校核可以使用经典软件工程和面向 Agent 的软件工程中的一些静态校核技术,此外还有以下方法。

(1)跟踪测试法:对模型中不同类型实体的行为状态进行跟踪,以确定模型逻辑上是否正确。这一过程包括校核交互算法,即给定一个 Agent 和另外一个 Agent 的某种状态属性 es,分析其后续状态 en 的合理性。

(2)输出评估法:对模型某一 Agent 的各属性变量之间的依赖关系进行分析,模型运行结束后,根据之前的分析结果对该 Agent 属性变量之间的关系进行验证,若一致则说明模型运行的逻辑具有合理性。

4.1.2 检验模型的稳定性

在检验模型的稳定性之前,可先进行灵敏度分析和模型校准,具体介绍如下。

(1)灵敏度分析

在 ABMS 模型中,灵敏度分析用来检验模型参数对模型行为和模型输出结果的影响,在模型校准之前可以选出对仿真结果影响较大的参数,为模型校准确定一个最小的参数集。最通用的方法是保持其他参数值不变,一次只改变一个参数的值。对模型灵敏度的定性分析如下:

设 S_0 对应的仿真系统模型参数集合(假定 n 个参数),修改某一参数值得到 S_1 ,假定 Y 为仿真系统输出集,则 S_0 对应输出集 Y_0 , S_1 对应输出集 Y_1 ,判断是否有:

$$|Y_1 - Y_0| < \delta \quad (1)$$

其中, δ 为用户给定的可接受的允许值。若满足公式,则说明所修改的参数对仿真输出结果无明显影响,可以从参数集合中剔除,依次迭代,最后得到一组对仿真输出结果影响较大的参数集合。

(2)模型校准

模型校准旨在确定参数值的范围,调整模型以期与真实系统更相似。把整个仿真系统看成一个黑盒子,然后使用高效的优化策略来确定最优参数配置。借鉴 GA^[6] 对模型的输入参数进行优化,操作步骤如图3所示。

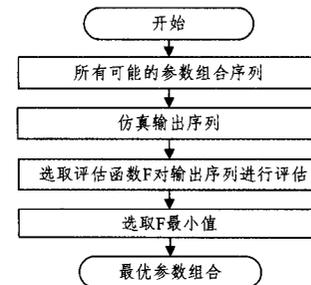


图3 参数优化流程图

若有 n 个参数,每个参数的取值情况有 $i_j (1 \leq j \leq n)$ 种,则共有 $m \leq \prod i_j (1 \leq j \leq n)$ 组可能的参数值,把每一组参数值配置到模型中进行计算,输出结果有 m 个序列。选取一个模型输出变量,且已知该变量系统数据序列 $\{X_{1,i}\}$ (样本大小为 s),模型输出的数据序列为 $\{X_{2,k}\} (1 \leq k \leq m)$ 。

选择评估函数如下:

$$F = \min(RMSE_j, 1 \leq j \leq m) = \min(\sqrt{\sum_{i=1}^s (x_{1,i} - x_{2,i})^2 / s}) \quad (2)$$

其中,均方根误差(Root Mean Squared Error, RMSE)是观测值与真值偏差的平方和,再求与观测次数值的平方根,均方根误差用以衡量观测值与真值之间的偏差,均方根误差对一组测量中的特大或特小误差反应非常敏感,所以均方根误差能够很好地反映出测量的精度。这里用于对模型输出的变量进行全局性评估,选择与 F 最小值对应的那组参数值做为模型的最优输入参数值。Agent 仿真模型的参数可能比较多,简单地依次执行特别浪费时间,可以选择使用批处理的方式来获取最优参数。

对于检验模型的稳定性,即检验多次运行模型输出结果的一致性,可以采用统计比较(仿真结果与真实系统数据)和一致性检验(多次运行仿真结果观测其相似性)两种方法。

由于基于 Agent 仿真模型的特殊性及其真实数据的不可得等原因,这里使用的确定模型稳定性的一致性检验方法如下:在同一输入参数的配置下,运行模型 n 次,得到模型某一输出变量的 n 个随机样本 $\{x_1\}, \{x_2\}, \dots, \{x_n\}$ (样本容量为

n), 每个随机变量的样本均值 $\bar{x}_i (1 \leq i \leq n)$ 构成一个新的随机变量 $\bar{x}(n)$, 求 $\bar{x}(n)$ 的方差 $S_2(n)$, 分析 $S_2(n)$, 其值越小越好。若 $S_2(n) \leq ARA$ (Acceptable Range of Accuracy, 设计者根据模型设计目标给定一个可接受的精度范围), 特别是当 $S_2(n) = 0$ 时, 认为模型在同一条件下的多次运行结果是相似的或相同的, 说明模型运行稳定。为了更好地验证模型的稳定性, 可多选择几个输出变量进行分析。

4.2 模型验证

模型验证是仿真模型开发中最重要的一环, 因为它确定了仿真模型的输出行为是否满足模型在其应用领域的预期精度要求, 即验证了仿真模型的有效性, 这里主要指输出验证。

影响仿真模型验证的主要因素是问题实体(系统)是否可观察, 这里可观察指的是可以收集到真实系统数据。根据问题实体的可观察性和决策方法对输出验证方法进行分类, 如表 1 所列。“对比”指的是仿真结果与真实系统或者另外一个仿真模型的输出结果使用图形化展示或者统计检验的方法进行比较; “探索模型行为”指的是使用恰当的验证技术, 包括参数的灵敏度分析来检验仿真结果。以下实验均在同等的实验条件下进行。

表 1 输出验证分类

方法类别	可观察的系统	不可观察的系统
主观方法	图形化方法进行比较 探索模型行为	探索模型行为
客观方法	统计检验方法	使用统计检验法与参考模型比较

仿真模型为了得到高可信度, 通常需要在多种实验条件下对仿真结果和真实系统的数据进行对比。通常不可观察的系统很难获得高可信度, 这种情况下, 尽量寻找有效的参考模型, 把仿真结果与参考模型相比较。

4.2.1 可观察的系统中的输出验证

在可观察的系统中, 因为真实数据可得, 可以将仿真模型结果与真实系统数据进行输出验证。验证方法主要分为两类: 一类是定性分析, 主要通过主观的判断比较, 如图灵测试、图形化展示等对仿真结果进行评价; 另一类验证方法是客观定量分析, 一般利用置信区间和假设检验来比较仿真实验和真实系统的输出, 包括静态输出和动态输出特性的比较两种方式。定量分析方法主要包括假设检验与置信区间法, 这方面的文献论述有很多, 以下内容介绍基于 Agent 仿真模型的验证中所使用的统计验证方法, 但并不限于此。

(1) 定性分析方法

主观判断模型与系统数据的偏差大小, 以此判定仿真模型的有效性。若偏差比较大则表明模型不理想, 反之若模型与系统数据能够较好地吻合, 则可认为仿真模型是有效的。定性分析主要采用以下 4 种方法:

a) 数据分布图: 直观展示数据的变化趋势及总体特点;

b) 箱线图: 反映一组或多组连续型定量数据的中心位置和散布范围, 若仿真模型的箱线图的第一分位线高于另外一个的中位线, 则表明仿真结果无效;

c) 行为图/样本直方图: 行为图是用于判断模型与系统变量之间的关系, 比如线性关系等;

d) 样本特征值分析: 用于比较两组数据的吻合程度, 样本特征值一般选取样本均值、样本标准差、最大值和最小值。

(2) 定量分析方法

这里使用置信区间法。置信区间法是仿真模型有效性定

量分析中的一种重要的统计检验法。在比较模拟数据序列与实际数据序列时, 采用假设检验法, 虽然能判断出两组数据有无差异, 但无法给出优劣程度大小的信息。而采用两组数据参量之差的置信区间估计方法时, 若置信区间包含零, 则可判定在给定置信度下两组数据性能无差异, 而且还可根据置信区间的大小来判断两组数据的优劣程度。若事先能给定仿真模型可接受的精度范围 ARA, 且计算出来的置信系数为 α 的置信区间满足 ARA 要求, 则可认为模型是有效的, 且模型的可信度为 $Pa = 100(1 - \alpha)$ 。

具体方法采用工程上较适用、可靠的韦尔奇置信区间估计法。

假设模拟数据序列 X_m 和实际数据序列 X_s 是独立同分布的, 它们的总体方差未知, 且认为它们是不相等的。则置信度为 $(1 - \alpha)$ 的置信区间为:

$$\bar{x}_1(n_1) - \bar{x}_2(n_2) \pm t_{\alpha/2, \nu} \sqrt{\frac{S_1^2(n_1)}{n_1} + \frac{S_2^2(n_2)}{n_2}} \quad (3)$$

式中, $\bar{x}_1(n_1)$ 、 $\bar{x}_2(n_2)$ 为样本均值; S_1^2 、 S_2^2 为样本方差; $t_{\alpha/2, \nu}$ 为自由度为 ν 的 t 分布的 $\alpha/2$ 分位数, 按下式计算:

$$\nu = \frac{[S_1^2(n_1) + S_2^2(n_2)]^2}{\frac{S_1^2(n_1)}{n_1 - 1} + \frac{S_2^2(n_2)}{n_2 - 1}} \quad (4)$$

根据置信区间可得以下两点结论性意见:

a) 若置信区间中包含零, 则可判断在置信度 $(1 - \alpha)$ 下两组数据性能上无差异;

b) 若置信区间满足模型的可接受精度范围 ARA, 则可判定该模型是有效的, 且仿真模型的可信度 P_s 在 $(1 - \alpha)$ 之上。

4.2.2 不可观察的系统的输出验证

对于不可观察的仿真模型的验证, 主要使用特征事实法(stylized facts)这一理论验证方法, 特征事实主要指现实中存在的具有广泛的规律性的现象和事实, 如货币量的增加和通货膨胀相随发生的现象。

在不可观察的仿真系统中, 验证所需要的真实数据不易获得, 在社会科学领域, 如经济科学、金融科学、人工社会等, 所建立的仿真系统不是目标系统的完全复制, 宏观层次数据不可得、不易得, 因而对仿真系统的输出进行实证研究比较困难。但是, 社会科学领域中存在大量的可观察数据, 例如经济领域存在很多被普遍认可的现象或规律, 所以可以使用特征事实来对仿真系统进行输出评估。

5 校核与验证框架在 EEPSS 中的应用

EEPSS 模型在开发中已在概念模型的基础上完成了计算机的正确实现, 下面由以下两个步骤对建立的模型进行校核。

5.1 模型校核

首先, EEPSS 通过采用系统流程图、设置系统快照(以 XML 脚本的形式保存、输出系统在某个时刻所有主体的状态属性以及宏观变量如 GDP 等统计结果)来进行表面验证, 并可对仿真结果进行实时动态显示, 另外还通过对主体适应性学习算法的观察来验证模型的合理性。

其次, 确认模型运行的稳定性, 在当前模型中输入参数运行时间 5 年, 家庭个数 3000, 食品、非耐用品、汽车、房地产企业的个数分别为 10、4、6、4, 政策场景为现行排污收费政策 0。选取输出变量中的 GDP 增长率进行测试, 模型运行次数 $m = 5$, 求 $\bar{X}(5)$ 的样本方差 $S^2(5)$ 并据此判断模型的稳定性, 若样

(下转第 472 页)

[4] FLAME. Projects[EB/OL]. <http://www.flame.ac.uk/projects>

[5] Coakley S, Gheorghe M, Holcombe M, et al. Exploitation of High Performance Computing in the FLAME Agent-Based Simulation Framework[C]// IEEE Computer Society. 2012, 2012: 538-545

[6] Chin L S, Worth D J, Greenough C, et al. FLAME: an approach to the parallelisation of agent-based applications. RAL-TR-2012-013. [R]. Science and Technology Facilities Council, 2012

[7] Chin L S, Worth D J, Greenough C, et al. FLAME-II: a redesign

of the Flexible Large-scale Agent-based Modelling Environment. RAL-TR-2012-019. [R]. Science and Technology Facilities Council, 2012

[8] Park S, Zeigler B P. Distributing Simulation Work Based on Component Activity: A New Approach to Partitioning Hierarchical DEVS Models [C]// IEEE Computer Society CLADE. 2003:124-131

[9] 伍选, 文中华, 汪泉, 等. 多 agent 规划领域中的观察信息约简[J]. 计算机科学, 2014, 41(6): 176-179, 192

(上接第 463 页)

本方差值在用户指定的可接受范围之内, 则认为模型是运行时稳定的; 否则, 若样本方差过大则应检查模型的逻辑合理性、算法或修改模型的参数等, 调整模型反复进行测试。

5.2 模型验证

两组数据序列如下($n=8$):

模拟数据序列 $\{X_1\}$: 0.994570, 0.989843, 0.992003, 0.996662, 0.987338, 0.97718, 0.987935, 0.996443;

实际数据序列 $\{X_2\}$: 0.995421, 0.985820, 0.995916, 0.988551, 0.986710, 0.985244, 0.978873, 0.981318。

5.2.1 定性分析

(1) 数据分布图

如图 4 所示, 分布图中除了一个异常点以外, 其他的数据偏差不大。

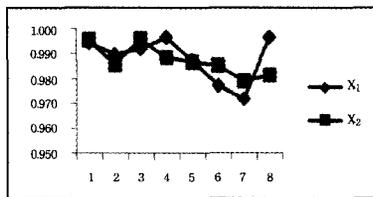


图 4 数据分布图

(2) 箱线图

模型数据范围包括实际数据范围, 且 model 的中位线高于 system 的第一分位线, 仿真模型无效, 如图 5 所示。这与总体结论相悖, 原因是数据太少, 箱线图要求数据至少在 10-30 个或者更多, 这样才能合理表现出统计特性。

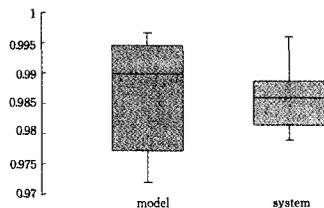


图 5 箱线图

(3) 样本直方图

如图 6 所示, 除了最后一组数据外, 其他各组数据的差别不大, 在 0.015 的范围内, 从而得出数据吻合得较好。

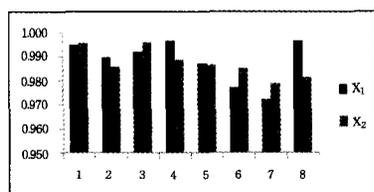


图 6 样本直方图

(4) 样本特征值分析

从表 2 可看出两组数据吻合得较好。

表 2

	$\{X_1\}$	$\{X_2\}$
样本容量 n	8	8
最大值 X_{\max}	0.996662	0.995916
最小值 X_{\min}	0.971935	0.978873
样本均值 \bar{X}	0.988239	0.987232
样本标准差	0.008554	0.005648

5.2.2 定量分析——置信区间估计

$$\bar{x}_1(n_1) = 0.988247$$

$$\bar{x}_2(n_2) = 8.34E-5$$

$$S_1^2 = 0.987232$$

$$S_2^2 = 3.65E-5$$

$$\nu = 12.14, \text{取 } \nu = 12, \text{查表得: } t_{\alpha/2, \nu} = 2.68099.$$

根据式 (4) 计算得置信度为 95% 的置信区间为 $(-0.00936, 0.011394)$, 分析置信区间可得出如下两点结论性意见:

a) 由于该置信区间包含零, 则可判断在 95% 置信度下两组数据性能无差异。

b) 假定该模型的可接受精度范围 ARA 为 $(-0.1, 0.1)$, 由于该置信区间满足 ARA 要求, 则可判断该模型是有效的, 且模型的可信度 Pa 在 95% 以上。

结束语 本文提出了一个基于 Agent 仿真模型的校核与验证框架, 包括具体的表面验证技术、灵敏度分析、模型校准与运行时验证, 并从定性和定量两个角度进行运行时验证, 适用于一般的 ABMS, 其缺点是没有恰当的例证。基于 Agent 模型的校核与验证是 Agent 仿真开发中关键的环节, 目前甚至在将来的一段时期内仍是 ABMS 领域的一大挑战。

参考文献

[1] 宋承龄. 关于仿真模型验证[J]. 计算机仿真, 2000, 17(4): 8-11

[2] 唐见兵, 查亚兵, 李革. 仿真 VV&A 研究综述[J]. 计算仿真, 2006, 23(11): 82-98

[3] 廖守亿, 陈坚, 陆宏伟, 等. 基于 Agent 的建模与仿真概述[J]. 计算机仿真, 2008, 25(12): 1-7

[4] Sargent R G. Verification and validation and simulation models [C]// Proceedings of the 2010 Winter Simulation Conference. 2010:166-183

[5] Janssen M A, Ostrom I. Empirically based, agent-based models [J]. Ecology and Society, 2006, 11(2)

[6] Ngo A, See L, A J. Heppenstall et al. Agent-based models of geographical systems[D]. Vietnam: Hanoi University of Agriculture, 2012