

复杂系统的计算实验方法

——原理、模型与案例

薛 霄 著

科 学 出 版 社

北 京

科学出版社
www.abook.cn

内 容 简 介

本书以计算模型的设计和分析为主线,从方法理论、计算模型和应用案例3个方面介绍计算实验方法的知识体系,为该方法的理论与具体应用提供指导。书中所覆盖的课题包括计算实验的方法框架与系统平台,基础的计算模型(如状态自动机、马尔可夫模型、系统动力学、博弈模型、排队论等),经典的计算模型(如元胞自动机、糖域模型、社会生态系统模型等),以及不同领域的应用案例。

本书适用于复杂系统、社会计算、计算机仿真、计算实验等领域从事教学科研的师生以及相关的专业人员;也可供政府相关部门的决策者,从事社会学、经济学、组织学和管理学等领域的相关人员参考。

图书在版编目(CIP)数据

复杂系统的计算实验方法:原理、模型与案例/薛霄著. —北京:科学出版社, 2020.9

ISBN 978-7-03-064374-2

I. ①复… II. ①薛… III. ①计算机辅助设计 IV. ①TP391.72

中国版本图书馆CIP数据核字(2020)第021806号

责任编辑:孙露露 王会明 / 责任校对:马英菊
责任印制:吕春珉 / 封面设计:东方人华平面设计部

科学出版社 出版

北京东黄城根北街16号
邮政编码:100717
<http://www.sciencep.com>

印刷

科学出版社发行 各地新华书店经销

*

2020年9月第 一 版 开本:787×1092 1/16

2020年9月第一次印刷 印张:19 1/4

字数:437 000

定价:79.00元

(如有印装质量问题,我社负责调换〈 〉)

销售部电话 010-62136230 编辑部电话 010-62135397-2010

版权所有,侵权必究

举报电话:010-64030229; 010-64034315; 13501151303

科学出版社
www.abook.cn

序：为什么要计算实验？

非常高兴看到薛霄教授的这本《复杂系统的计算实验方法——原理、模型与案例》专著，相当程度上，这是迄今为止我读过的关于复杂系统计算实验分析方面最为优秀的一部著作。相信专著的出版，一定会更进一步地推动相关的研究和应用，为解决复杂系统的设计、分析、管理、控制和综合等问题提供新的且更加有效的计算理论和手段。

众所周知，实验是从事科学研究所必备的手段。在经典物理时代，一项研究成果是科学的，首先必须有实验支撑，其次别人还必须能够重复这项实验，这是对科学成果最基本的要求。然而，这些实验一般是物理性的实体实验，对于复杂系统，往往无法进行。特别是涉及人和社会因素的系统，传统的实验方法更是难以实施，以至于许多社会学问题只能求助于“反事实实验（counterfactual experiments）”。这就是要进行计算实验的原因，本质上，就是将“反事实（counterfactual）”算法化（algorithmization），成为量化分析复杂系统的一种数字化和计算化方法。

为什么不能用传统的实验方法对复杂系统进行分析？最直接的是经济方面的原因：由于成本和时间等因素，代价太大，无法承受。就算解决了实验的经济问题，又遇到了法律方面的原因：由于国家和社会安全等因素，法律条文不允许进行实验。就算修改法律可以实验了，又面临道德方面的原因：由于社会文化和伦理等因素，在道德上无法接受实验。最后，就算有人或组织愿意承担生命财产的风险甚至牺牲，却发现无法分解还原复杂系统，就是说目前的科学理论及手段根本就完不成复杂系统的实验。因此，如果希望科学地分析复杂系统，必须另辟蹊径，计算实验应运而生。

计算实验是计算仿真的自然扩展与升华。不同之处在于数据驱动、涌现机制和多重世界的解释与引导理论。在计算机仿真的传统的描述和预测功能之外，计算实验强调结果的引导（prescription）功能，从牛顿时代的“大定律、小数据”，迈向默顿时代的“大数据、小定律”。经过十余年的发展，计算实验方法已成为分析复杂系统的主流方法之一，除代理技术之外，其基本人工系统、软件定义系统，特别是近来的数字孪生理念，已得到学界甚至是社会上的广泛认同，成为从小数据生成大数据，再从大数据提炼“小智能”的深度智能或精准知识的新兴手段，与人工智能方法相辅相成。

我自己关于计算实验的想法源于 1982 年从事断裂力学和材料损伤的实验研究，以及 1988 年和 20 世纪 90 年代初关于太空和外星无人系统的仿真要求，主要是经济和时间方面的考量。最早文献是 1982 年在浙江大学撰写的报告——《用计算模拟疲劳实验的设想》，以及 80 年代末和 90 年代初所写的英文 NASA 工作报告和公开发表的论文。2000 年前后，参加斯塔菲关于复杂系统研究的讨论会才使我真正地认识到计算实验方法和计算实验学必须作为一门独立于计算机仿真的科目单列出来。非常高兴的是，自从 2004 年发表《计算实验方法与复杂系统行为分析和决策评估》之后，计算实验得到空前的发展，甚至连平行仿真也被作为一种新的独立仿真方法提出并发展成长起来。

本书的完成，将使计算实验方法的完善与普及向前迈进一大步。特别令我高兴的是，作者从 2002 年到 2007 年在中国科学院自动化研究所硕博连读，与我的研究生关系密切，对当时我们关于复杂系统的计算方法研究十分关注。他认为，这段时间和经历对他的科研之路起到了非常重要的启蒙作用，激发了他对复杂系统与计算实验的浓厚兴趣。毕业工作后，坚持不懈，长期钻研，终于有现在的成果。作为一名教师，这是最令自己欣慰的成就。

从多 Agent 软件开发方法，到人工社会模型的构建，再到计算实验系统方法的基础理论，本书是薛霄教授对他前期工作的一个总结与回顾。在此，他以计算模型的设计和分析为主线，从原理、模型和应用三个方面介绍了计算实验方法的知识体系，提出了他自己对计算实验方法的许多深刻的思考，相信可以为该领域的研究者和应用者提供许多有益的启发和帮助。

特别令我期望的是，在与薛霄教授就本书进行交流的过程中，我们对如何加强计算实验方法的实验设计问题有了许多共识。第一次交谈之后，薛霄教授就立即找来我凭记忆告知的几本参考书，并决心在相关方向上深入研究。我非常愿意尽自己的力量相助，使计算实验能够在薛霄教授及其团队的努力之下，理论更加完备，方法更加齐全，工具更加有效，应用更加普及！

卡尔·雅斯贝思曾说：“人如果不想垮掉，就必须面对虚无。”计算实验就是一门将“虚无”科学地变为“现实”的学问。所以，“人如果希望准确，就必须计算实验”！

王飞跃

怀德海学院

中国科学院自动化研究所

复杂系统管理与控制国家重点实验室

2019 年 12 月

前 言

近年来，随着互联网的普及、物联网的渗透、大数据的涌现以及社会媒体的崛起，产生了越来越多的工程复杂性与社会复杂性高度融合的具有不定性、多样性和复杂性特征的社会物理信息系统（cyber physical social systems, CPSS）。CPSS 的复杂性主要源于智能群体间的协调网络以及所受到的外部干预，其建模、有效控制和管理等问题正面临前所未有的挑战和发展机遇。计算实验为复杂系统的分析、推理、预测和控制提供了一种新颖的、可行的研究工具和手段，目前已经成功应用于很多领域，包括计算组织学、计算社会学、计算流行病学、计算经济学等。

20 世纪末，复杂性科学兴起。这一新兴科学不仅是系统科学的前沿，而且是整个科学体系的前沿，因此有人称其为 21 世纪的科学。面对生命系统、互联网系统、社会系统、经济系统等复杂的系统，复杂性科学放弃了还原分解的方法，选择了一种自下而上的系统集成方法。复杂性科学强调复杂系统是由大量自治个体构成的，其基本思路就是通过观察现实的复杂系统，抽象出每个个体所遵循的简单规则，构建出这些简单个体的模型，并让它们相互作用，从而观察整体作用的涌现结果，最后通过观察涌现属性找到规律，并用这些规律解释和理解现实系统中的宏观现象。

自然科学非常强调科学实验，然而对于很多复杂系统而言，由于经济、法律、伦理等方面的限制，传统实验方法几乎不可能实现。在这种背景下，计算实验方法开辟了一条认识、理解和干预复杂系统的新路。基于计算实验，人们可以研究现实系统在信息世界中的“数字孪生”，通过修改系统所遵循的规则、参数，从而进行各种各样的系统实验。短短几十年的发展，复杂性科学已经成为一个热门的研究领域，计算实验方法已经广泛应用于经济学、社会学、生态学、环境学、组织理论、语言起源、文化传播等广阔的领域。

计算实验方法作为复杂性科学研究的一种主要手段，集成了计算机科学、社会科学、系统科学、计算机模拟技术、多 Agent 系统技术、人工智能技术等诸多学科。尽管科研人员付出了很多努力，目前计算实验方法还不能算是一套成熟的体系，也没有一套完整成形的理论，其理论发展和实际应用之间还存在着巨大的差距。为了能够对解决该问题有所帮助，本书梳理了计算实验方法的来龙去脉，希望能够帮助读者构建起计算实验方法的完整知识体系，从而为后续的方法应用奠定坚实的基础。本书的内容包括方法理论、计算模型和应用案例三个方面，重点回答三个问题：计算实验方法的系统框架是什么、计算实验方法的核心技术是什么，以及计算实验方法如何实现领域应用。

本书的内容基于我们多年的研究成果。在研究过程中，先后得到了天津大学 2019 年“研究生创新人才培养项目”（YCX19057）、国家自然科学基金项目“基于知识图谱的云制造服务跨界融合关键技术研究”（61972276）、国家自然科学基金项目“社会移动云环境下的 App 自主协同及个性化混搭机制研究”（61572350）、国家重点研发计划“健康养老跨界服务应用示范”（SQ2017YFB1401200）、河南省科技创新杰出青年支持计划“企业协同服务系统的自主进化机制研究”（174100510008）的支持。在这些课题的支持下，我们进行

了有关计算实验方面的研究和开发工作，取得了一些初步的成果，希望通过本书与广大科技人员和读者交流与共享，从而推动计算实验方法与其他科学领域的交叉融合与发展。

需要强调的是，本书凝聚了计算实验领域很多研究人员的工作成果，在此感谢他们为计算实验方法的发展所做出的显著贡献。特别感谢天津大学智能与计算学部的冯志勇教授与陈世展副教授在百忙之中抽出时间审阅了全书，提出了很多有价值的修改意见，这对于全书的形成至关重要。尤其要感谢中国地质大学（北京）的研究生周德雨，全程参与了本书的撰写、修订、校对和作图等工作，为书稿的最终出版付出了非常多的时间和精力，祝愿她早日学业有成。

此外，本书的完成还要感谢天津大学服务计算与数据科学实验室的研究生徐砚伟、陈红旗、张璐、张瑜芳、陆敏、杨玉莹、侯佛之、王渐旭、周显、王颖、于慧雪、汪步鹏和冯轶然，他们为本书的勘误付出了大量的时间和精力；感谢我的历届研究生高佳佳、郭耀丹、皇甫帅、陈召杰、王纪才、韩红芳、施曼等，他们为本书的实验设计和代码实现进行了大量辛苦的工作。在撰写本书的过程中，我们参考了大量国内外相关文献，在此对每位专家学者表示感谢。

由于作者水平有限，书中难免存在疏漏、错误与片面性，恳请读者不吝赐教。

薛 霄

天津大学智能与计算学部

天津大学新媒体研究院

2019年10月

目 录

第 1 部分 方法理论篇

第 1 章 导言	3
1.1 复杂性科学	3
1.1.1 复杂系统的概念	3
1.1.2 复杂性科学研究的历史阶段	7
1.2 计算实验方法	13
1.2.1 计算实验的起源	13
1.2.2 计算实验的应用	15
1.3 本书的组织结构	18
1.4 重要的经典文献	20
1.5 小结	22
参考文献	22
第 2 章 计算实验的理论基础	25
2.1 引言	25
2.2 复杂性的实验分析	25
2.2.1 计算实验的目的	25
2.2.2 计算实验的模式	26
2.2.3 计算实验的质量维度	28
2.3 系统的复杂性的规律	29
2.3.1 复杂性的成因分类	29
2.3.2 结构复杂性的规律分析	31
2.3.3 分布复杂性的规律分析	34
2.4 系统复杂性的解释	38
2.4.1 系统复杂性理论：解释	38
2.4.2 基于顺序逻辑的建模解释	39
2.4.3 基于反向逻辑的建模解释	41
2.5 小结	43
参考文献	44
第 3 章 计算实验的方法框架	45
3.1 引言	45
3.2 现实建模	46
3.3 系统开发	48
3.4 实验执行	50
3.5 结果分析	51

3.6 模型评估	53
3.7 结果应用	54
3.8 小结	55
参考文献	56

第 2 部分 计算模型篇

第 4 章 基于数学变量的模型	59
4.1 引言	59
4.2 系统动力学	59
4.2.1 系统动力学概念	59
4.2.2 描述模型	61
4.2.3 模型示例	64
4.2.4 分析	66
4.3 博弈模型	67
4.3.1 博弈的概念	67
4.3.2 博弈中的行为推理	69
4.3.3 最佳应对与严格占优	73
4.3.4 动态博弈与进化博弈	75
4.4 排队模型	78
4.4.1 排队系统概述	78
4.4.2 单服务台排队系统	81
4.4.3 单级多服务台排队系统	84
4.4.4 多级多服务台排队系统	85
4.4.5 其他类型的排队系统	86
4.5 马尔可夫模型	88
4.5.1 基本概念	88
4.5.2 系统建模	88
4.5.3 模型示例	89
4.5.4 隐马尔可夫模型	90
4.6 小结	93
参考文献	93
第 5 章 元胞自动机模型	95
5.1 引言	95
5.2 相关文献	96
5.3 元胞自动机的概念	97
5.3.1 定义	97
5.3.2 研究问题	98
5.3.3 概念和形式模型	99
5.3.4 元胞自动机的实现	105

5.3.5 分析	107
5.4 谢林隔离模型	107
5.4.1 背景介绍	107
5.4.2 模型设计	108
5.4.3 隔离实验	109
5.5 康威生命模型	112
5.5.1 人工生命	112
5.5.2 康威生命游戏	113
5.6 小结	115
参考文献	116
第 6 章 多 Agent 系统模型	118
6.1 引言	118
6.2 相关文献	119
6.3 研究问题	120
6.4 概念模型设计	120
6.4.1 Agent	121
6.4.2 组织	123
6.4.3 环境	123
6.4.4 规则	124
6.4.5 定制	124
6.5 多 Agent 系统实现	126
6.5.1 如何感知	126
6.5.2 如何决策	129
6.5.3 如何行为	130
6.5.4 如何交互	134
6.5.5 如何协作	135
6.6 小结	137
参考文献	137
第 7 章 人工社会: SugarScape	139
7.1 引言	139
7.2 基础的糖域模型	140
7.2.1 简单的糖域模型	140
7.2.2 糖域内的人工社会	142
7.2.3 Agent 种群内的财富分布	145
7.2.4 Agent 波浪的移动	148
7.2.5 邻居的社交网络	149
7.3 性别、出生与死亡	151
7.3.1 有性繁殖	152
7.3.2 生命进化论	154
7.3.3 人口的变化	156

7.4	文化与冲突	158
7.4.1	文化过程	158
7.4.2	朋友网	162
7.4.3	季节迁徙	163
7.4.4	战斗	164
7.5	糖与香料的交易	169
7.6	小结	173
	参考文献	173
第 8 章	社会生态系统：SLE 模型	175
8.1	引言	175
8.2	社会学习演化模型：SLE	176
8.2.1	SLE 框架	176
8.2.2	个体进化	177
8.2.3	组织演化	179
8.2.4	社会演化	180
8.3	社会生态的实验系统	182
8.3.1	总体框架设计	182
8.3.2	供应侧设计	184
8.3.3	需求侧设计	187
8.3.4	供需匹配设计	189
8.4	社群化制造的用例分析	190
8.4.1	社群化制造的概念	190
8.4.2	计算实验的参数初始化	192
8.4.3	计算实验案例	193
8.5	讨论	197
8.6	小结	199
	参考文献	200

第 3 部分 应用案例篇

第 9 章	互联网的跨界影响力分析	205
9.1	引言	205
9.2	互联网影响评估方法（服务桥）	206
9.2.1	“服务桥”模型	206
9.2.2	供应侧的能力模型	208
9.2.3	需求侧的特征模型	209
9.2.4	“服务桥”的工作流程	210
9.3	计算实验系统构建	212
9.3.1	实验系统的运行机制	212
9.3.2	需求侧 Agent 模型的构建	213

9.3.3	供应侧 Agent 模型的构建	213
9.4	互联网对 O2O 产业的影响分析	215
9.4.1	案例研究	215
9.4.2	实验参数设置	217
9.4.3	实验过程分析	219
9.5	讨论与预测	221
9.5.1	互联网对销售行业的影响	221
9.5.2	互联网对金融行业的影响	222
9.5.3	互联网对制造行业的影响	223
9.5.4	互联网对医疗行业的影响	223
9.6	小结	224
	参考文献	224
第 10 章	新零售的跨界竞争力分析	227
10.1	引言	227
10.2	相关工作	228
10.2.1	跨界服务的概念	228
10.2.2	跨界服务分析方法	229
10.2.3	跨界服务的应用领域	229
10.3	基于价值链的跨界服务分析框架	230
10.4	新零售案例研究	232
10.4.1	传统电子商务的运作模式	232
10.4.2	传统电子商务的能力分析	233
10.4.3	新零售商实施策略	235
10.5	新零售实施策略的实验分析	236
10.5.1	计算实验系统设计	236
10.5.2	平台模式的策略实验	238
10.5.3	自营模式的策略实验	240
10.6	讨论	243
10.7	小结	244
	参考文献	244
第 11 章	O2O 服务推荐策略的影响因素分析	247
11.1	引言	247
11.2	相关工作	248
11.3	基于反馈机制的 O2O 服务推荐框架	249
11.4	服务推荐策略设计	251
11.4.1	基于服务节点的服务策略 (Service-CF)	251
11.4.2	基于服务路径的服务策略 (Path-CF)	253
11.4.3	基于持续反馈的服务策略 (O2O-CF)	255
11.5	构建 O2O 服务推荐计算实验模型	256
11.5.1	系统环境模型	256

11.5.2	服务 Agent 模型	257
11.5.3	用户 Agent 模型	258
11.6	计算实验结果	259
11.6.1	计算实验的初始化	259
11.6.2	计算实验案例	261
11.7	小结	265
	参考文献	265
第 12 章	智慧矿山应急服务的实验研究	268
12.1	引言	268
12.2	研究背景和动机	268
12.2.1	研究背景	268
12.2.2	研究问题	270
12.3	基于计算实验的情境感知服务评估方法	271
12.3.1	情境感知服务策略的定制	271
12.3.2	计算机实验系统的构建	272
12.3.3	服务策略效果评估	273
12.4	煤矿环境感知服务的实现	274
12.4.1	不同的报警服务实现策略	274
12.4.2	不同的疏散服务实施策略	275
12.5	计算实验系统建模	277
12.5.1	系统实体模型	277
12.5.2	系统运行模型	277
12.6	服务策略实验评估	279
12.6.1	服务策略实验平台简介	279
12.6.2	服务策略实验案例	280
12.7	分析和讨论	286
12.8	小结	286
	参考文献	287

第 4 部分 总结与展望篇

第 13 章	总结与展望	291
13.1	总结	291
13.2	展望	292
	参考文献	295

第 1 部分

方法理论篇

第 1 章 导 言

1.1 复杂性科学

1.1.1 复杂系统的概念

“复杂”是诞生于秩序与混沌边缘的科学，是一场激动人心的科学观念的革命，“复杂”的原则和思想影响深远，以至于绝大多数学科都卷入其中。在科学发展进程中，人们曾经把宇宙看作是一架机械的钟表，认为只要明白了支配宇宙的规律，就能推察过去，预言未来。拉普拉斯（P.S.Laplace）曾经写道：“假如有一位智者在一给定时刻都能洞见所有支配自然界的力和组成自然界的存在物之间的相互位置，假如这一智者的智慧巨大到足以对自然界的所有数据进行分析，他就能将宇宙最大天体和最小原子的运动统统纳入单一的公式之中。对这样的智者来说，没有什么是不能确定的，未来同过去一样都历历在目。”

经典科学基本上是指由培根（F.Bacon）、牛顿（I.Newton）、伽利略（G.Galilei）、笛卡尔（R.Descartes）等开创的，近三四百年内发展起来的一整套观点、方法、学说。由培根首先倡导的分析和实验方法，经过伽利略等许多学者的实现与发展，深入到几乎所有的学科领域；牛顿归纳的物理学的定律以及相应的数学方法，则成为人类对于宇宙（包括社会）认识的基本框架；而笛卡尔则从思想方法上进行了深刻的概括与总结。工程技术正是在由此形成的庞大、完整的近代科学体系的基础上形成的，并成为创建近百年来人类文明空前繁荣的基石。

人们曾经很陶醉，以为已经认识了世界，用数学给出了世界完美的描述。但是，20世纪伊始，相对论和量子力学的出现就对牛顿力学所代表的宇宙观提出了尖锐的挑战。很快人们就发现这个世界其实很复杂，在很多领域，诸如物理学、化学、生物学、信息科学、社会学、经济学、城市科学等领域，具有与简单系统不同的特性，如开放性、非线性、混沌、自组织、涌现等特性，充满了不确定性。系统的发展不仅与其初始条件紧密相关，还与其发展过程中的偶然事件密切相关，初始条件的微小变动就会导致最终的结果失之毫厘、谬以千里。这个时候，人们开始意识到：近代科学的思想框架已经无法容纳人类对世界越来越广泛、越来越深入的认识。

德国心理学家德尔纳（D.Derner）曾经提出一个问题：假定人们具备所有的智能、经验和信息条件，为什么仍然会犯错误，有时甚至引起灾难性的后果。在著作《失败的逻辑》中，德尔纳给出了自己的答案^[1]：人们思维模式中的某些倾向适合于过去的简单世界，但对于现在所生活的复杂世界却有着灾难性影响。当今世界，一切事物都是相互关联的，不能一次只做一件事情，因为每件事都有多重结果；不能用孤立的因果模式考虑问题，因为所有的情况都存在副作用和长远影响。

德尔纳找到了不少例子，例如为什么阿斯旺水坝的规划者们，只想到会给埃及带来廉

价用电的好处，却没有意识到他们也将中断一年一次的洪水漫灌，而这维持了尼罗河谷地几千年来的肥沃富饶？为什么人们在开垦荒地，希望能够增加粮食产量的同时，却没有意识到对生态系统的保护的重要性，最终导致土地荒漠化与粮食的颗粒无收？这些都说明：在未了解一个复杂系统中所有连锁因素之前就采取行动，即使初衷很好，也难免铸成大错；面对力不能及的问题，小错误累积起来，最后就会酿成严重的错误；过于频繁地忽视问题的大局，却只在局部范围内寻求权宜之计，这只能是杯水车薪、无济于事。

为了尽可能地避免失败，就必须去认识这个世界的复杂性，了解整个复杂系统背后运行的规律到底是什么。阿肯巴克（J.Achenbach）指出：“科学正逐渐走出易解问题构成的领域，开始接触真正难解的问题。科学已达到了一个新的转折点。”20世纪80年代中期以来，一些有远见的科学家就已开始探索这一新的转折点，一门被誉为“21世纪的科学”的学科——复杂性科学，应运而生。

1999年4月，美国 *Science* 杂志出版了一期以“复杂系统”为主题的专辑，这个专辑分别就化学、生物学、神经学、动物学、自然地理学、气候学、经济学等领域中的复杂性研究进行了报道。由于各学科对复杂性的认识和理解都不一样，所以该专辑避开术语上的争论，采用了“复杂系统”这个名词。如图 1-1 所示，复杂系统都有一些共同的特点：系统与外部环境进行物质、能量、信息的交换，在外部激励和内部驱动的共同作用下，不断发展演化，甚至发生突变。其复杂性主要表现为以下几个方面^[2]。

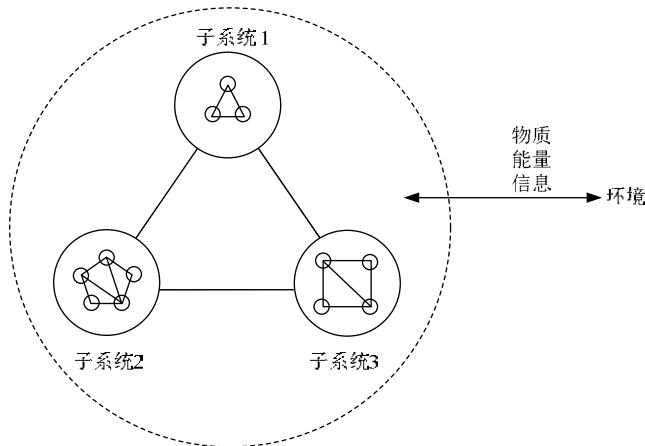


图 1-1 复杂系统示意图

1. 开放性

复杂系统与外部环境存在输入/输出接口，能够进行物质、能量、信息的交换，引入负熵，使得系统熵为负值，系统逐步进入有序状态。通过引入新的资源，系统调整自身的结构和行为，以适应外部和内部的变化，推动系统演化发展。系统内部层次众多、结构错综、关系复杂，存在不均匀性和不可预测性，导致系统远离平衡状态，具有耗散结构特性。当复杂系统远离平衡状态后，将继续开放式地与外界发生交换，内部自组织调整，进而不断趋于有序、稳定的状态。

2. 自组织

复杂系统中，其内部主体具有自治性，能够发挥主观能动作用和形成自我学习机制。与外界交换资源时，无须外界指令，可以利用自身的主动调整、自组织调解等方式，向有序、稳定的状态演化。通过竞争与协同实现资源优化配置、流程治理再造、组织架构完善、系统升级提升，达到系统有序化、优化的目的。

3. 非线性

在复杂系统中，信息海量、分散且流动，业务流程错综复杂，各主体层次众多且相互关联。因此，其信息关系、业务关系、组织关系、子系统之间、系统与外部环境之间等关系均不是清晰的线性关系，而是非线性关系。非线性作用使得系统内部相互制约，系统内部与系统外部环境协调同步，从而使系统自组织并与外部环境自适应，在各种条件下可能会引发无序（混沌）和有序（规则）的转化问题。

4. 涌现性

系统内部子系统和元素通过多种模式交互，并与外部环境存在资源交换，从而改变了内部结构和系统行为。在外界各种信息、物质等进入系统内部后，由于内部各元素之间的不断碰撞反应，加入新的外部物质后，内部结构、内部物质与外部物质之间可能将发生从微观到宏观的质变。一定时间后，经过放大和发展，在整体上演化出新的形式和特性，产生了涌现性。涌现性体现了宏观和微观的关系，是复杂系统微观行为产生的宏观效应。

5. 层次性

组成成分的多样性和差异性造成组分之间相互关系的多样性和差异性，是系统复杂性的根本源泉。对于产生复杂性，结构效应比规模效应更重要。因为组分的差异越大，把它们整合起来的难度就越大。特别地，等级层次结构是复杂性的主要根源之一，每一层次均构筑其上一层的单元，同时也有助于系统某一功能的实现。在系统变化无常的活动背后，呈现出某种捉摸不定的秩序，应当承认并区分不同层次上的复杂性。复杂性科学试图探索掌握不同层次的复杂系统活动背后的东西。

复杂系统为适应外界环境变化和自身结构行为调整，在系统内部以及系统与外部环境之间发生各种交互作用。系统的自组织、自适应、自聚集、自加强、自协调等特性，使得系统呈现复杂性，不断推动系统功能和结构不断演化。复杂系统之所以复杂，主要源于如下原因。

(1) 相互作用（关联）

事物的发展不是孤立的、割裂的、互不联系的，而是相互联系、相互作用、相互制约的统一整体。组织的特性并不只取决于单体的特性，更取决于单体如何相互关联形成组织。因为这类系统共同的特点是长程关联，关联往往导致 $1+1>2$ 或 $1+1<2$ ，也称为非线性。市场中出现的价格是受网络相互作用导致的，就像人们也都受到邻居的影响那样。相互作用使得物理系统无法轻松地由整体拆成部分，非线性即其根源。

一旦这种相互作用的维度增加，就会产生一个新的现象，即混沌。混沌是指由于系统

内自由维度的增加，系统的动力学属性不再归于闭合轨道，而是开放或成为不可预测的轨迹。初始条件的轻微变化在未来的影响远未可知。

相互作用导致协同效应。两个人相互作用的结果可以是 $1+1>2$ ，也可以是 $1+1<2$ ，但基本不会是 $1+1=2$ ，前两者可以看作非线性的体现。例如，为什么会有公司产生，那一定是某种合作导致 $1+1>2$ 的效应使得公司可以产生。

(2) 循环反馈

复杂系统多用于描述一个系统的时间变化过程，例如市场价格的波动、神经网络随时间的活动等。研究这个时间变化过程，往往要考虑此刻的结果对下一刻系统输出的影响。股市的反身性就是反馈机制的一种。

依据琼斯（S.Jones）对复杂系统涌现机制的划分，在系统整体和个体行为之间具有前馈和反馈机制（feedforward and feedback）的涌现称为二阶涌现（second order emergence），只有单向反馈机制的涌现称为一阶涌现。具有社会复杂性的信息系统中的涌现都属于二阶涌现，如图 1-2 所示。在复杂信息系统自组织下，个体行为因受到系统反馈而产生新的个体行为，复杂信息系统通过汇总涌现，在原有的一阶涌现基础上，引入了影响个体行为的反馈，从而形成信息系统的闭环回路，实现了复杂信息系统的双向反馈，构成了反馈环，增加了信息系统的复杂度。

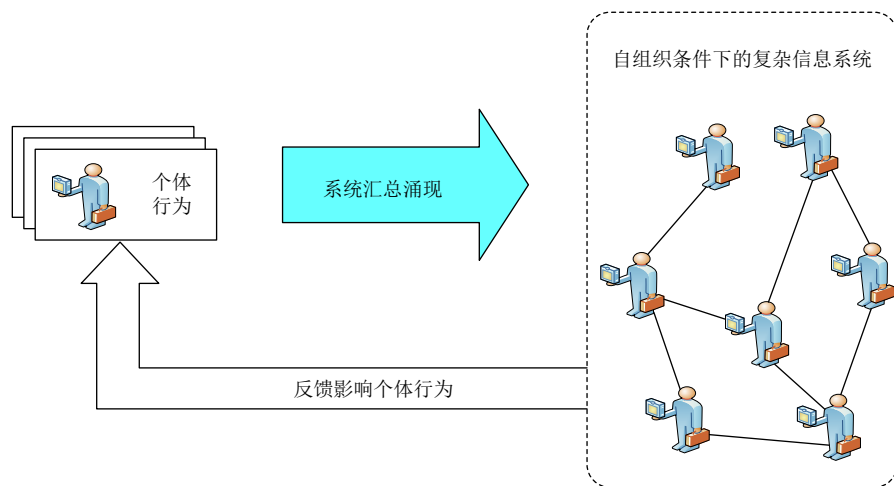


图 1-2 二阶涌现示意图

反馈分为正反馈和负反馈两种：正反馈导致不稳定性，如雪崩、股市崩盘等；负反馈导致定点平衡态。在所有复杂系统中，都有正反馈和负反馈两种状态。反馈带有回路的概念，一个单元通过相互作用传递给另一个单元，反过来另一个单元又可以把信息传递回来。反馈往往是指此刻活动对下一刻活动的影响，如市场价格永远围绕着均衡波动，价格高，导致市场中消费的人数变少，消费人数减少后又导致价格降低，这是典型的负反馈。负反馈把系统维持在稳定位置，通过 $dx = -x$ 表示。

(3) 临界相变

临界相变是复杂系统的第三个重要特质，而且是组织形成的核心。当系统主导反馈的性质发生变化时，则经历一个相变。相变在自然界和社会中无处不在，自然界中的相变既

包括冰和水之间的转化，也包括磁铁从一种相到另一种相的变化。而相变，就是通过外部变量使整个系统从一个相到达另一个相的过程。

相变理论是复杂系统研究的重要对象。以水到冰的相变为例，可以控制的外部变量就是温度。温度越高，熵越大。当温度为 0°C 时，系统自由度最小的状态是一致有序的态；随着温度升高，无序的态的自由度逐渐减少，直到到达某个点，成为更具有优势的态。在某个温度上，无序和有序交替，这个温度就称为临界。所谓临界，就是相变时候的状态，因为这个时刻最为特别。临界点上的系统属性特别复杂，大部分和人们息息相关的系统事实上都在某种程度上处于临界态（或靠近临界态），包括大部分生物系统和经济系统。涌现性和相变点也有着千丝万缕的联系，有兴趣的读者可以关注自组织临界（self organized criticality）理论。

1.1.2 复杂性科学研究的历史阶段

复杂性科学有三个主要特点：①研究对象是复杂系统，涉及自然现象、工程、经济、管理、军事、政治和社会等领域，从一个元胞呈现出来的生命现象到大脑的结构及心智、股票市场的涨落、社会的兴衰及人体的免疫系统；②研究方法需定性判断与定量计算相结合，微观分析与宏观分析相结合，还原论与整体论相结合，科学推理与哲学思辨相结合；③研究深度不限于对客观事物的描述，而是着重于揭示客观事物构成的原因及其演化的历程，并力图尽可能准确地预测其未来的发展。作为一门交叉学科，复杂系统的研究大致分为以下几个历史阶段，如图 1-3 所示。

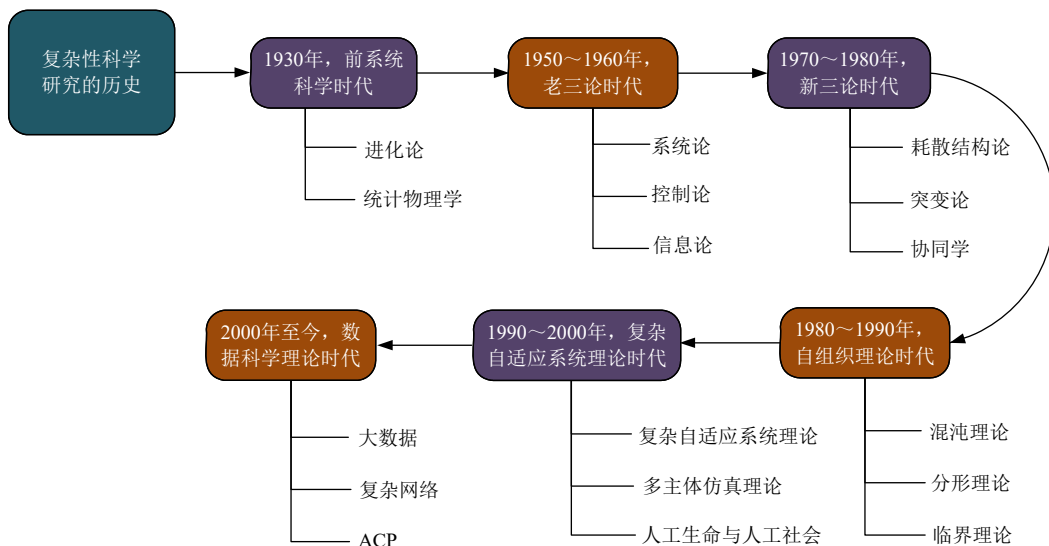


图 1-3 复杂性科学研究的历史脉络图

1. 前系统科学时代（1930年）

前系统科学时代的代表性理论主要包括英国生物学家达尔文（C.R.Darwin）的进化论^[3]和奥地利物理学家玻尔兹曼（L.E.Boltzmann）的统计物理学^[4]。1859年，达尔文在《物种起源》^[3]一书中提出了物竞天择、适者生存的生物进化思想，核心内容为：生物界的

个体之间存在差异,面对复杂多变的自然环境,适者生存,不适者被淘汰,且种群的有利特征通过世代遗传和变异保留下来。这是人们首次从群体和系统的视角观察生物群体的进化现象。

玻尔兹曼是热力学和统计物理学的奠基人之一,提出了玻尔兹曼 H 定理、热力学第二定律和著名的玻尔兹曼熵公式。通过分析大量的观察结果,热力学第二定律的规律总结如下:在孤立系统中,体系与环境没有能量交换,体系总是自发地向混乱度增大的方向变化,总是使整个系统的熵值增大,即熵增原理。这个定律描述了自然系统的演化趋势,同时给出了一种评估系统复杂度的指标。

2. 老三论时代(1950~1960年)

(1) 系统论

系统论(general system theory)是由美籍理论生物学家贝塔朗菲(L.V.Bertalanffy)创立的^[5]。1937年,贝塔朗菲提出了一般系统论,奠定了系统论的基础。1945年,贝塔朗菲将一般系统论的相关原理以论文的形式公开发表^[6]。1949年,贝塔朗菲出版了著作《生命问题》,该书的出版标志着系统论的正式诞生^[7]。1968年贝塔朗菲的专著《一般系统论——基础、发展和应用》发表,该著作被认为是系统论的代表^[5]。

按照系统论的观点,任何系统都不是各部分的简单组合,而是一个有机的整体,具有整体性、关联性、层次性和统一性等特征。一般系统论的提出是为了研究不同系统的共同特征和规律,主要包括三个方面:①“系统科学”,又称数学系统论,采用精确的数学语言描述系统,研究适用于一切系统的根本学说;②“系统技术”,又称系统工程,用系统思想和系统方法研究工程系统、生命系统、经济系统和社会系统等复杂系统;③系统哲学,研究一般系统论的科学方法论性质,并将其上升到哲学方法论的地位。

20世纪40年代,美国贝尔电话公司首次提出了“系统工程”这一概念,并采用系统工程的思想和方法研究微波通信网络。1957年,美国密歇根大学的古德(H.H.Goode)和麦考尔(R.E.Machol)合作出版了第一部以《系统工程》命名的著作^[8]。

(2) 控制论

1948年,维纳(N.Wiener)发表了著名的《控制论——或关于在动物和机器中控制和通信的科学》一书^[9]。在书中,维纳创造性地提出一个英语新词“Cybernetics”命名控制论科学,他认为控制论是一门研究机器、生命社会中控制和通信的一般规律的科学,更具体地说,是研究动态系统在变化环境条件下如何保持平衡状态或稳定状态的科学。

控制论自20世纪40年代末诞生至今,经历了三个发展时期^[8]:①经典控制论时期(20世纪40年代末至50年代),该阶段主要研究单输入和单输出的线性控制系统的一般规律,其主要应用于军事与工业生产领域的各种自动调节系统;②现代控制论时期(20世纪60年代),主要研究多输入和多输出的非线性控制系统,广泛应用于军事与工业生产和科研领域,如导弹系统、人造卫星系统、航天系统等;③大系统理论时期(20世纪70年代至今),该时期控制论已经从工程控制领域向生物、经济和社会等领域渗透,主要研究对象发展为各种复杂的控制系统,如宏观经济系统、生态和环境系统、能源系统等。

(3) 信息论

信息论(information theory)是一门采用概率论和随机过程研究信息传输和处理中一般

规律的学科。1948年,美国数学家香农(C.E.Shannon)在《贝尔系统技术学报》上发表论文《通信的数学理论》^[10];1949年发表《噪声中的通信》^[11],为信息论奠定了基础。信息论的主要内容包括信息熵、信源编码、信道编码、信道容量等。信息熵的提出是香农的伟大贡献之一,其主要具有三个性质^[12]:①单调性,即事件的发生概率越高,所携带的信息量越少;②非负性,即信息熵不能为负;③累加性,即多随机事件同时发生存在的总不确定性的量度可以表示为各事件不确定性的量度的和。信息熵可以用来量化评估一个系统信息含量的多少,其计算公式为

$$H(X) = -\sum_{x \in X} p(x) \lg p(x) \quad (1-1)$$

3. 新三论时代(1970~1980年)

(1) 耗散结构论

比利时物理学家普里戈金(I.Prigogine)在他出版的《从混沌到有序:人与自然的新对话》一书中提出了“复杂性科学”的概念^[13]。他认为“在经典物理学中,基本过程被认为是决定论的和可逆的。但是可逆性和决定论只适用于有限的简单情况,我们自己所处的世界却是由不可逆性和随机性占统治地位的。因此,物理学正在从决定论的可逆过程走向随机的和不可逆的过程。”普里戈金抓住的核心问题就是,经典物理学在静态的、简化的研究方式中从不考虑“时间”这个参量的作用,从而把物理过程看成是可逆的。实际上,普里戈金并没有提出一个明确的“复杂性”的定义,这里的复杂性理论主要是揭示物质进化过程中理化机制的不可逆性,即耗散结构论(dissipative structure theory)。

1969年,普里戈金在“理论物理学和生物学”国际会议上发表论文《结构、耗散和生命》^[14],首次正式提出了耗散结构的概念。耗散结构理论可以概括为:一个远离平衡态的非线性的开放系统,当不断地与外界交换物质和能量达到一定阈值时,系统可能发生突变即非平衡相变,由原来的混沌无序状态转变为一种在时间上、空间上或功能上的有序状态^[15]。

(2) 突变论

1972年,法国数学家托姆(R.Thom)出版著作《结构稳定性和形态发生学》,标志着突变论(catastrophe theory,也称灾变论)的诞生^[16]。20世纪70年代,英国数学家塞曼(E.C.Zeeman)对突变理论的发展应用做了进一步的研究。托姆将系统内部从某一稳定状态向另一稳定状态的整体性转变现象称为突变,突变论研究的是从一种稳定组态跃迁到另一种稳定组态的现象和规律,主要被用来认识和预测复杂的系统突变行为。

(3) 协同学

协同学(synergy theory),又称“协同理论”或“协和学”,是以系统论、信息论、控制论、突变论等多种理论为基础发展起来的。1969年,联邦德国物理学家哈肯(H.Haken)首次提出“协同学”这一概念,并于1971年与雷格厄姆(R.Graham)合作撰文对“协同学”进行了介绍。1972年,联邦德国召开了第一届国际协同学会议。1973年,随着《协同学》这一论文集的公开出版,协同学也宣布正式诞生^[17]。1977年,协同学进一步研究从有序到混沌的演化规律。1979年,协同学的研究对象扩大到了生物分子方面。目前,协同学在生物学、物理学、管理学、经济学、社会学等学科均有广泛应用。

协同学认为,系统是由大量的子系统组成的,各子系统之间相互关联、协同作用。协同学主要包括三方面内容^[18]:①协同效应,即大量子系统之间相互协同作用而产生的集体效应;②伺服原理,即系统接近不稳定状态或临界点时,序参量支配子系统行为;③自组织原理,即系统内部子系统之间能够自发地按照某种规则形成一定的结构或功能。

4. 自组织理论时代(1980~1990年)

(1) 混沌理论

很多专家学者认为混沌学是继相对论、量子力学之后的又一次物理学革命。1913年,法国数学家庞加莱(J.H.Poincaré)对微分方程解的灵敏性问题进行研究,得出了两个结论:①在某些情况下,“预测”是不可能的;②在确定性系统中,存在随机性现象。直到1963年,美国气象学家洛伦兹(E.N.Lorenz)偶然通过一个实验错误导致了混沌现象的发生,验证了庞加莱提出的初值敏感性。1975年,美籍华人李天岩(Li Tianyan)与他的导师约克(J.A.Yorke)在《美国数学月刊》上合作发表了论文《周期3蕴含混沌》,给出了著名的Li-Yorke定理,正式定义了混沌的概念^[19]。1976年,美国数学生态学家梅(R.May)将差分方程应用于生物种群的研究,采用形象的分支理论描述了Li-Yorke定理及混沌现象^[20,21]。

混沌是指系统内部看似随机的不规则运动。混沌理论(chaos theory)则指宇宙本身是处于混沌状态的,其中某一部分似乎并无联系的事件间的冲突,可能会给宇宙造成不可预测的结果。复杂系统内部也存在混沌现象,因此需要在系统内部看起来无规律、不可预测的现象背后,寻求混沌现象的规律。

(2) 分形理论

分形理论(fractal theory)是美国数学家曼德布罗特(B.Mandelbrot)创立的一门新几何学。1975年,曼德布罗特首次提出了分形的概念。1977年,曼德布罗特出版了第一本著作《分形:形态、偶然性和维数》,标志着分形理论的正式诞生^[22]。1982年,曼德布罗特出版了《自然界的分形几何》,分形理论初步形成^[23]。分形理论可以描述、计算和思考那些不规则、破碎、参差不齐和断裂的几何形状,包括从雪花的结晶曲线到星系中互不联系的尘埃。分形曲线意味着深藏在这些惊人的复杂的形状中的有组织的结构。分形最重要的特性是标度不变性,即适当放大或缩小几何尺寸,分形的整个结构并不改变。

如今,分形已经成为理解非线性动力学的关键结构,分形理论也已经成为自组织和复杂性理论的重要理论构成。分形概念提出虽然只有短短30多年的时间,但其发展之迅速却超出人们的想象。分形观念已经深入科学,扎根于社会,渗透到各个领域之中。海岸线、闪电、松花蛋、树枝、凝聚体、胶体、岩石、山脉、云彩、星系、粗糙的表面和界面、聚合物和股票市场,无不存在分形。

(3) 临界理论(critical theory)

1987年,巴克(P.Bak)、维森菲尔德(K.Wiesenfeld)和汤超(Tang Chao)提出“自组织临界性”的概念,采用著名的沙堆模型形象地说明自组织临界性的形成和特点^[24]。所谓自组织临界性指的是:一类开放的、动力学的、远离平衡态的、由多个单元组成的复杂系统,能够通过一个漫长的自组织过程演化到一个临界态;当系统处于临界态时,一个微小的局域扰动可能会通过类似多米诺骨牌效应机制被放大,其效应可能会延伸到整个系统,形成一个大的雪崩。临界性的特征是,处于临界态的系统会出现各种大小的“雪崩”事件,

并且“雪崩”的大小（时间尺度和空间尺度）均服从幂律分布。

5. 复杂自适应系统理论时代（1990~2000年）

美国圣塔菲研究所（Santa Fe Institute, SFI），1984年由三位诺贝尔奖获得者盖尔曼（M.Gell-Mann）、阿罗（K.J.Arrow）、安德森（P.W.Anderson）等倡议建立。它是一个专门研究复杂性的机构，由各学科的一流科学家参与，被誉为复杂性研究的世界中枢。其研究课题涉及地球上出现生命之前的化学演化和之后的生物演化、哺乳动物的免疫系统理论、人类与动物个体的学习和思维、人类文化和语言的演变、全球经济复杂的演化系统、计算机和程序设计的全新战略等。个人观点，圣塔菲研究所最具代表性的贡献在于提出了复杂性研究的计算机模拟方法，开辟了复杂性研究的新路径。其研究路径如下。

（1）复杂自适应系统理论

复杂自适应系统（complex adaptive system, CAS）是圣塔菲研究所霍兰（J.H.Holland）于1994年提出的一种复杂性理论，其基本思想可以用一句话概括：“适应性造就复杂性。”系统中的成员称为具有适应性的主体，简称为主体。所谓适应性，就是指它能够与环境以及其他主体进行交互作用。主体在这种持续不断的交互作用的过程中，不断地“学习”或“积累经验”，并且根据学到的经验改变自身的结构和行为方式。整个宏观系统的演变或进化，包括新层次的产生、分化和多样性的出现，新的、聚合而成的、更大的主体的出现等，都是在这个基础上逐步派生出来的。

霍兰在研究CAS时，就是在隐喻的基础上采用模型方法建构CAS的模型的，从而建立起复杂适应系统理论。他通过选择积木块和重组这些积木块的不同方法，建立起一些规则，用来创建易于理解的受某些规则支配的系统模型：刺激-反应模型。这个模型反映了CAS中具有主动性的主体的基本行为模型，即对个体是怎样适应和学习的理解和描述。以此为基础，霍兰建立整个系统的宏观模型，即回声模型（echo model）：主体在系统中可以寻找交换资源的其他主体，与其他主体进行资源和信息的交流。通过回声模型，霍兰清晰地解释了CAS（特别是基于计算机的CAS）的重要性质，探讨了CAS如何演化、适应、聚集、竞争、合作，以及如何创造极大的多样性和新颖性等。回声模型是使用很少的原理构建出极其优美的模型典范，为复杂性如何涌现和适应设定了一个路标。

（2）多主体仿真理论

多主体仿真（multi-agent simulation）理论是20世纪90年代中期以多主体系统理论和计算机仿真技术为基础发展而来的，是研究复杂系统的重要方法。多主体仿真中的主体（Agent）指具有一定的自治性、智能性和适应性的个体，多主体仿真的主要思想是通过建模模拟仿真复杂系统，主要流程如下^[25]：观察实际系统、建立概念模型、建立仿真模型、仿真运行和结果分析。学术界对多主体仿真本身的研究比较少，多是借助仿真工具实现多主体仿真。目前，应用比较广泛的多主体仿真工具包括Swarm、Repast、Ascape和NetLogo。这些工具各有优劣，需要根据具体情形进行选择。

（3）人工生命与人工社会

朗顿（C.Longton）开创了人工生命理论，推动了元胞自动机（cellular automata, CA）理论的发展，提出了复杂吸引子的概念。他认为在不动点吸引子、周期吸引子和奇怪吸引子之外还有一类吸引子为复杂吸引子，即所谓的混沌边缘状态，在这种状态下，系统表现

出永恒的新奇性。朗顿在“混沌边缘”概念的基础上，与其他学者一起建立了探索人工生命生成演化的各种模型，如自繁殖元胞自动机^[26]、鸟群（Boids）模型^[27]、蚁群模型^[28]、Tierra 模型^[29]、Avida 模型^[30]、“阿米巴世界”^[31]等。正是通过这些模型，朗顿等发现，生命的本质在于物质的组织形式而不在于具体的物质本身；如果在某种媒质中创造出产生混沌边缘的条件，那么就可能在這種媒质中创造出生命。

人工社会是在人工生命与仿真社会结合的基础上进行研究的。同人工生命一样，人工社会的核心方法是基于主体的建模、模拟和分析方法^[32]。人工社会的主体方法主要由三部分组成：主体、环境和规则。利用主体模型，芝加哥大学社会学研究所构建了一个糖域（SugarScape）模型模拟人工社会^[2]。人工社会拓展了社会学、人类学研究的手段，在虚拟的社会中可以观察战争、文化等因素的形成，可以检查各种社会科学的假设和方法，例如货币和市场经济的形成过程。

6. 数据科学理论时代（2000 年至今）

（1）大数据

大数据是指无法在一定时间内采用常规数据库工具对其内容进行获取、处理和分析的海量数据集合。大数据具有四个典型特征：①数据体量巨大，大数据的量级从 TB 级发展到 PB 级；②数据类型繁多，大数据的数据类型包括工程数据、网络日志、视频、图片、位置信息等；③数据蕴含的价值巨大，对大数据进行分析处理能够获取数据中包含的重要价值信息，如精准营销；④数据处理速度要求很高，要求实时给出分析结果，处理时间太长数据就失去了价值，如网上购票。大数据的处理需要依靠大数据处理平台，目前常用的大数据处理平台包括 Hadoop、Spark 等，大数据平台技术已成为现代建模仿真领域的重要支撑技术之一。采用大数据平台技术对复杂系统数据进行分析，有利于复杂系统规则的提取和发展预测。

（2）复杂网络

复杂网络是构成复杂系统的基本结构，具有自组织、自相似、小世界、无标度等部分或全部性质。每个复杂系统都可以看作系统各组成元素之间的相互作用网络，研究复杂网络的特性有助于了解复杂系统的性能。1998 年，美国康奈尔大学的博士生 D.J.Watts 及其导师在 *Nature* 杂志上发表了题为《“小世界”网络的集体动力学》的论文^[33]；1999 年，美国圣母大学教授 A.L.Barabási 及其博士生 R.Albert 在 *Science* 杂志上发表题为《随机网络中标度的涌现》的论文^[34]。两篇论文分别揭示了复杂网络的小世界特征和无标度性质，是复杂网络研究新纪元的开始。复杂网络主要包含四种网络模型：规则网络、随机网络、小世界网络和无标度网络。这四种网络模型在计算机、社会经济、交通等领域应用广泛。

（3）ACP

近年来，随着互联网的普及、物联网的渗透、大数据的涌现以及社会媒体的崛起，产生了越来越多的工程复杂性与社会复杂性高度融合的具有不定性、多样性和复杂性特征的社会物理信息系统（cyber-physical-social systems, CPSS）^[35]。其建模、有效控制和管理等问题正面临前所未有的挑战和发展机遇。王飞跃在 2004 年提出了一套利用计算手段研究复杂系统的平行系统理论和方法体系，即“人工系统+计算实验+平行执行”（artificial systems+computational experiments+parallel execution, ACP）方法。通过构造与实际系统对

应的软件定义模型——人工系统，采用在线学习、离线计算、虚实互动等手段，使人工系统成为可实验的“社会实验室”，以计算实验的方式为实际系统运行的可能情况提供“借鉴”、“预估”和“引导”^[36]。

平行系统采用复杂系统研究的“多重世界”观点，即对复杂系统进行建模时，不再以逼近某一实际的复杂系统的程度作为唯一标准，而认为模型是一种“现实”，是实际复杂系统的一种可能的替代形式和另一种可能的实现方式。实际复杂系统只是可能出现的现实中的一种，其行为与模型的行为“不同”，但却“等价”。简单来讲，人工系统是对实际系统的软件化定义，不仅是对实际系统的数字化“仿真”，也是为实际系统运行提供可替代版本（或其他可能的情形），从而实现对实际系统在线、动态、主动的控制与管理，为实际复杂系统管理运作提供高效、可靠、适用的科学决策和指导。

1.2 计算实验方法

1.2.1 计算实验的起源

系统科学是对现实世界中各类系统进行有效描述与求解的科学，通过适当的技术求解各类输入变化与系统输出的关系。但是，许多真实系统不能实验或者难以实验，不可以人为地改变系统输入或系统运行环境，原因可归结为如下几个方面：①经济方面，由于复杂系统的规模和成本因素，实验代价太大，以致经济上因无法承受而不可能进行实验；②法律方面，许多复杂系统涉及国家防卫、军事战备、社会安全等问题，受立法保护，以致无法对研究的系统进行实验，也无法重建这些系统；③道德方面，许多复杂系统，特别是复杂社会系统，往往有大量人员的参与，对这些系统进行实验，有可能冲击人的正常生活，甚至危害人的生命和财产，以致在道德上无法接受这类实验。

因此，如何对复杂系统进行实验研究，成为推进这一领域研究工作的关键问题之一。把复杂科学的系统理论与计算机模拟技术结合起来，将研究对象看作一个有机系统去试图把握其进化演化的一般规律，并在把握一般规律的基础上进行有效的引导和调控，就成为一种可能的思路 and 手段。传统的计算模拟认为实际系统是唯一存在的，并将仿真结果与实际系统是否一致作为检验实验结果的唯一标准；而复杂系统模拟则是希望把计算机作为“人工实验室”，来“培育”实际系统中可能出现的宏观现象，探究背后的规律，这为分析复杂系统行为和评估各种干预效果提供了一种可行方式。基于情景的策略评估如图 1-4 所示。

计算实验技术的发展可追溯至 20 世纪 60 年代，大致可以归为以下三条线索。

1) 基于微分方程对参照系统进行建模，包括状态自动机与系统动力学仿真，旨在通过传统的数学方法来描述确定性或随机性的系统。系统动态（system dynamics, SD）模型以离散时间差分方程的形式表示参照系统的状态和动态。因此，这种模型非常适合根据状态网络和变化率描述复杂系统。SD 模型可能是完全确定性的或部分随机的。

2) 基于随机过程与概率分布对系统的状态变迁进行建模，包括排队论、马尔可夫模型（Markov model）与离散事件仿真，旨在描述参照系统的流输入与流处理。例如，顾客在队列内的等候时间、车辆抵达目的地需要花费的时间等。

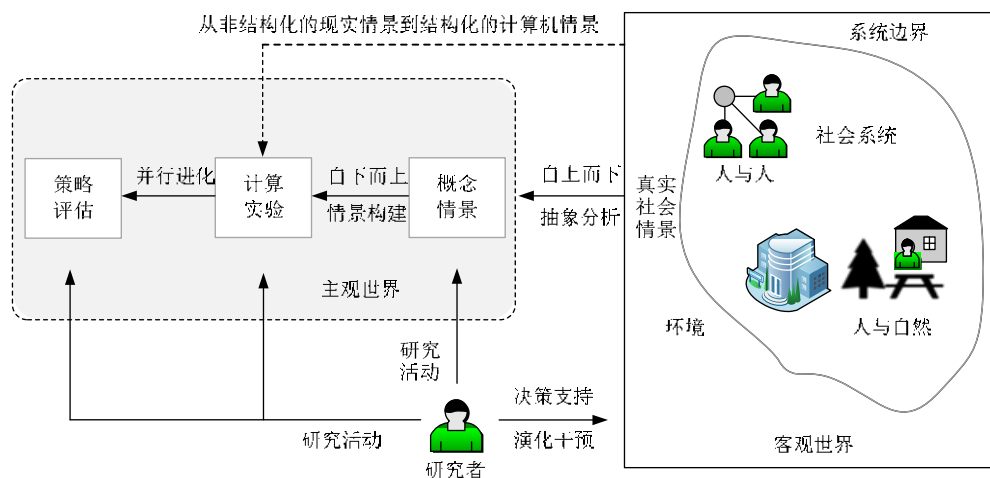


图 1-4 基于情景的策略评估

3) 基于自主个体间的相互作用进行建模，着眼于参照系统中实体的提取策略，包括元胞自动机模型、糖域模型与社会生态模型等，更加关心微观个体的简单行为在宏观层次上的涌现属性，旨在为复杂系统的模拟提供通用型的框架。元胞自动机模型首先强调实体元胞与拓扑结构，然后着眼于属性/变量。基于 Agent 的模型也与此有点类似。

在 20 世纪 90 年代，圣塔菲学派首先开始采用仿真实理论解决社会系统模拟的问题，开创了计算社会学的先河。仿真不是把计算机简单地作为一种仿真工具，而是把它作为“生长培育”实际系统替代版本的“人工社会实验室”，并进行各类有关系统行为和决策分析的“实验”。从此以后，社会领域不只是人文学者的园地，也是计算机专家大有作为的领域。计算实验所带来的最重要也最有影响力的一个方面，是实现真正的跨学科协同合作。从方法和路径上看，它提供了连接计算机领域、应用工程、人文学科、社会科学和经济研究的可能性。计算机研究者从自身的学科立场出发，可以把施展研究力量的范围扩展到社会领域，并能集体地从“问题导向”出发，与其他领域的专家合作研究许多有趣而新颖的课题，处理非常有特点的社会大数据。随着计算技术和分析方法的进一步发展，计算仿真迈上了一个更高台阶——计算实验。

2004 年 4 月王飞跃在文章《计算实验方法与复杂系统行为分析和决策评估》^[2]中，系统化地提出计算实验方法的基本思想、概念和方法，强调人工系统与实际系统之间的循环反馈关系。同年 5 月发表文章《关于复杂系统研究的计算理论与方法》^[37]，在讨论了对复杂系统及研究方法的一些认识基础上，提出了利用平行系统的概念和方法研究复杂系统的设想。其中，所设计的计算实验过程模型如图 1-5 所示，其中 $\{x_1, x_2, \dots, x_m\}$ 是人工系统的输入， $\{y_1, y_2, \dots, y_n\}$ 是人工系统的输出， $\{u_1, u_2, \dots, u_p\}$ 是可控因素或策略， $\{v_1, v_2, \dots, v_q\}$ 是不可控因素或事件。计算实验的目的可包括以下内容：①通过计算实验确定最能影响系统输出的因素集；②通过计算实验确定最有效的可控因素 u_i ，使输出结果集更小，且更接近理想的水平；③通过计算实验确定最有影响的可控因素 u_i 集，使不可控因素或事件 v_i 对系统的影响作用最小。

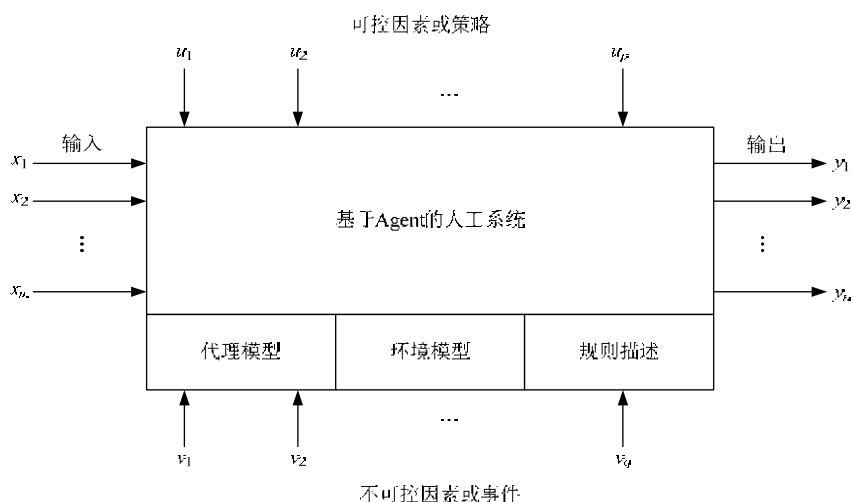


图 1-5 计算实验的过程模型

2009年2月，哈佛大学、麻省理工学院等世界一流大学的15位顶尖学者在 *Science* 杂志上联名发表了论文《计算社会科学时代的到来》^[38]。2012年，14位著名欧美学者又联合发布了《计算社会科学宣言》，计算社会科学（computational social science）的声音就盖过了其他名称^[39]。计算社会科学为观察和分析人类社会的复杂行为模式提供了重要的科研方法，具有两大特点^[38]：①强调综合与合成在行为产生中的作用，通过利用人工组件构造复杂行为模式，而不是将自然系统形式拆成部件来理解系统行为，因此采用的是整体而不是还原的方式；②通过把人工创造的系统置于实际、仿真或混合环境下，产生复杂的互动方式和相关行为，利用涌现方法进行观测总结，了解、分析和理解复杂系统的行为及其各种影响因素。

目前，计算实验方法已经用于一些风险较大、成本较高或者现实中无法进行直接实验的系统研究，包括交通系统^[40]、战争模拟系统^[41,42]、社会经济系统^[43]、生态环境系统^[44,45]、生理/病理系统^[46,47]、政治生态系统^[48]等。一旦建立了复杂系统的计算实验室，就可以在这个实验室中进行复杂系统的各种研究工作，甚至包括开放复杂巨系统^[29]。通过计算实验，可以分析复杂系统的行为，对复杂系统的干预策略进行量化评估，为实现有效干预提供新的工具和手段。

1.2.2 计算实验的应用

随着相关研究的不断发展，计算实验方法在多个领域显示出巨大的应用前景，特别是复杂系统管理与控制相关的诸多重要领域，如社会安全、电子商务、金融市场等。计算实验的核心研究领域如下：①社会变化和政府失灵的出现；②适应、学习、扩散以及散布；③联盟、销售协议及收购和技术评估；④风险评估及政策评估；⑤信息和网络安全。以下是计算实验与不同领域结合而形成的新兴学科。

1. 计算流行病学

计算流行病学是一门涉及多学科领域，利用计算机理解和控制疾病在群体中的空间扩散规律的科学。计算模型有助于理解流行病的空间与时间的动态关系，包括描述型模型和

生成型模型。描述型模型的功能是对大型数据库的相关性进行静态估计，生成型模型的功能则是计算疾病在大量群体中如何通过个体与个体之间的互动进行扩散。“疾病”一词可以指一种实际的传染病，也可以指一个更普遍的反应扩散过程。群体类型取决于疾病性质，包括人、动物、植物和计算机等。与此类似，必须通过模型予以体现的各种互动也取决于疾病和群体，包括气雾传播疾病的身体接近、性传播疾病的性接触，以及蚊虫传播疾病的昆虫取食方式。

2. 计算经济学

计算经济学是以计算机为工具而研究人和社会经济行为的社会科学，是经济学的一个分支。计算经济学是一个介于资讯科学、经济学与管理学之间的学科，主要采用计算经济模型求经济问题的解析解与统计解。这里的计算建模主要包括代理人模型、一般均衡模型、总体模型、理性预期模型、计算计量与统计模型、计算金融模型、网络市场的设计演算工具，特别是为计算经济学设计的规划工具，以及计算经济学的教学工具等。代理人计算经济学是其中一个比较经典的研究方向，专门研究将整体经济过程视为代理人间互动的动态系统，因此它是复杂适应系统的经济适应方式。在这里，“代理人”被视为根据规则互动的演算个体，而不是真的人群。代理人包括社会个体、生物个体和实质个体。

3. 计算社会学

社会科学也是一门比较复杂的学科。一方面，很多至关重要的社会过程都是复杂的，无法简单地拆分为单独的子过程，经济、人口统计、文化、空间等因素将这些子过程整合在一起；另一方面，某种可控的社会实验较困难，难以测试个人行为的多样化微观世界如何产生总体的社会宏观规律。计算社会学的新领域可被定义为关于多阶社会空间的跨学科调查，从个人到最大群体，均通过计算媒介实现。计算社会学立足于社会的信息处理范例，这意味着在理解社会系统和进程如何运转时，信息扮演着至关重要的作用，尤其当解释和理解社会复杂度时，计算模型发挥着基础性作用，包括贸易、移民、群体形成、战争、与环境交互作用、文化传播，以及人口动态等社会现象。

4. 计算政治学

政治理论、比较政治、公共政策和国际政治等，一直是社会科学领域的热点研究方向。研究政治科学的方法通常包括正式的理论构建、叙事分析、统计分析和个案分析等。随着社交媒体、电子商务、移动终端等新平台的兴起，依靠传统的统计定量分析方法去挖掘这些数据背后的政治问题已经非常乏力，有必要结合计算机技术进行多学科、多角度的研究与分析。马萨诸塞大学阿姆赫斯特分校政治系给出了计算政治学的定义：不但包含了对网络、传感器、通信、电子媒体或电子数据库等计算机生成数据的分析，而且使用计算形式及语言描述和分析政治现象，研究方法包括社会网络分析、文本分析、基于主题的建模、动态关系或集群模型、数据挖掘等。以下是计算政治学研究的几个经典问题^[49]。

(1) 政治倾向性及政治观点识别

对于某类政治事件，不同个人、不同群体、不同阶层都有着不同的观点及看法。如何从微博、博客等社交媒体的数据中，判断个人的政治倾向，进而分析不同群体、不同阶层

的政治主张，如选民支持率与地域及阶层的关系、个人政治信息的刻画等，成为国内外学者的主要研究内容。

（2）冲突观点检测

医疗、教育、住房等这些关乎民生的话题，在社交媒体中经常会受到关注。对于某项政策的颁布或修改，不同的个人或者团体都有着各自的观点与看法，不同的观点在表达的同时，会有不同的响应。有些观点切中要害，与部分人看法吻合，会被大量地转发和支持；某些观点违背了部分人的意愿，也会被大量地转发与质疑。对于某类政治事件，往往会有多方观点共存，这些观点有的相互补充，有的却大相径庭。针对这类问题的研究，包括冲突观点的识别、检测与分析，也成为计算政治学主要研究的内容。

（3）选举预测

淘宝可以推荐人们需要的商品，微博可以了解人们的兴趣、爱好，腾讯 QQ 可以猜测出人们认识的人是谁，并准确地将其分类。试想，如果人们拥有大量的电子病历档案，是否可以提前预测疾病，以及推荐合适的治疗措施？所以说大数据的核心就是预测。通过以往的数据来预测即将发生的事情，这是多么美好的一件事情。在计算政治学的研究中，选举预测也毫无例外地成为其研究的核心。

5. 计算金融学

1987年，SFI研究小组的各学科专家建立了基于 Agent 的股票市场仿真模型。该模型是基于 Agent 的金融市场建模研究过程中的突破性成果，开基于 Agent 的计算金融学（Agent-based computational finance, ACF）之先河，成为基于 Agent 金融市场建模研究的真正开端。

金融市场的复杂性和金融市场固有的投机性决定了金融市场经常处于非均衡的波动状态，金融监管政策的实施效果也因此面临着巨大的不确定性。金融市场是一个复杂适应系统，它由大量并行的、相互作用的适应主体通过经济行为联系在一起，每个主体只能获取整个系统的局部信息，但主体具有“适应性”，能够通过适应环境而不断地学习和改变自己的投资策略。因而，主体与主体之间以及个体与外部环境之间的交互作用使得整个金融市场形成了一个动态的演化系统。计算金融学以复杂适应系统理论为立论基础与研究方法，以金融市场的建模为立题背景，将金融市场复杂适应系统看作是由系统中 Agent 机制不断地与环境中的资源机制及其他 Agent 机制相互作用、相互聚集、层层构建而涌现出来的动态机制。

计算金融学通过计算和样本途径，利用大量具有交互影响的 Agent 构造与模拟复杂的金融系统，这些 Agent 不仅具有判断、更新和学习等功能，而且被赋予多样化的“个性”，从而使整个金融仿真系统具有自适应（adapting）和演化（evolving）的特征。这种分析工具比传统金融经济学的方法有更独到之处，它强调计算性、实验性、行为性、演化性与自下而上的结构，适用于对金融复杂系统分析的需要。更为重要的是，这一方法提供了进行经济实验的工具与平台，通过建立贴近现实情况的仿真市场模型，不仅可以考察不同投资者的市场行为对市场波动的影响，以及它们的预期收益和风险状况，而且可以模拟市场参与者和监管部门的行为，考察各种监管措施和政策可能产生的市场反应，由此评估政策效果和制度建设的可行性。（注：在真实市场中进行经济实验代价巨大，几乎是不可能的。）

计算经济学与政策模拟研究旨在通过经济过程的模拟计算和经济政策的模拟分析，为经济政策的制定提供科学基础。

1.3 本书的组织结构

计算实验方法作为复杂性科学研究的一种主要手段，集成了计算机科学、社会科学、系统科学、计算机模拟技术、多 Agent 系统技术、人工智能技术等诸多学科。尽管科研人员付出了很多努力，计算实验方法目前还不能算是一套成熟的体系，也没有一套完整而成形的理论，其理论发展和实际应用之间还存在着巨大的差距。对于初学者，很容易湮没在浩如烟海、纷繁芜杂的文献资料中，难以剥茧抽丝，理出自己的思路。而且，人工智能技术等新技术的蓬勃兴起，应用领域的千差万别，更加剧了这种困境。

为了能够对解决该问题有所帮助，本书梳理了计算实验方法的来龙去脉，希望帮助读者构建起计算实验方法的完整知识体系，从而为后续的方法应用奠定坚实的基础。本书的内容包括原理、模型和案例三个方面。

1) 计算实验方法的系统框架是什么？计算实验方法作为一个多学科的交叉领域，涉及的知识非常多。初学者刚开始接触计算实验时，很容易湮没在浩如烟海的文献中，难以厘清计算实验方法与各类技术之间的关系，而各类层出不穷的新技术更加剧了这种困境。基于此，本书首先就是对纷繁芜杂的文献资料剥茧抽丝，对计算实验方法本身进行完整的系统化阐述，形成通用的方法框架，包括方法的适用领域、执行步骤、支撑平台，以及关键支撑技术。鉴于多 Agent 系统是大多数计算实验的实现平台，所以本书详细讲述了多 Agent 系统设计与开发所涉及的相关技术。

2) 计算实验方法的核心技术是什么？计算模型是计算实验方法的核心，也是不同领域知识发挥作用的容器。计算模型可以根据应用问题进行定制化，其中 Agent 智能程度的高低、自治单元间的交互关系、实验系统的抽象程度，完全取决于所要解决的问题，并不需要过度追求智能化。本书首先对常用的传统仿真模型进行了梳理，例如状态自动机、马尔可夫模型、排队模型等；然后重点介绍了几类人工社会的构建模型，包括元胞自动机、糖域与社会学习模型。现在关注的重点，就是如何利用各类新涌现出来的人工智能技术来提升计算模型设计水平，包括博弈学习、增强学习、群智学习等。

3) 计算实验方法如何实现领域应用？由于影响因素的多样性和不确定性，加上经济、法律、伦理等方面的限制，复杂系统的分析和评估对于传统分析方法是一个严峻挑战，有时甚至是不可能完成的任务。相比于传统分析方法，计算实验方法通过改变内外部因素的组合方式，可以建立各种各样的实验环境，对不同因素在演化过程中的影响进行全面、准确、及时、量化地分析评估，甚至可以对系统演化进行各种压力实验以及极限实验。因此，计算实验是非常适合复杂系统研究的手段与工具。本书从社会生态系统的角度出发，主要以电子商务领域和智慧矿山为例，具体给出了计算实验方法的应用案例。

本书内容的组织结构如图 1-6 所示，主要分为四部分：方法理论篇、计算模型篇、应用案例篇和总结与展望篇。全书采用“理论+验证”的研究思路，从计算实验的基础知识出发，使读者在理解计算模型的基础上对实验案例进行进一步剖析，从而能够对计算实验方法体系有一个完整的把握。每个章节的具体内容如下。

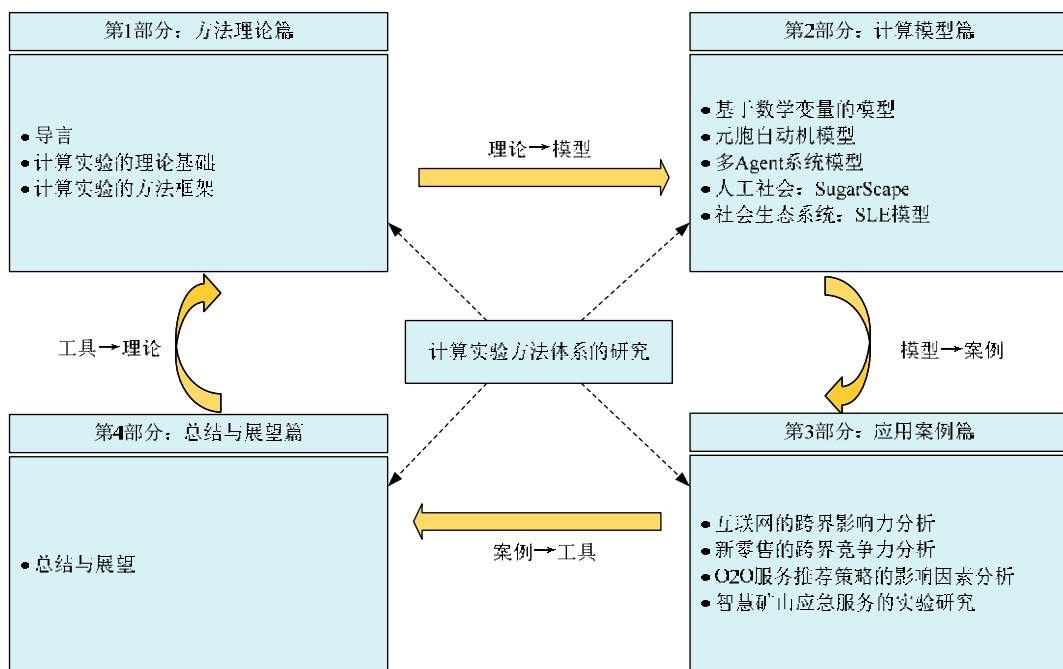


图 1-6 本书内容的组织结构

1. 方法理论篇

第 1 章主要介绍计算实验方法的起源与思想精髓，并介绍了本书的组织结构与内容。本章明确指出计算实验是计算仿真随着计算技术和分析方法的进一步发展而迈上的一个更高台阶，是分析复杂系统行为和评估各种决策效果的一种可行方式，也是弥补传统方法很难甚至无法对复杂系统进行实验的一种有效手段。

第 2 章主要介绍计算实验方法的复杂性理论基础，能够为系统复杂性的理解提供一些独到的见解，进而支持跨文化和跨学科领域的研究。本章首先从计算实验的目的出发，让读者理解计算实验的特有属性以及相应的运作机制；然后介绍了系统复杂性的规律描述与理论解释原则，为计算实验的应用提供方法指导。

第 3 章详细阐述了计算实验方法包含的六个步骤：①方法的适用性判定；②计算实验系统的开发；③计算实验的执行；④计算实验结果的分析；⑤计算模型的评估；⑥实验结果的应用。本章帮助读者对整个计算实验方法的应用过程有完整的了解，为后续章节的学习奠定基础。

2. 计算模型篇

第 4 章首先讲述了系统研究方法和模型分类，然后详细介绍了计算实验方法中一些常用的基础模型，包括状态自动机、马尔可夫模型与隐马尔可夫模型、离散事件仿真、系统动力学模型、社会网络模型及排队论模型等。每种模型都给出了其基本概念、建模过程，以及相关的案例分析。

第 5 章详细介绍了非常经典的元胞自动机模型，包括模型的起源、构成、应用与扩展。

这是一种时间、空间、状态都离散的、能够模拟复杂系统时空演化过程的动力系统，能够以几条简单的演化规则涌现出无法预测的复杂性，这正是复杂性科学的研究焦点。

第 6 章以多 Agent 系统作为构建计算实验系统的基础，介绍了 Agent 的概念模型以及所涉及的各种主流关键技术，包括如何感知、如何决策、如何行动、如何交互、如何协作等，并给出了多 Agent 系统的分析技术。

第 7 章详细介绍了另一类经典的计算模型：人工社会模型，包括人工生命模型、捕食者-猎物生态圈模型以及糖域模型。这类模型给出了人工社会的概念及基本思路，通过观察 Agent 整体作用的涌现属性找到人工社会的规律，并用这些规律解释和理解现实人类社会中的宏观现象。

第 8 章详细介绍了一种社会生态模型框架 SLE，通过描述个体、组织与社会之间的循环反馈交互作用，为社会经济技术生态系统的模拟提供了一条通用的技术路径。该模型框架具有定制化特点，能够利用各种流行的智能技术为构建计算实验系统提供支撑。

3. 应用案例篇

第 9 章以评估互联网对传统行业的跨界影响力为例，从供需匹配的角度出发，提出了“服务桥”方法，并采用计算实验方法对 O2O 在美容行业 and 外卖行业的应用进行了实验分析，并进一步对互联网能否颠覆几个传统行业进行了预测，如零售业、金融业、制造业和医疗行业。

第 10 章以评估新零售的跨界模式是否可行为例，提出了基于价值链的跨界服务分析框架，包括需求识别、数据融合、能力集成和实施策略四个模块。在此基础上，采用计算实验方法对阿里和京东的新零售策略进行了分析，并分别给出了平台模式与自营模式最合适的跨界策略。

第 11 章以 O2O 服务推荐策略的评估优化为例，提出了一个基于“Cyber-Social”循环反馈机制的 O2O 服务推荐策略设计框架，并据此设计了三种推荐策略：基于节点信息反馈的推荐策略（service-CF）、基于路径信息反馈的推荐策略（path-CF）和基于持续反馈的推荐策略（O2O-CF）。在此基础上，采用计算实验方法对三种策略在不同场景下的表现进行了对比分析，为策略选择提供决策支持。

第 12 章针对矿山特殊环境下难以评估应急策略效果的难题，提出了一种基于计算实验的情境感知服务评估方法，主要包括三部分：服务实现策略的定制、计算实验系统的构建和服务策略效果的评估。在此基础上，采用计算实验方法对煤矿环境中的报警服务和疏散服务进行了对比分析，为策略选择提供决策支持。

4. 总结与展望篇

第 13 章针对全书的内容进行总结，并给出下一步的研究重点。

1.4 重要的经典文献

本书凝聚了计算实验领域很多研究人员的努力成果，在此感谢他们对计算实验方法的发展所作出的显著贡献。表 1-1 给出了各章相关的重要经典文献，感兴趣的读者可以很方便地查阅原始文献，深入了解相关的技术细节。

表 1-1 各章相关的重要经典文献

章序	相关的重要经典文献	备注
第 1 章	王飞跃, 2004. 计算实验方法与复杂系统行为分析和决策评估[J]. 系统仿真学报(5): 893-897. 王飞跃, 2004. 关于复杂系统研究的计算理论与方法[J]. 中国基础科学(5): 5-12.	王飞跃关于 ACP 方法的开山之作
	WANG F Y, 2007. Toward a paradigm shift in social computing: The ACP approach[J]. IEEE intelligent systems, 22(5): 65-67.	针对计算实验方法以及相关应用进行系统化论述的几篇代表性论文
	WEN D, YUAN Y, LI X R, 2012. Artificial societies, computational experiments, and parallel systems: an investigation on a computational theory for complex socioeconomic systems[J]. IEEE transactions on services computing, 6(2): 177-185.	
	WANG F Y, 2010. Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications[J]. IEEE transactions on intelligent transportation systems, 11(3): 630-638.	
第 2 章	CIOFFI-REVILLA C, 2014. Introduction to computational social science-principles and applications[M]. London: Springer-Verlag.	书中对于社会领域中复杂性产生的原因和规律进行了系统化的阐述
第 3 章	XUE X, WANG S F, GUI B, et al., 2016. A computational experiment-based evaluation method for context-aware services in complicated environment [J]. Information sciences, 373: 269-286.	针对计算实验方法的具体应用问题给出了具体的路线图, 对于方法初学者很有价值
	李大宇, 米加宁, 徐磊, 2011. 公共政策仿真方法: 原理, 应用与前景[J]. 公共管理学报 8(4): 8-20.	针对计算实验在公共政策推演领域的应用进行了总结, 其中有关实验方法步骤的某些观点非常值得借鉴
	盛昭瀚, 张军, 杜建国, 等, 2009. 社会科学计算实验理论与应用[M]. 上海: 上海三联书店. 盛昭瀚, 李静, 陈国华, 等, 2010. 社会科学计算实验基本教程[M]. 上海: 上海三联书店. 盛昭瀚, 张军, 刘慧敏, 等, 2011. 社会科学计算实验案例分析[M]. 上海: 上海三联书店.	书中对于计算实验方法在社会领域的应用进行了系统化的阐述, 对于社会学及相关领域的科研人员非常有参考价值
第 4 章	GILBERT N, TROITZSCH K G, 2005. Simulation for the Social Scientist [M]. 2nd ed. Berkshire: Open University Press.	书中解释了为什么会在社会科学状况中用到模拟, 同时非常详细地列出了一系列社会模拟方法
第 5 章	梅拉妮·米歇尔, 2011. 复杂[M]. 唐璐, 译. 长沙: 湖南科学技术出版社.	书中第二和第三部分对元胞自动机和人工生命做了非常精彩的论述, 尤其是有关遗传算法的例子, 令人拍案叫绝
第 6 章	BELLIFEMINE F, CAIRE G, GREENWOOD D, 2007. Developing multiagent systems with JADE[M]. New York: John Wiley & Sons Ltd.	有关多 Agent 系统开发的著作, 其中 Michael Wooldridge 的《多 Agent 系统引论》是学习 Agent 技术一本很好的入门书籍
	WOOLDRIDGE M, 2009. An Introduction to multiagent systems[M]. New York: John Wiley & Sons Ltd.	
	毛新军, 2005. 面向主体的软件开发[M]. 北京: 清华大学出版社.	
	薛霄, 2009. 面向 Agent 的软件设计开发方法[M]. 北京: 电子工业出版社.	
	宣慧玉, 张发, 2008. 复杂系统仿真及应用[M]. 北京: 清华大学出版社. 方美琪, 张树人, 2011. 复杂系统建模与仿真[M], 2 版. 北京: 中国人民大学出版社.	有关复杂系统建模仿真的书籍, 都包含关于多 Agent 系统的相关章节

续表

章序	相关的重要经典文献	备注
第 7 章	EPSTEIN J M, AXTELL R, 1996. Growing artificial societies: social science from the bottom up[M]. Washington: Brookings Institution Press.	糖域模型的发明者, 详细介绍了该模型的来龙去脉。其中的细节对于如何将现实世界映射为计算实验, 非常有启发意义
第 8 章	XUE X, WANG S F, ZHANG L J, et al., 2019. Social learning evolution (SLE): Computational experiment-based modeling framework of social manufacturing[J]. IEEE transactions on industrial informatics, 15(6): 3343-3355.	SLE 模型的价值在于: 一方面, 给出了一个社会生态系统建模的层次化模板; 另一方面, 定制化的特性将计算实验与各种智能算法联系起来
第 9 章	XUE X, GAO J J, WANG S F, et al., 2018. Service bridge: trans-boundary impact evaluation method of Internet[J]. IEEE transactions on computational social systems, 5(3):758-772	针对网上各种 O2O 颠覆传统的言论提出了服务桥模型, 可以分析互联网是否能够改变某个行业, 以及多大程度上能够改变这个行业
第 10 章	XUE X, GAO J J, WU S, et al., 2019. Value-based analysis framework of crossover service: A case study of new retail in China[J]. IEEE transactions on services computing, (99):1-14.	分析的目的是为了预测, 预测的目的是为了控制。这个章节解读了针对不同的电商企业公司, 什么样的新零售实施策略最合适
第 11 章	XUE X, HUANG F S, WANG S F, et al., 2019. Research on escaping the big-data traps in O2O service recommendation strategy[J]. IEEE transactions on big data: 1-14. DOI: 10.1109/TBDATA.2019.2915798.	大数据分析给人们的生活带来了便利, 但是也会造成信息茧房。如何规避所谓的大数据陷阱, 本文给出了一些思路
第 12 章	XUE X, WANG S F, et al., 2016. A computational experiment-based evaluation method for context-aware services in complicated environment[J]. Information sciences, 373:269-286.	危险环境是计算实验另一个非常适用的领域。文中以煤矿中的报警和疏散为例, 给出了计算实验方法的使用流程

1.5 小 结

本章首先给出了复杂系统的概念, 即计算实验方法产生的原因。然后, 介绍了计算实验方法的思想和应用领域, 明确指出计算实验是计算仿真随着计算技术和分析方法的进一步发展而迈上的一个更高台阶, 是分析复杂系统行为和评估各种决策效果的一种可行方式, 也是弥补传统方法很难甚至无法对复杂系统进行实验的一种有效手段。另外, 本章交代了本书的组织结构, 为本书的后续学习提供了向导。

参 考 文 献

- [1] DERNER D. The logic of failure[J]. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 1990, 327(1241): 463-473.
- [2] 王飞跃. 计算实验方法与复杂系统行为分析和决策评估[J]. 系统仿真学报, 2004(5): 893-897.
- [3] DARWIN C. The origin of species[M]. New York: PF Collier & son, 1909.
- [4] 汪志诚. 热力学·统计物理[M]. 北京: 高等教育出版社, 1993.
- [5] BERTALANFFY L V. General System Theory: Foundations, Development, Applications[M]. New York: George Braziller, 1968.
- [6] BERTALANFFY L V. Zu einer allgemeinen systemlehre[J]. Biologia Generalis, 1949(19), 139-164.
- [7] BERTALANFFY L V. Das biologische Weltbild, Bern: Europäische Rundschau[J]. Problems of Life: An Evaluation of Modern Biological and Scientific Thought, New York: Harper, 1952.
- [8] GOODE H H. MACHOL R E. System Engineering: An Introduction to the Design of Large-Scale Systems[M]. New York:

- McGraw-Hill, New York, 1957.
- [9] WIENER N. *Cybernetics: Or Control and Communication in the Animal and the Machine*[M]. Paris: Hermann & Cie & Camb. Mass, 1961.
- [10] SHANNON C E. A mathematical theory of communication[J]. *Bell System Technical Journal*, 1948, 27(3): 379-423.
- [11] SHANNON C E. Communication theory of secrecy systems[J]. *Bell System Technical Journal*, 1949, 28(4): 656-715.
- [12] NUNEZ J A, CINCOTTA P M, WACHLIN F C. Information entropy[J]. *Celestial Mechanics & Dynamical Astronomy*, 1996, 64(1-2): 43-53.
- [13] PRIGOGINE I STENGERS I. *Order out of Chaos: Man's new dialogue with nature*[M]. New York: Bantam Books, 1984.
- [14] PRIGOGINE J. Structure, dissipation and life[J]. *Theoretical physics and biology*, 1969: 23-52.
- [15] LIU Z. Dissipative structure theory, synergetics, and their implications for the management of information systems[J]. *Journal of the American Society for Information Science*, 2010, 47(2): 129-135.
- [16] THOM R. *Structural stability and morphogenesis*[M]. Boca Raton: CRC Press, 2018.
- [17] HAKEN H. *Synergetics-cooperative phenomena in multi-component systems*[C]// Vieweg+ Teubner Verlag, Wiesbaden, 1973.
- [18] HAKEN H P J. *Synergetics*[J]. *IEEE Circuits & Devices Magazine*, 1977, 28(9): 412-414.
- [19] LI T, YORKE J A. Period three implies chaos[J]. *The American Mathematical Monthly*, 1975, 82(10): 985-992.
- [20] RAY T S. An evolutionary approach to synthetic biology: zen and the art of creating life[J]. *Artificial Life*, 2014, 1(1-2): 179-209.
- [21] RAY T S. *Evolution, ecology and optimization of digital organisms*[R]. Technical Report 92-08-042, Santa Fe Institute, Santa Fe, NM, 1992.
- [22] MANDELBROT B B. *Fractals, form, chance and dimension*[J]. *The Mathematical Gazette*, 1978, 62(420): 130-132.
- [23] MANDCLBROT B B. *The Fractal Geometry of Nature*[M]. 2nd ed. San Francisco: Chance and Dimension, Freeman and Co., 1977.
- [24] BAK P, TANG C, WIESENFELD K. Self-organized criticality: an explanation of $1/f$ noise[J]. *Physical Review Letters*, 1987, 59(4): 381-384.
- [25] 张发, 宣慧玉, 赵巧霞. 复杂系统多主体仿真方法论[J]. *系统仿真学报*, 2009, 21(8): 2386-2390.
- [26] WOLFRAM S. Random sequence generation by cellular automata[J]. *Advances in Applied Mathematics*, 1986, 7(2): 123-169.
- [27] TERUYA M, NAKAO Z, CHEN Y, et al. A boid-like example of adaptive complex systems[C]// *IEEE International Conference on Systems*. Hawai'i: IEEE, 1999: 1-5.
- [28] STUTZL E, MACRO D, THOMA S. *Ant Colony Optimization*[M]// *Ant Colony Optimization*. Holland: Bradford Company, 2004.
- [29] LAW A M, KELTON W D. *Simulation Modeling and Analysis*[M]. 3rd ed. New York: The McGraw-Hill Companies, 2000.
- [30] FISHWICK P. *Simulation Model Design and Execution* [M]. Kent: Prentice Hall, 1995.
- [31] GILBERT N, TROITZSCH K G. *Simulation for the Social Scientist*[M]. Berkshire: Open University Press, 2005.
- [32] WANG F. Linguistic dynamic systems for modeling and analysis of coupled natural and human systems[R]. Santa Fe, NM: NSF/SFI Workshop on the Robustness of Coupled Natural and Human Systems, 2003.
- [33] WATTS D J, STROGATZ S H. Collective dynamics of "small-world" networks[J]. *nature*, 1998, 393(6684): 440.
- [34] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509-512.
- [35] WANG F. The emergence of intelligent enterprises: from CPS to CPSS[J]. *IEEE Intelligent Systems*, 2010, 25(4): 85-88.
- [36] 王飞跃. 平行系统方法与复杂系统的管理和控制[J]. *控制与决策*, 2004, 19(5): 485-489.
- [37] 王飞跃. 关于复杂系统研究的计算理论与方法[J]. *中国基础科学*, 2004(5): 5-12.
- [38] LAZER D, PENTLAND A, ADAMIC L, et al. Computational social science[J]. *Science*, 2009, 323(1): 721-723.
- [39] CONTE R, GILBERT N, BONELLI G, et al. Manifesto of computational social science[J]. *The European Physical Journal Special Topics*, 2012, 214(1): 325-346.
- [40] 黄海军. 城市交通网络动态建模与交通行为研究[J]. *管理学报*, 2005(1): 18-22.
- [41] 胡晓峰, 司光亚, 罗批, 等. 战争复杂系统与战争模拟研究[J]. *系统仿真学报*, 2005(11): 200-205.
- [42] CARLEY K M, FRIDSMA D B, CASMAN E, et al. BioWar: scalable agent-based model of bioattacks[J]. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 2006, 36(2): 252-265.
- [43] LEIGH T. Agent-based computational economics[J]. *ISU Economics Working Paper*, 2003, 249(1): 2002-2016.

- [44] ACEVEDO M F, CALLICOTT J B, MONTICINO M, et al. Models of natural and human dynamics in forest landscapes: Giorgio-site and cross-cultural synthesis[J]. *Geoforum*, 2008, 39(2): 846-866.
- [45] 龚承柱, 李兰兰, 柯晓玲, 等. 基于 multi-agent 的煤矿水害演化模型[J]. *煤炭学报*, 2012, 37(6): 1005-1009.
- [46] HUANG C, SUN C, HSIEH J L, et al. Simulating SARS: small-world epidemiological modeling and public health policy assessments[J]. *Journal of Artificial Societies and Social Simulation*, 2004, 7(4): 2.
- [47] CIOFFI-REVILLA C, ROULEAU M. MASON RebeLand: an agent-based model of politics, environment, and insurgency[J]. *International Studies Review*, 2010, 12(1): 31-46.
- [48] 钱学森, 于景元, 戴汝为. 一个科学的新领域: 开放的复杂巨系统及其方法论[J]. *自然杂志*, 1990, 13(1): 3-10.
- [49] HEMASPAANDRA E, HEMASPAANDRA L A. Computational politics: electoral systems[J]. *Lecture Notes in Computer Science*, 2000, 1893: 64-83.

第2章 计算实验的理论基础

2.1 引言

科学研究的目的是主要有两个：规律描述与理论解释。规律用以理解系统复杂性“如何”产生；理论用以回答系统复杂性“为何”会发生。规律就像变量之间的映射；理论则是对所观察到的系统复杂性进行因果关系的说明。为什么将计算实验用作系统复杂性的科学研究方法呢？简言之，与其他研究方法（如统计学、数学或史学）相比，基于计算实验的系统复杂性理论有时更为可行，能够提供有关复杂性的一些独到见解，从而支持跨文化和跨学科领域的研究。

计算实验将复杂性理解为涌现现象^[1]，从规律描述和理论解释两个方面出发，描述信息-物理-社会相互耦合的复杂系统。通过精心设计的计算模型与模拟仿真环境，可提供用于系统复杂性推理^[2]、实验以及最终理解的强大工具。但是，计算实验同样面临着一些限制，其中比较重要的一点就是，计算实验的描述并不如其他形式和统计模型那么直接，有时可能对结果的可复制性产生重大影响。另外，计算实验的代码实现形式，可能随着时代和技术的发展，也需要不断地更新换代。

为了开发有效的计算模型，研究者有必要理解系统复杂性的基本研究模式。本章的重点是解释计算实验与复杂性的关联关系，帮助读者理解系统复杂性的成因（规律描述），并进而扩展到系统复杂性的解释理论（理论解释）。

2.2 复杂性的实验分析

2.2.1 计算实验的目的

复杂性理论^[3]关注大量个体之间的相互作用，个体可能是原子、鱼、人、组织或者国家。它们的互动模式可能包括集聚、冲突、交流、贸易、协作或是竞争。互动模式的研究要涉及大量个体，其数学求解往往非常困难。所以，复杂性理论的一个主要研究工具就是计算机模拟实验，其关键在于指定个体互动的方式，然后观察整个宏观层面上涌现出的特征。通过对模拟实验的非常规分析以了解复杂系统的特性，这就是计算实验模型的设计目的。采用计算实验来模拟系统复杂性有很多理由，包括但不限于以下方面。

1) 多功能性：有许多很复杂的社会系统和流程，统计或数学建模可能都无法奏效，但是却可以通过计算实验方法进行研究。也就是说，可以用计算实验对统计或数学模型进行替代，但并不是所有计算实验模型都可采用数学形式予以表示。

2) 高维度：系统复杂性的一个共同特征是必须分析大量的变量以及变量之间的交互作用，即所谓的高维度特性。例如，集体行动是涉及众多实体和变量的过程（包括环境参数、

目标、特征、资源等)。高维系统在社会复杂性领域中很常见。

3) 非线性: 实体间的动态交互作用通常为非线性的, 不受维度影响。人类感知、交互作用以及竞合模式, 都是非线性相互作用的示例。计算实验可处理复杂的非线性动态系统, 但会受制于相应的计算资源。

4) 耦合系统: 复杂性的另一个特征是信息世界、物理世界和人类社会之间的耦合。该耦合事实上暗示着高维度、非线性的交互作用。计算实验模型提供了表示 CPSS 的有效方法, 可用来表示社会机构、物理世界以及关键基础设施之间的耦合动力机制。

5) 随机性: 随机性不仅在社会系统和流程中无处不在, 并且是非常重要的。这里可以通过概率分布定义多种形式的随机性, 分析不同的随机性如何生成系统复杂性模式是使用计算实验的另一个主要原因。

6) 实验: 实验方法是所有科学研究的基础, 但在诸多原因(如经济的、伦理的、安全的等)的影响下针对复杂系统开展实验并不可行。因此, 计算实验被认为是可行的替代品, 且具有该类方法的所有典型特征, 包括实验组、对照组和许多不同的实验设计。例如, 计算实验可用来探索和测试各种假设之下的群体行为、群体动态以及管理假定。

7) 策略剖析: 由于系统的复杂性, 对于其他方法而言, 策略分析基本上属于不可能完成的任务, 而这恰恰是计算实验所擅长的领域。现实中一些难以评估测试的“邪恶问题”, 例如, 建模分析各类缓解通货膨胀的经济政策(如控制工资津贴或物价), 在这里都不再成为问题。

这里把计算实验^[4]方法与归纳、演绎这两种标准方法进行对比。归纳^[5]是在实证数据中发现模式, 在科学研究中, 归纳广泛地用于观点调查和宏观规律分析; 而演绎^[6]则是定义一套公理, 并证明由这些公理所能推导出的结论, 例如, 博弈均衡的结果就是使用理性选择公理得出的。计算实验则是进行科学研究的第三种方法, 与演绎一样, 它起始于定义一套清晰的假设; 但与演绎不同的是, 它并不证明定理。计算实验产生模拟数据, 这些数据可以用归纳的方法进行分析; 但是它又和典型的归纳法不同, 模拟数据来自严格设定的规则, 而不是对真实世界的直接测量。归纳旨在从数据中发现模式, 演绎希望发现假设的结果, 而计算实验旨在帮助和验证人们的直觉。

计算实验是进行思想试验的一种方法。虽然假设可能过于简单, 但是结果并非就是显而易见的, 局部互动的个体会产生全局范围的影响。不同于计算机仿真, 计算实验并不要求计算模型完全再现实际系统的行为, 如何证明计算模型与实际系统的有效性与等价性目前还没有取得共识。一种较为流行的解决方法是, 比较计算模型与实际系统的输出统计特征, 以此证明计算模型的等价性; 另一种解决方法则是在假设计算模型无法验证的前提下进行分析, 尝试挖掘计算模型所隐含的潜在结论。这种方法的应用较少, 但是其思路值得借鉴。

2.2.2 计算实验的模式

计算实验是一个迭代的过程, 始于现实世界中的参照系统, 然后通过抽象、形式化、编程和合适的数据构建可行的计算模型。计算实验的基本术语和通用过程如图 2-1 所示, 这个通用过程并不受计算模型的特定类型影响。接下来, 对过程中的一些必要项进行定义。

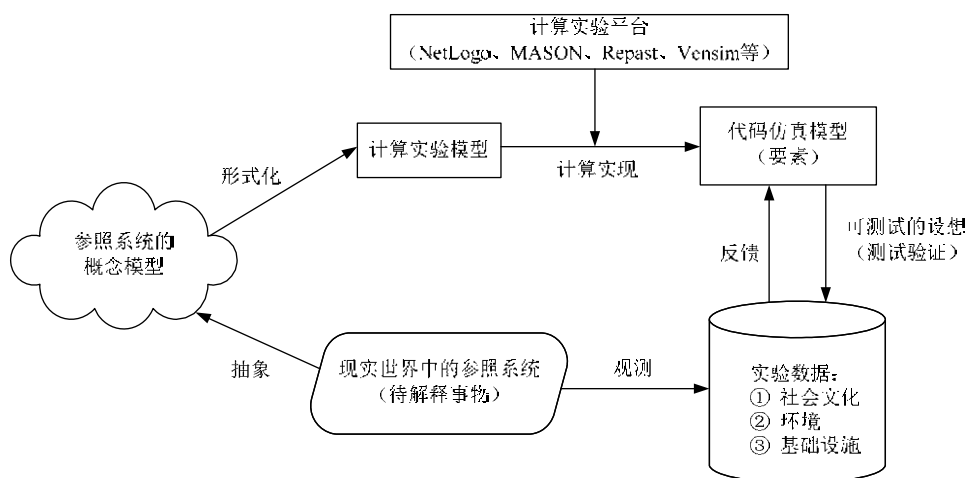


图 2-1 计算实验的基本术语和通用过程

定义 2.1 (参照系统)^[7] 作为研究对象的真实世界系统或过程。同义词：目标系统、重点系统、经验或历史世界。

复杂性研究的参照系统包含全部的实体、系统和流程，例如，人类思维、认知过程、个体决策、群体行为、社会、国际和全球网络。目前，复杂的参照系统大都为社会系统 (social system)、物理系统 (physical system)、信息系统 (cyber system)，以及它们的两两组合。最复杂的参照系统是 CPSS。客观地说，“现实”具有无限细节且规模巨大，因此，科学研究只能集中关注某些选定的现实子集，即根据研究问题来定义参照系统。

定义 2.2 (计算实验平台) 用于构建计算实验模型的计算工具组或代码库。

计算实验平台是一个高度复杂的计算平台，用于构建具体领域的计算实验模型，NetLogo、DYNAMO、Stella、Vensim、Swarm、MASON、Repast 及其前身都是计算实验平台的示例。计算实验平台与计算实验模型的关系，就如同汽车工厂与汽车的关系，可以使用前者制成后者。可以完全使用本地代码构建计算实验模型，这就类似于独立制造一辆汽车，而不是购买工厂制造的汽车，但是其性能和可靠性可能不如工厂制造的汽车。在科研过程中选择使用计算实验平台 (Vensim、MASON、Repast 等) 是因为仅依赖本地代码无法达到所需要的模型性能和可靠性水平。有时，对于一组给定的研究问题，使用计算实验平台是最好的解决方法。

计算实验平台的一些常见工具组包括：①构建模型的常用代码库或基本模块，如数学函数、分布、基本的 Agent 结构、场景、调度器、常见的数据字段以及构造函数方法等；②随机数程序，计算实验经常需要用随机数生成器表示各种随机形式的实体或过程（平均化、程序性、幂律等）；③可视化工具，绘制用于理解实验输出的直方图、时序图、网络图、地图和其他视觉手段。更专业的工具通常由模型开发人员自行开发，如自相关图、光谱图、差异地图、热图、动态网络、洛伦兹曲线图和各種非笛卡尔坐标系（如球形、圆柱形）等。

定义 2.3 (计算实验模型) 采用给定编程语言对参照系统进行形式化描述的模型。

在现实世界或正常情况下，人类寻求目标的行为并不完全依赖于理性选择，而往往是根据已知的有限理性做决定和行动。为了更好地理解计算实验模型与现实的差别，这里以完全理性模型为例，对比该模型的主要假设与实际人类的有限理性假设。理性选择模型无

论是单一假设还是作为一个集合的假设，都严格依赖目标、选择、结果、概率、预期效用和效用最大化 6 种假设。通过这个例子，也许并不难理解模型与现实之间的距离，以及如何根据需要修订模型，逼近现实。下面对两类模型的假设进行详细对比。

1) 目标：完全理性模型的目标清晰而精确；相比较而言，人类对于所追求目标的理解通常并不准确，尤其在压力之下做出的决策。

2) 选择：完全理性模型的可选择目标已知，而人类对于可选择的理解通常并不完整。很多状况会加剧该问题，包括压力、不完整信息和类似因素的存在。

3) 结果：对两种模型而言，各种选择均会引起一系列不同的演进路径与结果，因为涉及预测，所以这一点很难估计。人类的偏见会进一步加剧这种状况，如个人算计、集体思考及其他偏见。

4) 概率：在完全理性模型中，每种结果的出现概率已知，而人类在做选择时通常运用直觉，这对于估测结果的概率而言比较困难。

5) 预期效用：完全理性模型可以计算每种结果的预期效用，并可以在做选择的时候加以考虑。而人类推理无法进行预期效用计算，除非是最简单的状况或消耗大量的计算工作。

6) 效用最大化：完全理性模型通常选择预期效用最大的选项。而人类通常会决定做他们感觉必须要做的事情，但可能并不会使他们的利益最大化，或者他们通过一些其他原则做出选择，而这些做法可能并不会带来最大的预期收益。

2.2.3 计算实验的质量维度

计算实验的质量具有多重属性，而非单一维度。质量维度可用于评估开发计算实验模型时所需注意的属性清单。首先，从模式的角度看，给定模型的综合质量评估取决于多维度属性的累加或者乘积，而非取决于某个属性。其次，从建模的角度看，质量评估应包括整个建模过程，最终计算模型的质量取决于一系列前期设计阶段的产物。因此，通行的模型质量评估框架^[8]应该综合评估两个方面：模型本身及其开发过程。这一框架提供了可行的质量维度检查流程，主要包括以下 6 个要素。尽管这些基本维度之间并非独立，但是它们的整体表现对于综合质量评估很重要。

1) 规则。从设定计算实验应解决的研究问题开始进行质量评估，包括研究问题或问题类型是否明确？参照系统是否明确？研究问题是否具有原创性？所有的计算模型都是为了解决研究问题，那么清晰、原创和意义至关重要。问题设定时需考虑动机。通过对现有文献的完整调研，从而评估该模型是否首创，同一领域是否存在类似的统计或数学模型。

2) 实施。计算实验以代码形式呈现抽象模型，许多与质量相关的因素都涉及其中，包括代码是否将相关理论实例化？代码编写质量是否良好？模型是用本地代码还是工具组编写？针对具体的研究问题，随机数发生器（random number generator, RNG）的质量如何？考虑到不同语义，应使用哪种类型的数据结构？就算法效率而言，通过模型解决问题的实现难点是什么？在实现设计理念方面，代码的效率如何？从设计的角度看，适当的代码结构与研究问题相符吗？模型是否将参照系统提炼为设定的抽象水平？

3) 验证。执行主动和被动测试验证模型是否以预期的方式运行。验证实验包括但不限于以下内容：代码审查、调试、单元测试、剖析及其他软件开发时常用的技巧。质量评估应包括对于验证过程的调查，原因是最终结果的涉及范围很广，在不提供很多验证过程信

息的情况下，大多数计算实验看起来“结果不言自明”，但事实上并非如此。

4) 有效。对计算实验模型执行一系列的外部测试，包括评估拟合度、时间序列、空间分析、网络结构等。这样的测试对于计算实验模型的质量评估很重要。除计算实验模型的测试外，在计算实验模型生命周期的所有其他阶段，验证测试结果往往是报告的焦点，包括进行了哪些验证测试，结果是什么。

5) 分析。判断模型是否推动了复杂性分析，哪些复杂性（涌现、相变、幂律或其他长尾分布、临界、远程动态、串并联系统或其他结构特点）特征与特定模型相关？如果空间特征重要，计算实验是否采用了适当的空间指标和空间数据作为统计工具？基于实验运行的整体分析计划是怎样的？计算分析如何推进复杂系统的基础理论或应用理解？就整体有效性而言，模型是否回答了最初的研究问题？就仿真工具而言，模型是否具有通过大量计算分析回答问题的必要功能？在实施关键问题探索方面（如进化参数的探索），模型到底有多强大？实验平台基础设施的质量如何？

6) 应用。最后，计算实验模型的质量应根据“life-beyond-the-lab”原则进行评估。例如，在交流清晰度和透明度方面，是否提供各种有用的流程图和图表（如UML类、序列、状态和用例图）用于模型理解？在可重复性方面，模型的复制潜力或可行性是怎样的？在可视化分析方面，是否根据高标准实现了可视化？从“长期维护”的角度来看，模型代码和辅助文档是否持续可用？一些计算实验模型被用作政策分析工具，是否向决策人提供了新的、有用的可执行信息（新见解、似有道理的解释、预测、误差幅度、估计、贝叶斯更新）？

2.3 系统的复杂性的规律

复杂性是现代科学中最难定义的概念之一，至今无法给出统一的定义，也许根本就不存在统一的复杂性定义。但物理、生物、社会、意识这些现实世界的不同层次，各有性质不同的复杂性，相互之间不可混淆，既不可以拿低层次的复杂性代替高层次的复杂性，也不可以拿高层次的复杂性否定低层次的复杂性。人们可以从不同的角度、不同的层次来认识复杂性。本节首先对复杂性成因进行分类概述，然后以结构复杂性和分布复杂性为例进行详细阐述。这里的公式主要来自克劳迪奥·乔菲·雷维利亚（C.C.Revilla）的著作 *Introduction to Computational Social Science-Principles and Applications*，更多细节可参看原书。

2.3.1 复杂性的成因分类

据美国记者约翰·霍根（J.Horgan）在其著作《科学的终结》中所讲，麻省理工学院的物理学家塞思·劳埃德（S.Lloyd）提供了一份关于复杂性定义的清单，一共有45种之多，如分层复杂性、算法复杂性、随机复杂性、有效复杂性、同源复杂性、基于信息的复杂性、时间计算复杂性、空间计算复杂性等^[9]。这里选择几种典型的复杂性成因进行介绍。

（1）源于系统规模的复杂性

系统组件的数目代表系统的规模。在一定范围内，规模增大并不足以造成现有方法无法处理的复杂性。复杂性的形成需要足够大的系统规模，规模巨大会带来描述和处理上的困难，传统方法就会无能为力。简单系统不存在源于规模的复杂性，规模足够大（圣塔菲

要求系统有中等规模，钱学森要求达到巨系统规模^[10])是产生复杂性的必要条件，但不是充分条件，即使巨系统也不一定是复杂系统。

(2) 源于系统结构的复杂性

系统组件的多样性和差异性造成了相互之间关系的多样性和差异性，这是系统复杂性产生的根本原因。对于复杂性的成因而言，结构效应比规模效应重要得多。因为组件的差异越大，把它们整合起来的难度就越大。特别地，等级层次结构是复杂性的主要根源之一，复杂性研究者几乎都强调这一点。非等级层次结构只有元素和整体两个层次，相对简单；等级层次结构在元素与系统整体之间还包含中间层次，在元素层次上不能完成全部整合任务，需要经过不同层次逐级整合才能形成最终的系统。层次越多，越容易产生复杂性。按照盖尔曼^[11]对简单性和复杂性的词源学考察，英文“简单性”的原意是“只包含了一层”；“复杂性”则来源于“束在一起”，作为简单性的相反概念，具有“不止一层”的含义。系统是把多样性或多个层次捆绑在一起，其词义就隐含了复杂性来源于层次结构的观念。

(3) 源于开放性(环境)的复杂性

封闭系统没有复杂性，复杂性必定出现于开放系统。家务机器人被当作复杂系统，其复杂性并非来自系统规模或结构，而是来自家务劳动环境的多样性和极不规则性。但对外部环境开放也不是产生复杂性的充分条件。即使外部环境对系统的影响不能忽略，只要可以近似地当作对系统的干扰因素看待，或者系统行为可以归结为输出对输入的响应关系，就还是简单系统，使用传统的封闭系统加扰动的方法，或者黑箱方法，足以有效地进行处理。只有当外部环境对系统的作用无法再当作干扰，而是作为系统自身特性的有机构成成分时，封闭系统加扰动方法或者黑箱方法都失效，这种系统必然呈现某种复杂性。开放性也是复杂性产生的重要原因，系统与环境相互作用的复杂性是系统复杂性的重要表现。

(4) 源于动力学特性的复杂性

简单系统中的动力学因素往往可以忽略不计或者看作干扰因素。动力学过程可能产生无穷的多样性、差异性、丰富性、奇异性(包括分叉、突变、混沌等)和创新性，是产生复杂性的重要机制。复杂性一定具有某种动力学特性，但动力学因素并不是产生复杂性的充分条件，许多动力学系统(如经典控制论和运筹学处理的系统)仍然是简单系统。

(5) 源于非平衡态的复杂性

处于平衡态的系统都是简单的，不可能产生复杂性。非平衡态也不一定产生复杂性，但是复杂性只能出现于远离平衡的状态。在这种条件下，系统通过自组织形成耗散结构，即自组织地产生出复杂性。处于平衡态的系统，其耗散结构具有最小的复杂性。物理化学层次的耗散结构不可能具有生物复杂性，但只有具备了这种最小的复杂性，才可能进化出更高级的复杂性。

(6) 源于非线性的复杂性

哲学家早已指出，事物发展变化的终极原因是相互作用，但相互作用具有线性和非线性之分。线性意味着单一、均匀、不变，不具备产生复杂性的条件。线性系统通常是简单系统，线性相互作用产生的是简单性，无法造就复杂性。非线性相互作用意味着无穷的多样性、差异性、可变性、非均匀性、奇异性、创新性。元素之间、子系统之间的非线性相互作用是系统产生复杂性的内在机制，但非线性自身包含极大的差异性。弱非线性，仍然不可能产生复杂性，只是作为扰动因素处理，用线性模型加微扰的方法往往可以有效描述。

只有强非线性，特别是本质非线性，才可能产生复杂性。

(7) 源于不确定性的复杂性

随机性是一种重要的不确定性。平稳随机过程属于简单系统，非平稳过程才可能出现复杂性。简单地说，随机性是最大的复杂性，并无事实依据。物理系统的随机性规律一般只服从大数定律，仍属于简单性范围，可以用统计方法处理。但是，生命系统、社会系统的组成个体具有智能性，个体之间具有复杂的相互作用，只靠大数定律并不能揭示其本质特征。宏观整体特性不能仅看作大量微观个体相互碰撞的结果，现在的概率统计方法也不足以处理这类系统中的随机过程。例如，不可能从元胞特征出发依靠统计获得人体系统的整体特性，也不可能从个人特征统计数据获得社会系统的整体特性。

另一种重要的不确定性为模糊性，它既是复杂性的来源，又是复杂性的表现或结果。模糊集理论就是为处理复杂性而提出来的。不相容性原理认为，当系统的复杂性超过一定阈值时，描述的精确性与描述的价值互不相容，二者不可兼得，但目前的复杂性科学尚未涉足模糊性问题。

(8) 源于主动性、能动性的复杂性

在简单系统中，作用者与被作用者、原因与结果是界限分明的^[12]。不同组成个体之间、系统与环境之间互为因果，互动互应（所有个体都既是被动作用者，又是主动作用者）。一连串的、相互交叉的、网络式的因果联系，才能产生复杂性。特别是当个体有一定的自适应能力时，在不断适应环境的行为过程中必然产生出整体的复杂性。圣塔菲的一个基本信念就是适应性产生复杂性，所谓复杂适应系统（complex adaptive system, CAS）^[13]，就是在不断适应环境的过程中产生出复杂的系统。

(9) 源于个体智能的复杂性

由非智能个体构成的系统，即使通过自组织产生出复杂性，一般也是较为初步、低级的，这种复杂性总有办法解决，属于初级复杂性。由智能个体所构成的系统（如圣塔菲研究的CAS）能够辨识环境、预测未来、在经验中学习，以形成好的行为规则，使自身发生适应性变化，因而必定是复杂的。个体的智能程度越高，系统的复杂性也越高。个体智能是复杂性的重要根源之一。例如，地缘政治系统一般只包含为数不多的个体，属于小系统，至多算作大系统，但由于其个体是具有高级智能的主体，系统整体行为异常复杂多变，往往难以预测，因此属于开放的复杂巨系统。

(10) 源于人类理性的复杂性

以人作为构成要素的系统，其行为必须考虑人的理性因素所起的作用。尤其在竞争性系统中，博弈者的理性（智慧、谋略等）是产生复杂性的重要来源。但在完全理性（无限理性）假设下，复杂性的根源被抛弃了，博弈方都采取“最大-最小策略”。这种系统仍然是简单的，可按照运筹学处理，不完全理性（即有限理性）才可能产生复杂性。

西蒙（H.A.Simon）^[14]认为“人工性和复杂性这两个论题不可避免地交织在一起”，其中的人工性同时包含人的理性和非理性因素。

2.3.2 结构复杂性的规律分析

系统的结构复杂性是指系统和进程的组织方式，包括社会-技术-自然耦合系统^[13]及其组成部分。图 2-2 和图 2-3 是用网络图或树状图表示复杂系统和进程中形式同构的例子，

分别用因果和逻辑的形式表示。这里将仔细研究因果结构的性质以及它们如何产生系统复杂性，包括通过因果关系连接的串行复杂性，通过因果关系连接的并行复杂性，以及整体串行连接但结构内具有某些并行组件的混合复杂性。

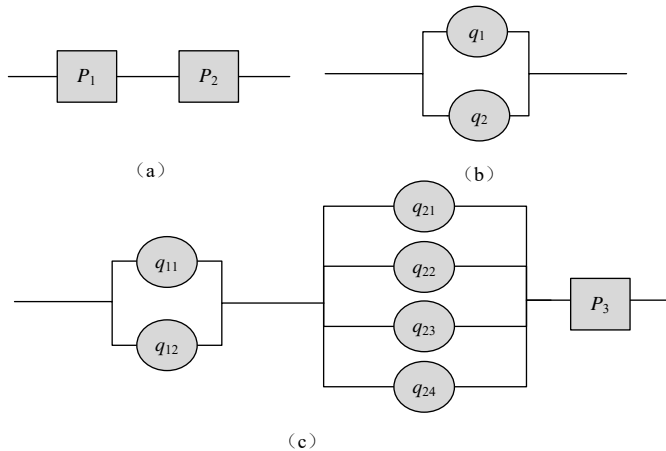


图 2-2 系统复杂性的因果关系结构模式

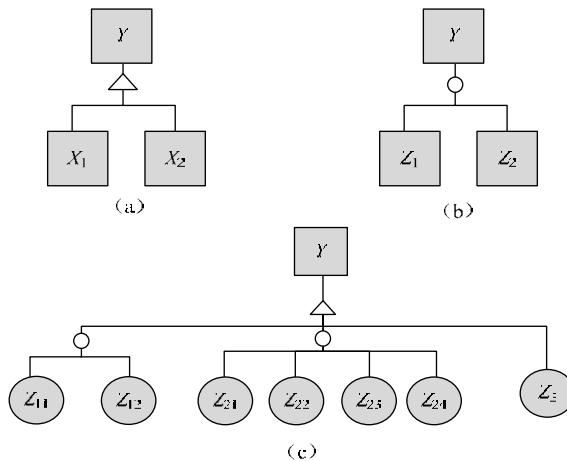


图 2-3 系统复杂性的逻辑关系结构模式

1. 连接造成的串行复杂性

在系统演化的进程中，复杂性往往是由复合事件产生的，而复合事件是从因果事件的关联中产生的。例如，在一个社会运行的标准模型中，成功的治理事件是由顺序过程产生的一个复合涌现：首先，集中解决重要部门的问题；其次，团体施压要求政府采取行动；再次，决策者通过颁布政策缓解社会压力；最后，该公共问题被缓解。

串行系统[图 2-2 (a) 和图 2-3 (a)]中复杂性产生的逻辑和因果关系概率表达式如下：

$$Y_s = \Psi_U(X_1, X_2, \dots, X_n) \Leftarrow X_1 \wedge X_2 \wedge \dots \wedge X_n \quad (2-1)$$

$$P_s = p_1 \cdot p_2 \cdot \dots \cdot p_n = \prod_{i=1}^n p_i \quad (2-2)$$

科学出版社
www.abook.cn

其中, Y_s 表示与所有必要因果条件整体关联的复合事件; Ψ_U 表示通过布尔逻辑的“与”运算符连接所有因果事件的结构化函数; X_i 表示第 i 个因果事件; 符号“ \wedge ”表示“与”连接; P_s 表示涌现事件发生的总体概率; p_i 表示第 i 个因果事件发生的概率。

串行连接的一个重要特点是当必要条件依次发生时, 则被称为顺序连接。注意, 发生概率是顺序因果事件的条件。在这种情况下, 式 (2-1) 和式 (2-2) 仍然需要考虑乘法的条件概率。但是, 不管是否考虑条件概率, 当系统复杂性被串行化时, 总体概率 P_s 还是会降低。一般而言, 如果复杂系统具有串行结构化特征, 它所拥有的执行概率要小于最差组件的执行概率。

2. 连接造成的并行复杂性

相比之下, 复杂系统也可以根据需要进行多个并发的行为。比如, 政府在抵御通货膨胀时, 通常采用以下几种政策的组合: ①价格控制; ②各种补贴 (食品、住房、药品); ③同时实施的其他计划。在图 2-2 (b) 和图 2-3 (b) 中, 相应示例包含了三个因果组成事件。

并行系统中复杂性产生的逻辑和因果关系概率表达式如下:

$$Y_p = \Psi_U(Z_1, Z_2, \dots, Z_m) \Leftarrow Z_1 \vee Z_2 \vee \dots \vee Z_m \quad (2-3)$$

$$P_p = 1 - (1 - q_1)(1 - q_2) \dots (1 - q_m) = 1 - \prod_{j=1}^m (1 - q_j) = 1 - (1 - Q)^{\Gamma} \quad (2-4)$$

其中, 符号遵循与式 (2-1) 和式 (2-2) 相同的约定。 Ψ_U 表示通过布尔逻辑的“或”运算符连接所有因果事件的结构化函数; P_p 表示涌现事件发生的总体概率; Q 表示因果事件出现的概率; Γ 表示因果事件出现的次数。根据 De Morgan 定律^[15], 可以很容易地证明并行化方程 (2-3) 和方程 (2-4) 是由串行化方程 (2-1) 和方程 (2-2) 演化而来。

并行系统的一个重要情形就是充分条件相互排斥, 这等同于布尔逻辑 XOR 运算符, 表示“任何一个”。在这种情况下, 因果事件的概率必须总计达到 1, 所以形成下列并行复杂性方程:

$$P_p = \Psi(Y_1, Y_2, \dots, Y_m) \Leftarrow Y_1 \vee Y_2 \vee \dots \vee Y_m \quad (2-5)$$

$$P_p = q_1 + q_2 + \dots + q_m = \sum_{j=1}^m q_j \quad (2-6)$$

不管因果概率是包含性 (OR) 还是排他性 (XOR), 当系统复杂性是基于并行结构时, 总体概率 P_p 总是在升高, 这在二阶因果关系和更高阶因果关系中也很常见。一般而言, 并行结构化的复杂系统所拥有的总体执行概率要大于最好组件的执行概率。

3. 混合结构造成的复杂性

现实世界的大多数复杂系统是通过串行和并行结构的某些组合运行的, 特别是那些复杂的人工制品或政策。图 2-2 (c) 和图 2-3 (c) 展示了这种结构复杂性的例子。串并行和并串行可以充当更复杂系统模型的构建模块, 以达到令人满意的结构复杂性。

串行-并行系统: 在图 2-2 (c) 和图 2-3 (c) 中所展示的结构复杂性, 包含了三个阶段的进程, 其中第一阶段和第二阶段分别执行两个和四个并行活动。现实中, 这种结构可以表示需要三个部门共同作用的社会系统 (如政府的立法、执行和司法部门), 其中第一个阶

段依赖于两个平行组件（如参议院和众议院）；第二个阶段依赖于四个平行组件（如安全、经济、卫生和基础设施）。

并行-串行系统：构建危机管理政策模型是现实中一个经典的例子，包括人道主义灾难、金融危机或网络安全等诸多问题领域。一阶复杂性是典型的串行结构，因为多个条件（如精确的情报、可用能力、执行方案等）必须一起发生。在灾害应急处理情况下，供应链管理也是一个突出的序列化结构，通信线路也是如此。在金融危机管理的情况下，立法和其他监管程序的过程具有类似的序列化结构。二阶复杂性通常是并行结构，因为所有条件都要通过多种不同的方法或战略来保证。例如，金融危机政策采用多重干预而不是单一的政府行为。

2.3.3 分布复杂性的规律分析

系统复杂性的特征还在于统计和概率分布，特别是非均衡分布^[16]和幂律分布^[17]。在过去的百年间，幂律分布已被证明存在于社会复杂性的多个领域。在几乎所有情况下，这些分布都是规模变量，而不是时间变量，这是一个有趣的特征，但目前仍没有很好的解释，依然有点科学之谜的意思。为了更好地鉴别和理解这一类复杂性，可以从定义幂律分布开始。幂律分布的不同表达形式如图 2-4 所示。

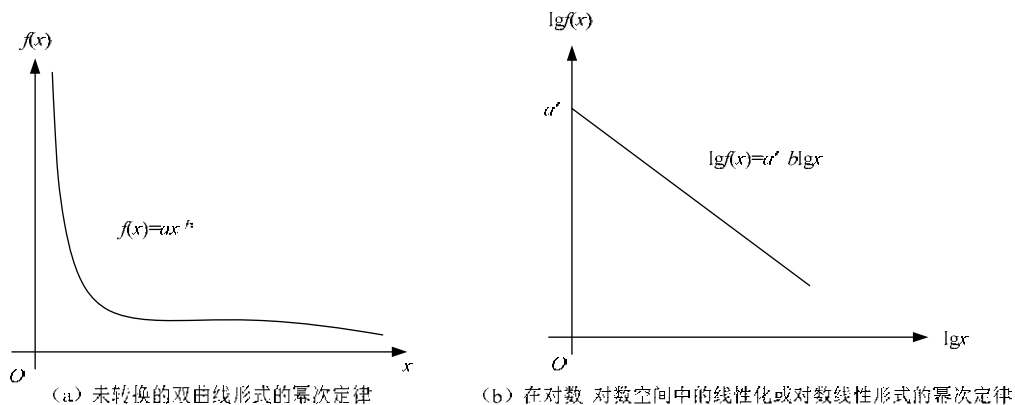


图 2-4 幂律分布的不同表达形式

定义 2.4（幂次定律） 令 X 是一个具有 $x \in \mathbf{R}$ 值集的实变量。在纯数学术语中，幂次定律的函数形式如式 (2-7) 所示。

$$f(x) \propto x^b = ax^b \quad (2-7)$$

其中，常数 a 和 b 可以采取任何值，使 $f(x)$ 在 X 中可以是增加的 ($b > 0$)、减小的 ($b < 0$) 或恒定的 ($b = 0$)，也可以是正数 ($a > 0$) 或负数 ($a < 0$)。但是，在社会复杂性理论的语境中，幂次定律总是表示为负指数 ($-b < 0$) 和正函数 ($a > 0$)。

在社会复杂性理论术语中，幂次定律是指具有以下形式的任意方程：

$$f(x) = ax^{-b} \quad (2-8)$$

利用以 10 为底数的对数转换对式 (2-8) 进行线性化，产生式 (2-9)：

$$\lg f(x) = a' - b \lg x \quad (2-9)$$

其中， $a' = \lg a$ ， a' 和 b 在对数空间中分别代表截距和斜率 [图 2-4 (b)]。

图 2-5 显示了幂律分布与其他分布模型。相比于正态（高斯或钟形）分布^[18]，幂律分布具有许多较小的值、一些（较少）中值以及几个稀有的极值。相比之下，正态分布中，最小值和最大值都非常稀少（趋于零的较小概率），而中值是常态。

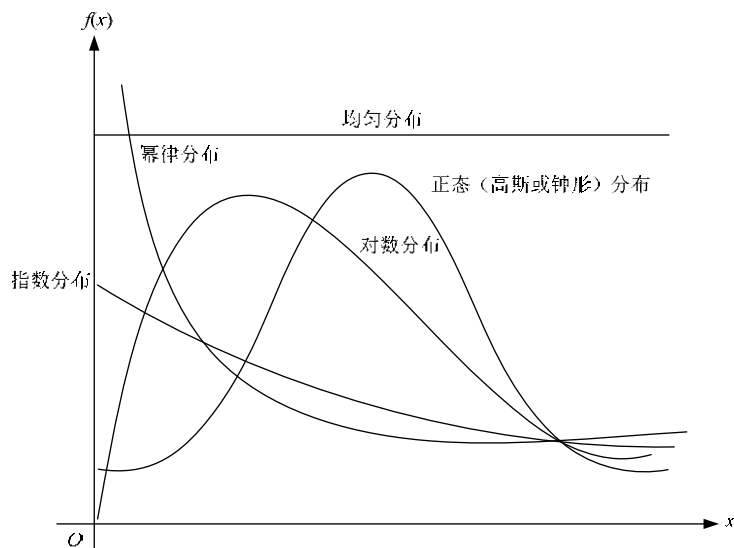


图 2-5 幂律分布和其他分布模型

如果采用不同的分布模型来理解社会复杂性，如指数分布、均匀分布、对数分布、正态分布等，会得出差异显著的结论。例如，极端事件在幂律分布中比在正态分布中出现的次数多出许多倍。社会复杂性的幂次定律适用于各种形式的分析，包括数据导向的实证分析以及数学导向的理论分析。这两种方法对于理解社会现象的复杂性是很有必要且互相协同的，都是假定一个给定的数据样本或一组观察值变量生成某种类型的幂律分布。根据现有文献中的常用做法，评估幂律模型与经验数据的拟合度经常采用以下两种方式：①判断对数坐标图中的数据点是否与幂律线性模型近似为一条直线；②采用不同的统计方法评估拟合度，如标准误差、 t 比率、Kolmogorov-Smirnov 实验、Anderson-Darling 实验等^[19,20]。

幂次定律建立了一种基于证据的经验规律，且具有强大的理论意义，可依据因变量类型对幂次定律进行分类，如图 2-6 所示。随着科学研究的发展，形式模型和经验数据之间的协同作用越来越重要。在从给定数据集寻找幂律模型的过程中，还需要注意到幂次定律在理解系统复杂性方面的一些潜在意义，这可能比给出幂律模型的公式更重要。下面给出了根据幂次定律理解系统复杂性的几个特性，这并不意味着每个幂次定律的实例都需要具有所有特性，但是对从哪些方面分析符合幂次定律的系统复杂性具有一定的启发性。

(1) 平均大小

幂律分布的平均值可由以下公式推导得出：

$$E(x) = \int_{\min(x)}^{\infty} xp(x)dx = \int_{\min(x)}^{\infty} (b-1)xf(x)dx$$

$$\begin{aligned}
 &= a(b-1) \int_{\min(x)}^{\infty} x^{1-b} dx \\
 &= \frac{a(b-1)}{2-b} x^{2-b} \Big|_{\min(x)}^{\infty} \\
 &= \frac{x_{\min}(b-1)}{b-2}
 \end{aligned} \tag{2-10}$$

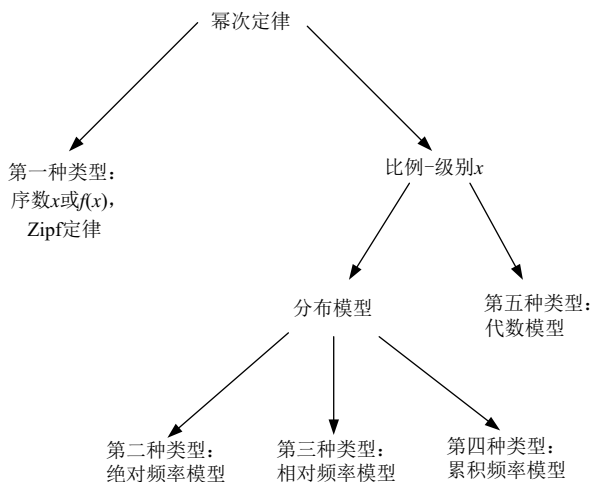


图 2-6 依据因变量类型对幂次定律进行分类

当 $b \leq 2$ 时，等式数值将趋近无限。换句话说，如果幂律的指数在 $0 < b < 2$ 或 $(b-1) < 1$ 范围内，其描述的系统现象不存在平均大小 [即不存在预期值 $E(x)$]。这是一个对组织规模、战争死亡和恐怖袭击等社会模式有深刻见解的理论结果。因此，阈值 $b = 2$ 在理论上是很关键的，因为它标记了一个边界值，用于区分具有有限平均和可计算大小 ($b > 2$) 的社会现象与那些缺乏预期值或平均大小 ($b \leq 2$) 的现象。

(2) 不等式

根据定义，幂律是一个不等式模型，所以每个幂律模型都有一个相关的洛伦兹曲线^[21]：

$$L(\phi) = 1 - [1 - \phi(x)]^{1/(b-1)} \tag{2-11}$$

和一个对应的 Gini 指数^[21]：

$$G(b) = 1 - 2 \int_0^1 L(\phi) d\phi = \frac{1}{2b-3} \tag{2-12}$$

这能通过经验方程式估算出^[22]：

$$\hat{G} = \frac{1}{n^2 E(x)} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \tag{2-13}$$

幂律的指数 b 与对应的 Gini 指数 G 之间存在着以下关联关系^[23]：

$$\begin{aligned}
 \text{重尾} \\
 (b \rightarrow 0) \Leftrightarrow \left\{ \begin{array}{l} \text{较多不等式} \\ \text{较少等式} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \text{较小的 } b \\ \text{较大的 } G \end{array} \right.
 \end{aligned} \tag{2-14}$$

$$\begin{aligned}
 \text{轻尾} \\
 (b \rightarrow \infty) \Leftrightarrow \left\{ \begin{array}{l} \text{较多等式} \\ \text{较少不等式} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \text{较大的 } b \\ \text{较小的 } G \end{array} \right.
 \end{aligned} \tag{2-15}$$

科学出版社
www.abook.cn

(3) 分形维数

如果幂律方程的指数 b 只能够取整数值 (1, 2, ...), 则与每个值相关联的频率将以此整数为比例实现反向递减。然而, 当 b 取分数值时, 比例范围本身是连续的, 不再像欧几里得空间 (Euclidean space)^[24] 中的数值那样是离散的。这就是为什么幂律的指数 b 通常被称为 Mandelbrot 分形维数^[25] 的原因。注意, 当 $b \rightarrow 0$ 时, 规模变化消失; 因为当 $b=0$ 时, X 的所有值都呈现相同的频率, 因此从规模变化的角度来看, 均匀随机变量存在于 0 维空间中。Zipf 定律^[26] ($b=1$) 产生一维空间。二次幂次定律 ($b=2$ 或临界值) 产生二维空间。

通常, b 幂律产生 b 维空间, b 分数值是嵌入在欧几里得空间内的分形维数。因此, 对于 $0 < b < 1$, 分形维数介于点和线之间; 对于 $1 < b < 2$, 分形维数介于线和平面之间; 对于 $2 < b < 3$, 分形维数介于平面和立方体之间; 以此类推。因此, 分形维数也为分析系统现象提供了一种新的分类方案, 物理学家已经开始探索这一方法^[27]。

(4) 临界和阈值驱动系统

规模变化可以由潜在过程产生, 通过输入的缓慢演变驱使系统处于临界状态。虽然可以通过连续输入驱动系统, 但是状态变量可能在关键区域内突然产生分叉和变动, 从而产生规模变化现象。几十年前, 数学家托姆^[28] 率先推出的突变理论为这一重要见解提供了参考。复杂性理论通过对动力学和亚稳态的新解释支持并拓展了突变理论。例如, 当利用幂次定律对给定的社会现象进行分析时, 应当注意一系列有关突变性的问题。

- 1) 这种现象是否受复杂性理论的阈值驱动系统支配?
- 2) 如何解释临界亚稳定状态的分叉集?
- 3) 限定在状态空间上的相关势函数 $p(x)$ 采用什么形式表示?

(5) 亚稳态

系统事件永远不会“出乎意料”——它们必须具有发展潜力, 才可能发生。从社会领域中给定的幂律公式可以产生另一个重要理论, 即“亚稳态”产生的理论条件。如果能够在一系列扰动情况下仍保持平衡, 则系统 (或者更准确地说, 给定状态为 $x \in X$ 的系统) 被称为是 Lyapunov 稳定^[29]。例如, 能够经受各种压力的社会关系 (如婚姻、友谊、联盟), 可以认为它是稳定的。相反, 如果社会关系在压力下分崩离析, 如在冲突或待解决问题的压力下结束的政治或联盟, 则认为它是不稳定的。

当系统中存在一个或多个潜在状态 ($x' \in X$) 或潜在运转机制 ($x \neq x'$), 在某些条件下会导致当前状态发生转变, 这种系统被称为亚稳态系统。亚稳态在许多社会系统中普遍存在。例如, 在经济学中, 金融市场在产生了能够导致市场崩溃的“泡沫”时, 它就会变成亚稳态。同样, 从更积极的角度看, 当恢复市场平稳的潜力增加时, 崩溃状态变为亚稳态。幂律模拟社会情境, 其中包括大范围、可实现的状态, 而不仅是现存的平衡状态或观察到的状态。系统演化的理论应当利用幂律中固有的亚稳态概念。

(6) 长程相互作用

大部分仅由最近邻相互作用控制的系统倾向于产生正态 (或高斯) 分布现象, 或者其他非幂律现象, 如在上、下分位点的尾部明显更短或更窄。相比之下, 由长程相互作用控制的“全局”系统受制于产生幂律的非平衡动力学过程。在这样的系统中, 极端事件的发生率比“正常” (高斯) 平衡系统高出几个数量级 (不仅是更大)。在社会系统中, 长程相

互作用的空间维度在社会或物理距离方面是相当简单的，时间长程相互作用则是指对过去的持久记忆以及对未来的期望。

2.4 系统复杂性的解释

理论的一项基本功能就是解释现象。因此，理论需要根据初始或驱动事件对所观察现象进行说明，理论的论点必须符合科学解释模式。当代模型和系统复杂性理论起源于18世纪，通过数学媒介将所积累的知识形式化。基本概率、决策模型和图形理论模型是最早运用的数学结构，其次是微分方程的动态系统、博弈论、差分方程、随机过程、模糊集合，最后是用于进行社会模拟和决策分析的计算模型。

2.4.1 系统复杂性理论：解释

科学理论的典型特性是其必须包含关联前因（诱因）和后果（影响）的故事或因果叙事。系统复杂性理论根据因果关系来解释系统出现的复杂性现象，其特点在于可解释所观察到的事实或可用数据下的经验模式。

定义 2.5（系统复杂性的出现） 假设系统复杂性表现为宏观层面上的综合事件 C ，它的产生源自于微观层面上基本事件（样本空间 Ω 内的样本点）与适应过程中个体间博弈的组合。

在系统复杂性的定义中，综合事件 C 包含两种构成成分：①一系列基本的宏观事件（由博弈结果和适应过程等相关自然状态构成）；②将此类事件关联在一起的操作规则。鉴于系统复杂性因个体博弈而显现（与大多数自然状态结果相对），系统复杂性出现的建模和解释都取决于个体博弈的结果。这种方法基于理论，同时也依赖于个体制定决策的微观基础。

根据因果事件在样本空间内出现或者未出现的情况，可采用事件发生的因果逻辑来解释系统复杂性的出现。例如，对于某种国际联盟组织的出现而言，与国家战略、领导能力、利益均衡、征募同盟等有关的事件必须以特定组合的形式出现或不出现，或者以一种等同有效组合的形式出现。因果事件必须以一种非任意形式出现，从而使系统复杂性能够显现。对于即将发生的群体行为来说，关键的因果事件组合必须以一种特殊的方式出现；否则群体行为便不会出现。通常情况下，系统复杂性的出现可归因于更加基本且有时无法观测的自然状态和博弈结果。将因果事件映射到当时情景中的工具可说明系统复杂性。

定义 2.6（事件函数） 鉴于复合事件 Y 的出现或未出现通常和一系列其他事件 $\{X\}$ 有关，映射 $\varphi:\{X\} \rightarrow Y$ 被称为 Y 的事件函数。因此，事件 $Y = \varphi(X)$ 。

事件函数 $\varphi(X)$ 定义了因果解释，其在实际意义上指因果论证的功能函数建模。从计算角度看，表示编写包含众多内嵌函数的代码。根据这些定义，系统复杂性出现的事件函数可作以下定义。

定义 2.7（系统复杂性出现的事件函数） 鉴于系统复杂性的综合事件 C 的出现或不出现通常和一系列其他事件 $\{X\}$ 有关，所以映射 $\varphi:\{X\} \rightarrow C$ 被称为 C 的事件函数。因此，事件 $C = \varphi(X)$ 。

在形式上，事件函数可以很详细地表示出精确因果逻辑，解释如何产生综合事件。但

是，存在哪些事件函数？不同事件函数如何解释综合事件的发生？如何根据事件函数确定综合事件发生的概率？为了回答这些问题及其他类似问题，就必须在微观水平上查看系统复杂性的逻辑，根据顺序逻辑和条件逻辑辨识两种因果解释模式。

根据西蒙理论^[14]的概念和假设（环境复杂性、目标寻求行为、有限理性、适应性、近乎可分解性），前进因果逻辑的顺序树可用于适应过程和复杂度的模型创建，详见图 2-7。在某些初始时间点 r_0 ，在特定环境中发生了事件 E 。在后续时间点 r_1 ，为了应对当前环境的挑战，个体可根据有限理性决定是否发生适应性改变，即事件 D 。若它们并未做出改变的决定（即事件 $\sim D$ ），则它们将继续忍受与时间点 r_0 之前相同的环境影响，无论是什么样的后果（结果 E ）。如果它们决定适应，则它们可选择是否实际执行决策和落实适应性措施（即事件 A ）。如果它们未能成功执行事件（即事件 $\sim A$ ），则它们仍需要忍受环境影响，只是此时已经过去了一段时间，产生结果 E^* 。可论证 $S(E) \approx S(E^*)$ 。

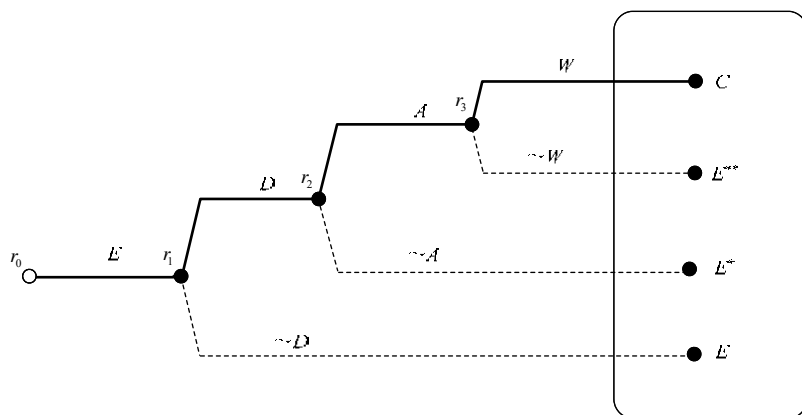


图 2-7 西蒙适应理论和系统复杂性出现的前进因果逻辑顺序树

如果个体已经采取了某些行为，则在 r_3 时刻，行为反馈可能有效，也可能无效。若反馈有效（事件 W ），则结果为成功且具有更高的复杂度，因为现在必须维护更为复杂的系统（结果 C ）。若反馈无效（事件 $\sim W$ ），则结果为失败后仍需忍受环境影响（结果 E^{**} ）。可论证 $S(E^{**}) > S(E)$ ，其中 $S(X)$ 代表与事件 X 有关的影响或无效性。

图 2-7 的模型提供了关于西蒙理论的一阶表征。主要结果为 Ω 空间内的各种结果均可通过组合产生。尤其，系统复杂性 C 的出现最少需要四种必要的顺序状况，具有显著的低概率；否则就无法成功出现。其他结果（失败 E 、 E^* 和 E^{**} ）发生概率相对较高，因此总体发生的概率更高。

2.4.2 基于顺序逻辑的建模解释

在顺序逻辑模式^[30]下，系统复杂性作为综合事件出现，可通过当时出现的序列或之前的事件路径进行解释。在这种模式下，系统复杂性 C 的出现被解释为经过多项博弈和决议的分支过程 P 是样本空间 Ω 内的一种结果，即可能事件之一。顺序逻辑通常将解释重点放在面向流程的因果论证上。在众多备选结果中，以顺序逻辑发生的综合事件通常被解释为可能的结果，而不是必定出现的结果。

顺序逻辑会引起扩展式博弈，即使用顺序事件树介绍因果过程。起始事件 I 代表顺序

过程 $P_N(I \rightarrow C)$ 开始, 在第 N 个分支节点之后产生某些事件 C , 所以 C 发生概率极小, 除非发生一系列特定事项。空间 Ω 内的起始事件 I 和结果之间的分支节点由博弈结果决定, 由个体选择、自然状态和博弈理论产生。因此, 根据顺序逻辑, 从起始状态演变的可能事件会产生系统复杂性。

假设 2.1(系统复杂性的顺序因果逻辑) 系统复杂性 C 在未来任意时间 t 内, 在自 $t-n$ 开始的分支程序的样本空间 Ω 内, 作为未来可能的结果而出现。事件节点导致系统复杂性 C 的出现:

$$C = X_{t-1} \wedge X_{t-2} \wedge \cdots \wedge X_{t-n+1} \wedge X_{t-n} \quad (2-16)$$

其中, 时间指数 t 表示 C 出现之前的时间, 而各事件均独立于之前事件而存在, 拥有基于条件概率的顺序概率:

$$\Pr(C) = p_{-n} \cdot p_{-n+1} \cdot p_{-n+2} \cdots p_{-1} = \prod_{i=0}^{n-1} p_{-n+i} = P^A \quad (2-17)$$

其中, $A = 0, 1, 2, \dots, n-1$; $p_{-n} = \text{Pt}(X_{-n})$, 针对第一个(初始)事件;

$p_{-n+1} = \text{Pt}(X_{-n+1} | X_{-n})$, 针对第二个事件;

$p_{-n+2} = \text{Pt}(X_{-n+2} | X_{-n} \wedge X_{-n+1})$, 针对第三个事件;

\vdots

$p_{-1} = \text{Pt}(X_{-1} | \text{所有之前事件})$, 针对 C 之前发生的最后一个事件;

A 表示 C 发生之前的事件数量, 或者事件流程的长度;

$P = p_n = p_{-n+1} = p_{-n+2} = \cdots = p_{-1}$, 表示各事件的发生概率均相同。

通常情况下, C 之前的所有事件, 例如式(2-16)中 X_i 之前的顺序, 其构成了 C 的发生可能性, 或者“ C 实现的可能性”。

定理 2.1(系统复杂性的顺序低概率) 当结果 C 的之前事件尚未发生时, C 的有限概率(突然发生的概率)始终: ①比个别概率 P 小; ②比之前事件的最小概率小。正常情况如下。

$$\text{Pt}(C) < \min(p_{-n}, p_{-n+1}, p_{-n+2}, \dots, p_{-1}) < P \quad (2-18)$$

从顺序逻辑视角看, 系统复杂性的另一个理论是必须处理之前事件变动和分支流程长度变动所造成的不同影响。判断哪一项对于 $\text{Pt}(C)$ 的影响较大: p_{-n+i} 发生变化, 或者 A 发生变化, 可通过下列原则得出精确答案。

定理 2.2(顺序概率对于之前事件发生概率的依赖性) 系统复杂性 C 的顺序概率与之前事件 P 发生概率的变化率, 可通过下列表达式计算得出。

$$\frac{\partial \text{Pr}(C)}{\partial P} = AP^{A-1} \quad (2-19)$$

其始终为正值, 所以 $\text{Pt}(C)$ 在 P 方向出现内凹。

定理 2.3(顺序概率对于过程长度变化概率的依赖性) 系统复杂性 C 的顺序概率与之前事件发生数量 A 的变化率, 可通过式(2-20)计算得出。

$$\frac{\Delta \text{Pr}(C)}{\Delta A} = P^{A+1} - P^A \quad (2-20)$$

其始终为负值, 所以 $\text{Pt}(C)$ 在 A 方向出现内凸。

两项依赖方程均为非线性方程, 与综合事件(如 C)的复杂度一致。这些定理构成了回答之前问题的基础。

定理 2.4 (顺序主导原则) 相比于分支流程内之前因果事件的概率 P , 系统复杂性结果 (C) 的顺序概率对于事件发生数量 A 更加敏感。正常情况如下。

$$S_p > S_A \quad (2-21)$$

因为,

$$\frac{\partial \Pr(C)}{\partial P} \cdot \frac{P}{\Pr(C)} > \frac{\Delta \Pr(C)}{\Delta A} \cdot \frac{A}{\Pr(C)} \quad (2-22)$$

很多情况下, 主导原则违反直觉, 因为直觉会将更大的关注点放在因果关系基数上, 而不是放在概率上。通俗地说, 导致 C 的之前因果事件更多地看重个体, 而不是看重个体总量, 或者更改之前因果事件发生概率比更改其发生总次数更加重要。在没有正常分析的情况下, 这种答案并不是一种直接答案, 可通过计算方法进行验证。

2.4.3 基于反向逻辑的建模解释

在条件逻辑模式下, 当系统复杂性作为综合事件 C 出现时, 可通过提供必要或充分条件进行解释。条件逻辑将解释重点放在因果论证的布尔结构上, 根据当前位点研究背景状况, 因此得到了术语反向逻辑。最好将条件逻辑模式下综合事件的发生概率描述为必须解释的前提条件, 而不是可能的处理结果。

假设 2.2 (系统复杂性的条件因果逻辑) 系统复杂性 C 以双重因果模式出现: 或者基于必要状况同时出现[事件 (X_1, X_2, \dots, X_N) 通过布尔逻辑连接词 AND (与) 连接]; 或者基于任意充分条件发生[事件 (Z_1, Z_2, \dots, Z_M) , 通过布尔逻辑连接词 OR (或) 连接]。正常情况如下。

AND 连接产生事件 C_X :

$$C_X = \phi \wedge (X_1, X_2, \dots, X_N) = X_1 \wedge X_2 \wedge \dots \wedge X_N \quad (2-23)$$

OR 连接产生事件 C_Z :

$$C_Z = \phi \vee (Z_1, Z_2, \dots, Z_M) = Z_1 \vee Z_2 \vee \dots \vee Z_M \quad (2-24)$$

1. 序列复杂性: 逻辑结合交集

系统复杂性理论的下列原则主要源于综合事件的概率定理运用。

定理 2.5 (系统复杂性结合原则) 系统复杂性 C 的概率取决于必要事件共同发生的概率。正常情况如下。

$$\Pr(C_X) = \Pr(\wedge X_i) = \prod_{i=1}^n \Pr(X_i) = p_1 p_2 \dots p_n = P^\theta \quad (2-25)$$

其中, θ 为导致 C 发生的必要事件的数量 ($2 < \theta < n$), 而 P 为这些事件的发生概率。当把系统复杂性的出现理解为微观事件产生的宏观综合事件时, 定理 2.5 构成了系统复杂性理论的计算基础。对比式 (2-25) 和式 (2-18), 可以看出顺序逻辑模式是很特殊的结合, 也将其称为基于顺序条件 (AND) 的结合。因此, 系统复杂性顺序模型的某些属性也适用于所有的反向逻辑模型, 低概率、依赖和主导原则为其中最重要的因素。

2. 并行复杂性: 逻辑分离集合

定理 2.6 (系统复杂性分离原则) 系统复杂性 C 的分离概率可通过式 (2-26) 给出。

$$\begin{aligned}
 \Pr(C_Z) &= \Pr(\vee Z_j) \\
 &= 1 - \prod_{j=1}^m [1 - \Pr(Z_j)] \\
 &= 1 - (1 - q_1)(1 - q_2) \cdots (1 - q_m) \\
 &= 1 - (1 - Q)^\Gamma
 \end{aligned} \tag{2-26}$$

其中， Γ 为导致综合事件 C 的因果事件出现的次数 ($2 < \Gamma < m$)，而 Q 代表其出现的概率。根据定理 2.6，可以很容易得出系统复杂性不出现的概率，即 $\neg C_Z$ 事件出现的概率为 $1 - \Pr(C_Z)$ ，具体如式 (2-27) 所示。

$$\prod_{j=1}^m [1 - \Pr(Z_j)] = (1 - Q)^\Gamma \tag{2-27}$$

定理 2.7 (超概率原则) 当多种因果事件导致系统复杂性 C 出现时， C 的出现概率总是：①比个别因果事件的出现概率 Q 要大；②比因果事件的最大出现概率要大。正常情况如下。

$$\Pr(C) > \max\{q_1, q_2, \dots, q_m\} > Q \tag{2-28}$$

超概率和低概率原则强调了系统复杂性基于因果结构分离与合并的特征属性。那么，冗余事件的发生概率和发生次数如何影响系统复杂性的出现概率呢？会产生哪些主要影响呢？分离系统复杂性的依赖性、敏感性和主导原则需要遵循定理 2.6，可采用与结合模式类似的多变量分析。

定理 2.8 (依赖冗余事件的出现概率 Q) 系统复杂性事件 C 出现的概率随着相应因果事件发生概率 Q 的变化而变化，可通过式 (2-29) 表达。

$$\frac{\partial \Pr(C)}{\partial Q} = \Gamma(1 - Q)^{\Gamma-1} \tag{2-29}$$

其始终为正值，所以 $\Pr(C)$ 在 Q 方向出现内凹。

定理 2.9 (依赖冗余事件的出现数量 Γ) 系统复杂性事件 C 的出现概率随着相应因果事件发生数量 Γ 的变化而变化，可通过式 (2-30) 表达。

$$\frac{\Delta \Pr(C)}{\Delta \Gamma} = Q(1 - Q)^\Gamma \tag{2-30}$$

其始终为正值，所以 $\Pr(C)$ 在 Γ 方向出现内凹。

这两种依赖公式均为非线性方程，与之前的结果一致。但是，注意冗余/充分性 Γ 与必要性 Λ 具有相反的效果。这些定理构成了复杂性解释的理论基础，表现出冗余事件对于系统复杂性事件 C 出现概率的主要影响。

定理 2.10 (分离系统复杂性的主要原则) 系统复杂性事件 C 的出现概率相比于冗余/充分因果事件的出现概率，比事件出现的次数 Γ 更加敏感。正常情况如下。

$$S_Q > S_\Gamma \tag{2-31}$$

因为，

$$\frac{\partial \Pr(C)}{\partial Q} \cdot \frac{Q}{\Pr(C)} > \frac{\Delta \Pr(C)}{\Delta \Gamma} \cdot \frac{\Gamma}{\Pr(C)} \tag{2-32}$$

在布尔 OR (和/或) 连接词条件下，系统复杂性之前的所有结果均有效。在逻辑上，

后面的连接词指“同或”，与其不同的是“异或”，也被称为布尔 XOR。

3. 混合复杂性：少数个别原因

前面的两种因果关系（即结合和分离），代表纯因果模式，也就是将系统复杂性 C 建模表示为必要或充分条件的结果。但是，这两种极端的因果关系中间，还有很多导致系统复杂性产生的部分必要条件或部分充分条件。这种情况会发生在（海量范围内）多种条件（超过一种）发生时。例如，群体行为，可能不是因为系统内个体构成整体而发生，也不是因为单个个体独自行为而发生，但可能因为某些核心组织而发生。

假定系统中包含 m 个个体，而系统复杂性至少需要其中的 v 个个体通过组合才能够产生，其中， $m > v > 1$ 。这意味着支持 C 的因果组合数量（即使并不是所有的组合都合理可行）可由式（2-33）给出。

$$\binom{m}{v} = \frac{m!}{(m-v)!v!} \quad (2-33)$$

其中， $m! = m(m-1)(m-2)\cdots 1$ 表示 m 的阶乘。其中一些复杂性非常重要，因为它在形式上会降至：①纯粹连接情况 $v \rightarrow m$ （定理 2.5）；②纯粹分离情况 $v \rightarrow 1$ （定理 2.6）。

因此，基数 v 是关键的模式变量：在上限水平（ $v \rightarrow m$ ）时，复杂性由必要条件（低概率）组合产生；而在下限水平（ $v \rightarrow 1$ ）时，条件彼此分离（超概率）。当过多个体属于同一联盟时，会出现超概率。若个体开始离开且联盟达到最小生存水平上，则低概率开始出现，直至达到关键阈值位置，此后联盟崩溃。

定理 2.11（少数个别原则） 在事件函数中，系统复杂性事件 C 的出现是因为集合（ m ）中一系列可能或可行的条件（ v ）发生所导致，其发生概率可通过式（2-34）表示。

$$\Pr(C) = \sum_{i=v}^m \binom{m}{i} P^i (1-P)^{m-i} \quad (2-34)$$

其中， P 为前因事件的发生概率，且 $i = 1, 2, \dots, v, \dots, m$ 。

系统复杂性的很多方面可归因于组合复杂性。当 $v \rightarrow m$ 且 $v \rightarrow 1$ 时，这种部分必要/充分原则可分别归结为比较简单的结合与分离原则。这是系统复杂性中的一种显性表现，包含结合因果结构和分离因果结构。

2.5 小 结

科学研究的目的是实现规律描述与理论解释。在系统复杂性科学中，某些情境难以在实际环境中验证某些科研想法，计算实验由此应运而生。计算实验能够为系统复杂性的理解提供一些独到见解，进而支持跨文化和跨学科领域的研究。

本章首先从计算实验的目的出发，让读者理解计算实验的特有属性以及相应的运作机制；进而在计算实验的模式中，对计算实验的参照系统、实验平台和实验模型进行了详细介绍，并对计算实验的质量维度评价进行了介绍。其次，从系统规模、结构和开放性环境等方面，介绍了复杂性的成因，并采用相关定理、公式，对结构复杂性和分布复杂性的规律进行了具体分析。最后，给出了系统复杂性的理论解释原则，并以顺序逻辑建模过程与反向逻辑建模过程为例，说明了复杂性产生的过程。

参 考 文 献

- [1] DRABEK T E, MCENTIRE D A. Emergent phenomena and the sociology of disaster: lessons, trends and opportunities from the research literature[J]. *Disaster Prevention & Management*, 2013, 12(2): 97-112.
- [2] TAINTER J A. Social complexity and sustainability[J]. *Ecological Complexity*, 2006, 3(2): 91-103.
- [3] HEMASPAANDRA L A . Complexity theory[J]. *Acm Sigact News*, 2001, 32(3): 40-52.
- [4] 王飞跃. 计算实验方法与复杂系统行为分析和决策评估[J]. *系统仿真学报*, 2004, 16(5): 893-897.
- [5] 唐建国. 数学归纳法在高等代数教学中的应用[J]. *湖南科技学院学报*, 1996(s1): 44-46.
- [6] 何斌, 王若恩. 物元演绎推理[J]. *系统工程理论与实践*, 1998, 18(1): 85-92.
- [7] MORITZ H. Geodetic reference system 1980[J]. *Journal of Geodesy*, 1980, 54(3): 395-405.
- [8] REBITZER G, EKVAL T, FRISCHKNECHT R, et al. Life cycle assessment: Part 1: Framework, goal and scope definition, inventory analysis, and applications[J]. *Environment international*, 2004, 30(5): 701-720.
- [9] HORGAN J. The end of science: facing the limits of knowledge in the twilight of the scientific age[J]. *Isis*, 1996, 20(4): 113-114.
- [10] 王寿云. 开放的复杂巨系统[M]. 杭州: 浙江科学技术出版社, 1996.
- [11] GELL-MANN M, RAMOND P, SLANSKY R. Complex spinors and unified theories[M]// Murray Gell-Mann: Selected Papers, 2010: 266-272.
- [12] 张彩江, 王春生. 简单系统决策与复杂系统决策: 对决策范式中分化和转换的一种理解[J]. *系统管理学报*, 2005, 14(5): 418-423.
- [13] MILLER F P, VANDOME A F, MCBREWSTER J. Complex adaptive system[J]. *Springer Encyclopedia of Pain*, 2007(DOI=10.1007/978-3-540-29805-2_820).
- [14] NEWELL A, SIMON H A. Human problem solving[M]. New Jersey: Prentice-Hall, 1972.
- [15] CAMELLO O, JOHNSTONE P. De Morgan's law and the theory of fields[J]. *Advances in Mathematics*, 2009, 222(6): 2145-2152.
- [16] BEZDEK J C, HATHAWAY R J. VAT: a tool for visual assessment of (cluster) tendency[C]// International Joint Conference on Neural Networks. Paris, 2002, 3: 2225-2230.
- [17] CLAUSET A, SHALIZI C R, NEWMAN M E J. Power-law distributions in empirical data[J]. *Siam Review*, 2009, 51(4): 661-703.
- [18] CHHIKARA R. The Inverse Gaussian Distribution: Theory: Methodology, and Applications[M]. Boca Raton: CRC Press, 1988.
- [19] BAUKE H. Parameter estimation for power-law distributions by maximum likelihood methods[J]. *European Physical Journal B*, 2007, 58(2): 167-173.
- [20] 彭作祥. 一类 Hill 型估计量的收敛性[J]. *西南师范大学学报 (自然科学版)*, 1998(2): 133-137.
- [21] KAKWANI N. Applications of Lorenz curves in economic analysis [J]. *Econometrica: Journal of the Econometric Society*, 1977, 45(3): 719-727.
- [22] KLEIBER C, KOTZ S. Statistical size distributions in economics and actuarial sciences[M]. New York: John Wiley & Sons, 2003.
- [23] GASTWIRTH J L. The estimation of the Lorenz curve and Gini index[J]. *Review of Economics & Statistics*, 1972, 54(3): 306-316.
- [24] HARRISON J. A HOL theory of Euclidean space[C]// International Conference on Theorem Proving in Higher Order Logics. Springer, Berlin, Heidelberg, 2005: 114-129.
- [25] MANDELBROT B B. Self-affine fractals and fractal dimension[J]. *Physica Scripta*, 1985, 32(4): 257-260.
- [26] VOGT P. Exploring the robustness of cross-situational learning under Zipfian distributions[J]. *Cognitive Science*, 2012, 36(4): 726-739.
- [27] SORNETTE D. Critical market crashes[J]. *Physics Reports*, 2003, 378(1): 1-98.
- [28] SUSSMAN H J, THOM R, FOWLER D H. Structural stability and morphogenesis: an outline of a general theory of models[J]. *Contemporary Sociology*, 1977, 6(5): 544.
- [29] SHEVITZ D, PADEN B. Lyapunov stability theory of nonsmooth systems[J]. *IEEE Transactions on automatic control*, 1994, 39(9): 1910-1914.
- [30] PSARAKIS M, GIZOPOULOS D, PASCHALIS A, et al. Sequential fault modeling and test pattern generation for CMOS iterative logic arrays[J]. *IEEE Transactions on Computers*, 2000, 49(10): 1083-1099.

第3章 计算实验的方法框架

3.1 引言

计算实验方法^[1-3]基于分布式思想,采用自底向上的方法,利用分散的微观智能模型模拟整个复杂系统,通过模拟系统中各类 Agent 实体的微观行为,再将这些实体通过相互作用集合起来反映整个系统演化发生的宏观规律,实现由简单元素互相作用而形成的复杂现象。与传统方法相比,计算实验方法可以模拟现实系统在各种场景中的演化过程,包括现实场景和从未发生过的虚拟场景,这使得研究者更容易探索复杂系统的运行规律并找到有效的干预措施。但是,计算实验方法并不适用于所有领域,对它的使用也有较高的技术门槛。计算实验方法的研究框架如图 3-1 所示,包含六个步骤:现实建模、系统开发、实验执行、结果分析、模型评估与结果应用。下面将对每个步骤进行详细阐述。

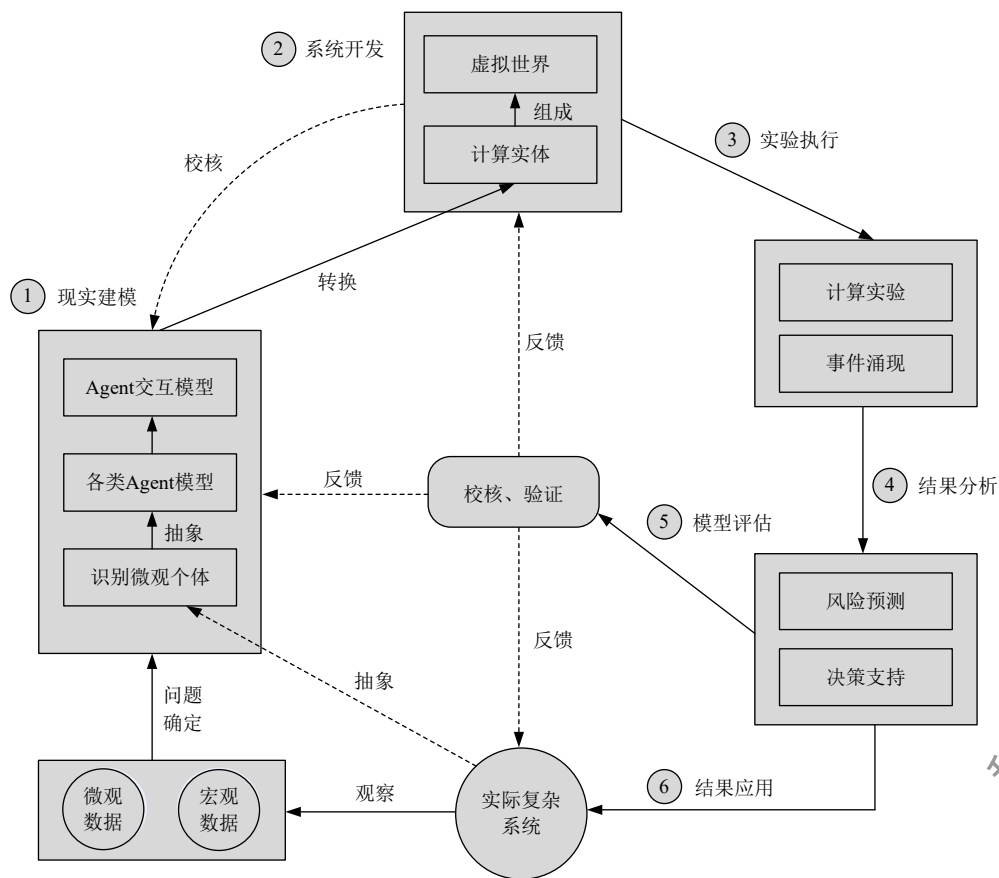


图 3-1 计算实验方法的研究框架

3.2 现实建模

所有复杂系统仿真都旨在解决基于参照系统所定义的一个或多个研究问题。事实上，参照系统在很大程度上由研究问题所定义，两者之间存在协同关系。研究问题是科学探索的主要引擎，在很大程度上确定了计算实验与现实世界的一致性程度。一般来说，计算实验的表现形式比较简单，而实际系统的运行比较复杂。两者之间的差距主要表现在两个方面：①如果计算系统抽象程度过高，就会难以反映现实世界的运行规律；②如果计算系统的抽象程度过低，建立模型的复杂度就会很高，会遇到数据缺乏、资源和时间不足、知识体系尚未建立等一系列问题。因此，在真实世界与计算系统之间实现适当的映射是非常重要的。在计算模型的设计中，模型设计原则应该与真实情况之间达成平衡，如图 3-2 所示。

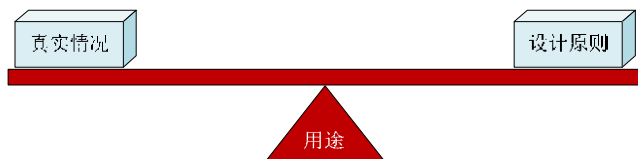


图 3-2 模型设计原则与真实情况的平衡关系

在计算模型设计过程中，以下方法违背了计算原则与真实情况的平衡关系。

1) 激进现实主义方法：该模型尽可能逼近现实，模型规模大、参数众多且很多模块未经测试。该模型设计原则与真实情况的平衡关系如图 3-3 所示。

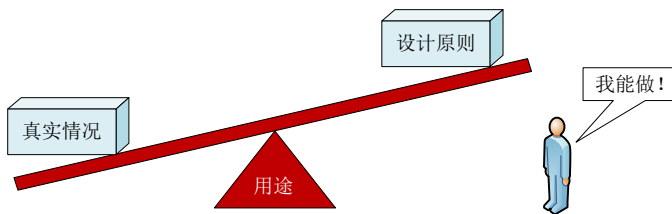


图 3-3 激进现实主义方法模型设计原则与真实情况的平衡关系

2) 简约性方法：该模型遵守 KISS (keep it simple and stupid, 尽量简单) 原则，模型规模小且参数较少。该模型设计原则与真实情况的平衡关系如图 3-4 所示。

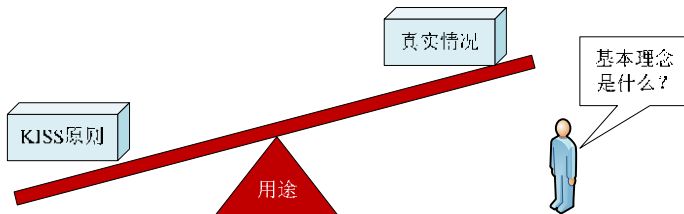


图 3-4 简约性方法模型设计原则与真实情况的平衡关系

3) 竞争性方法：该模型以实验为基础，以解决应用问题为导向，其特征是规模大、部分模块已测试。该模型设计原则与真实情况的平衡关系如图 3-5 所示。

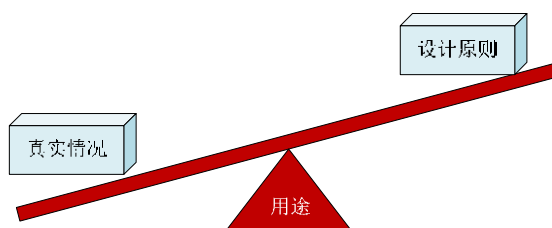


图 3-5 竞争性方法模型设计原则与真实情况的平衡关系

4) 搭积木法: 该模型的特征是规模小且模块化, 专注于必要的理论组成部分, 通过搭积木的方法增加已经成熟的模块。该模型设计原则与真实情况的平衡关系如图 3-6 所示。

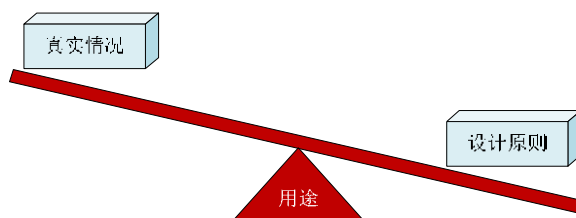


图 3-6 搭积木法模型设计原则与真实情况的平衡关系

模型设计是在某一个层面上对现实与抽象取得平衡, 所以无法预期该模型在不同层面上都表现良好。例如, 高度现实的模型可能有重大的政策价值, 却只有很少或根本没有理论价值; 反之, 高度抽象的模型可能提供深刻的科学见解, 但就政策贡献而言, 只能提供极少可以直接应用的结果。根据与所给定参照系统的逼近程度, 按照“抽象→现实”的路径划分计算实验模型的类型, 以便理解其价值和限制。

1) 高度抽象: 模型仅与参照系统有少数定性相似之处, 并没有试图复制任何量化特征。这些模型主要用于基础科学的理论分析, 而非运营策略分析。这类模型有时候被称为玩具模型, 因为它们代表了问题领域内非常简单的参照系统。早期的一些社会仿真模型^[4]就属于这一类型, 但是它们至今依然很重要, 提供了理解人类和社会基本运行规则的独特方法。例如, 热虫 (Heatbugs) 模型^[5]、鸟群 (Boids) 模型^[6]等, 对于理解社会仿真的基本原理具有重要价值。

2) 适度抽象: 模型能够显示令人信服的定性特征且符合某些定量标准。虽然这些模型仍主要是理论性的, 但是可以提供某些适用性的见解, 对政策产生有价值的影响。经典的谢林隔离 (Schelling segregation) 模型^[7,8] (详见第 5 章论述) 虽然相当抽象, 但却揭示了关于社会隔离模式的重要见解, 并为决策人提供了帮助。

3) 适度现实: 模型虽然属于定性范畴, 但是在重要特征方面均符合定量特征。以实验为基础的社会计算研究对这类模型最感兴趣。

4) 高度现实: 实验输出和经验数据之间的定量和定性特征最为符合。高度现实仿真可以从多个维度与参照系统进行比照, 包括立体性 (如众多详细的地理特征, 通过 GIS 和遥感数据实现)、时长 (定义为小时间增量, 如十年、年、季节、月、周、日、小时、分钟等), 或组织模式 (节点、子图和图表在水平分析时所匹配的网络) 等^[9]。此类模型在商业和政府组织中应用极为丰富。

复杂系统的计算模型^[10-15]通常由跨学科团队创建 (团队成员精通不同的专业), 能够实

现学科知识 (cyber-physical-social) 的耦合。这类模型通常还需要多年的开发, 涉及多个研究机构。复杂系统的计算实验方法一般将模型开发视为螺旋式的多阶段过程, 从初级简单模型开始^[16], 不断向系统中加入不同影响要素, 从而向复杂的最终模型推进。物理学历史上类似的著名案例是牛顿的行星动力学模型^[17]的构建过程。牛顿并非是一开始就构建出一个庞大的完整模型, 而是通过一系列渐进模型实现自己的最终目标。牛顿最开始的模型只是绕轴自转的一个球体, 后续逐渐添加了卫星、旋转斜轴、椭圆轨道和许多其他精心挑选的功能模块, 最终形成了完整的行星系统模型。

在对复杂系统进行渐进式建模的过程中, 有几点需要注意: ①有必要确定一个简单的初始模型, 足以清楚模型的全部细节, 同时代表了参照系统最终模型的核心要素。这个阶段的模型与实际系统可能并没有太多相似之处。②设计模型的序列并不是随意的, 必须为完成最终模型提供递进式的帮助。③在整个开发过程中, 模型验证是必要的, 但需要适度。因为中间阶段的模型尚不成熟, 可能由于缺乏足够的实证支持而被否决。④确定参照系统的最终模型是必要的, 否则模型的改进会无限期延续下去。因此, 明确核心研究问题的重点对于管理复杂系统的建模是必不可少的。

3.3 系统开发

基于 3.2 节所构建的计算模型, 将所研究的复杂系统抽象为一个多 Agent 系统^[18], 基本步骤包括: 划分系统边界, 分离系统和环境; 识别系统中的主体, 赋予主体属性; 建立各类主体的微观实现; 建立主体之间以及主体和环境之间的相互作用机制。在对复杂系统进行观察时, 应同时收集微观数据和宏观数据。微观数据和关于个体的先验知识是提出假设, 建立微观实现的基础。宏观数据反映系统的宏观特性, 用来推断实验系统和真实系统的微观-宏观联系是否一致。

如图 3-7 所示, 实验系统包括数据库模块、参数配置模块、实验执行模块、仿真显示模块和结果评估模块五部分。若不使用干预策略, 系统用于模拟没有干预情况下的自然演化过程; 若使用干预策略, 系统用于模拟在干预情况下的可控演化过程, 结论可以为干预策略的制定和修订提供信息依据。下面对各部分进行具体的说明。

(1) 数据库模块

1) 模型库。模型库用于存储实验系统运行所涉及的各种算法模型, 包括市场需求模型、企业个体模型、企业交互模型、服务策略模型 (用于实现产业集群的可持续发展问题求解, 例如企业的均衡化发展程度、企业的平均利润、企业的数量限制等)、多目标资源优化调度模型 (用于实现多阶段资源优化调度的方程表达式)、决策分析模型 (多属性群决策分析模型、模糊决策分析模型) 等。模型库的模型可以重复使用, 可以有针对性地筛选使用, 也可以根据实际情况予以扩展。

2) 知识库。知识库存储、组织、管理和使用互相联系的知识集合。知识集合包括与企业运营有关的定义、定理和运算法则以及常识性知识。知识库基于预案数据和案例数据, 制定知识规则, 产生行为方案; 或者根据统计数据, 生成知识推理产生式, 设置 Agent 参数。Agent 是知识的表现和实现形式, 知识库包括 Agent 原型建模规则、Agent 行为规则、Agent 学习规则、环境更新规则、反应方案制定规则以及知识推理产生式知识。

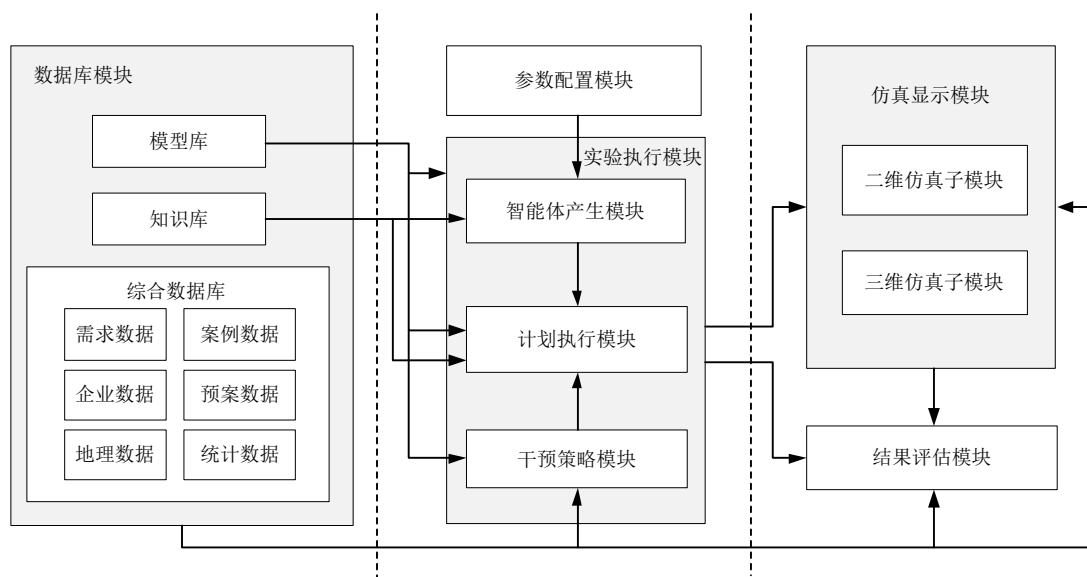


图 3-7 实验系统的结构

3) 综合数据库。在综合数据库中，需求数据、企业数据、地理数据为系统提供基础数据支持；统计数据、预案数据、案例数据为知识库和干预策略模块提供数据支持，为制定合理的干预策略提供依据和指导。具体的细节如下。

① 需求数据：包括市场需求的类型、订单的数量、随时间的波动情况、对产品的时间和价格限制、相关供应商等信息。

② 企业数据：包括企业种类、数量、分布位置、企业信誉、资金实力、最大产能、产品质量、产品价格、经营特点等信息。

③ 地理数据：包括产业集群的空间分布、地理状况以及环境数据。环境数据包括区域资源、气候条件自然环境，以及基础设施、商业传统、地方政策等社会环境信息。

④ 案例数据：包括典型案例信息以及案例过程中所采取的干预策略信息。典型案例的外部环境、实验对象、演化过程以及干预策略都存入关系型数据库中。

⑤ 预案数据：将设计好的干预策略以知识库的形式存入数据库。预案是指根据预测，对潜在的或可能发生的演化过程和结果进行事先制定的干预策略手段。

⑥ 统计数据：实验对象在前期演化过程中的各种业务统计信息，例如企业分布变化、企业数量变化、区域经济成分变化、协作联盟比例变化等。

(2) 参数配置模块

参数配置模块接受用户设置参数，将设置参数提交给 Agent 产生模块和仿真模块。这里描述的设置参数由 Agent 参数和仿真环境参数构成，Agent 参数包括各类企业 Agent 参数；仿真环境参数包括外部事件参数和区域环境参数，实现企业 Agent 实体以及环境参数的配置，例如产业集群的区域特征、企业 Agent 的数量、特征、位置分布等。通过参数配置模块，能够在虚拟环境中定制不同场景下演化事件、企业实体及环境参数的参数值。

(3) 实验执行模块

实验执行模块主要分为三个部分：智能体产生模块、计划执行模块和干预策略模块。

通过计算实验的执行，可以实现两方面的功能：一是在没有干预策略的情况下，根据实验的初始条件设置，实现产业集群的自然演化，主要用于识别不同的初始因素对于演化结果的影响程度；二是在有干预策略的情况下，根据实验的初始条件和干预策略设置，实现产业集群的可控演化，主要用于识别不同干预策略对于演化结果的影响。

(4) 仿真显示模块

仿真显示模块包括二维仿真子模块和三维仿真子模块。根据设置参数，在综合数据库的基础上建立符合真实情况的场景环境模型，将场景中的环境信息构造映射到仿真模块，为计算实验提供仿真环境。仿真环境将产业集群的产生、演化、发展和结束通过二维场景、三维场景、图像或者文字等多种媒体形式在计算机中直观展现。仿真模块作为展现仿真环境的平台，要实时动态更新企业 Agent 的状态和环境的变化，将仿真结果直观地反映给用户。

(5) 结果评估模块

结果评估模块在计划执行、仿真模块运行数据的基础上，利用模型库中的决策分析模型统计各类参数对演化过程产生的影响，例如统计信息（企业数量、企业利润等）、初始参数值（企业分布、组织模型等）、干预策略（服务收费、服务组合等）。所采取的分析方法包括多属性群决策分析模型（对演化结果建立评价向量矩阵，通过计算各属性的理想值得出综合评价建议）、模糊决策分析模型（根据模糊决策分析原理求解定性定量相结合的决策问题）等。根据分析结论，为制定最优干预策略提供依据。实验执行模块和仿真显示模块为结果评估模块提供信息，实验评估结论为知识库提供知识储备。

3.4 实验执行

复杂系统的演化过程是内部因素和外部因素共同作用的结果。外部环境因素是客观的、不可控的，主要包括初始条件以及外部环境。内部因素是可控的、可调整的，主要包括个体之间的组织形式、协作策略、协调机制等。如果没有外部因素的干预，实验系统将会进行自然演化，可以用来分析初始环境在系统演化中的作用。如果对实验系统进行外部诱导干预，则可以用来分析外界措施对系统演化的影响。通过对仿真系统设计不同的“实验”方案，用仿真实验的方式探索复杂系统的演化动态，各种事件的发生是在仿真过程中自发涌现出来的，并在必要时进行多次大量的重复模拟，完成对各种系统按不同指标体系进行全面、准确、及时、量化的分析与评估。

根据知识库中的知识规则和模型库中的模型，计划执行模块计算并执行智能体产生模块产生的各类 Agent 在仿真环境中的行为，同时动态更新 Agent 的状态属性及仿真环境，将计算结果输出到仿真模块。如图 3-8 所示，计算实验的执行过程包括预处理、计划执行和输出三个步骤。

(1) 预处理

首先，收集计算实验必需的数据，数据来源于数据库模块所提供的基础数据，以及参数配置模块所提供的 Agent 环境参数值。然后，在数据搜集工作完成的基础上，构建定制的 Agent 实体和虚拟仿真环境。创建 Agent 结构由智能体产生模块完成，创建虚拟环境由仿真模块完成。该步骤主要是核查 Agent 和虚拟环境的创建工作是否完成，在创建工作完成的基础上初始化各 Agent 的实体目标以及简单行为。

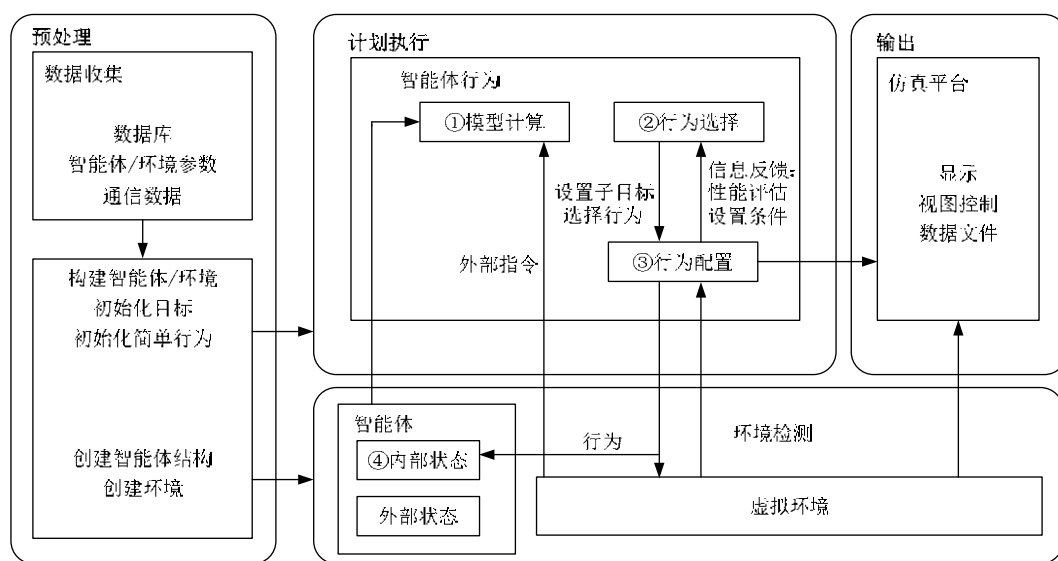


图 3-8 计算实验的执行过程

(2) 计划执行

实验系统分别计算出各个 Agent 的行为，以及 Agent 和虚拟环境的相互作用，经过合成配置后执行其行为模式。整个过程包括以下几个步骤：①模型计算，Agent 在预处理阶段收集的内部状态值、虚拟环境提供的外部状态值和外部指令的数据基础上，选择模型库中的模型和知识库中的知识进行模型计算，计算结果进入行为选择器。②行为选择，利用知识库中 Agent 行为规则进行行为计算并选择，以及设置下一时刻的子目标。③行为配置，首先，在行为选择结果和外部环境检测数据的基础上，利用知识库中 Agent 行为经验学习规则进行 Agent 行为的复合配置；其次，将行为配置计算信息进行信息反馈给行为选择步骤，对实体的行为选择参数进行修正，通过性能评估或设置条件的方式实现。④更新 Agent 的内部状态值，以及将行为状态信息传递给仿真模块中构建的虚拟环境，利用知识库中环境更新规则实现虚拟环境的动态更新，并进入下一个时刻的循环计算，直至各 Agent 达到目标。

(3) 输出

当 Agent 行为执行完成后或虚拟环境更新完成后，就形成信息输出，实时更新虚拟环境。执行结果输出到仿真模块中的仿真平台上，以虚拟仿真环境显示，视图控制以及数据文件的方式展现。通过多次有目的的实验，可以对系统的特征进行刻画，实验过程中所产生的数据可作为决策者分析系统行为的定量依据。基于计算模型构建出来的实验系统，可以作为“培育”各种复杂现象的手段。

3.5 结果分析

计算实验系统是数据发生器，能够产生海量的复杂系统模拟运行数据。计算实验数据分析根据已有的数据，进行科学分析，发现系统运行中的问题和薄弱环节，并找出原因，提出符合实际的解决问题的办法或建议。基于构建出来的实验系统，可以对感兴趣的实际

现象进行实验研究。这些新现象的出现是通过 Agent 之间的局部交互涌现出的系统全局行为，通过与实际现象的比对，可以得出微观行为与宏观现象之间的关联关系。计算实验数据的分析仍然采用通常的大数据分析手段，包括关系网络、聚类、分类、事件关联、异常检测等。二者的不同之处体现在：计算实验能够模拟整个系统在不同初始条件和不同场景下的演化过程，可以分析最后的数据结果是如何产生的；而单纯的大数据分析只能给出结果，却无法说明微观行为是如何与宏观现象进行关联的。

为了有效地分析复杂系统的演化过程，可以从“宏观机理”“中观结构”“微观行为”三个层次展开分析，具体方案如图 3-9 所示。

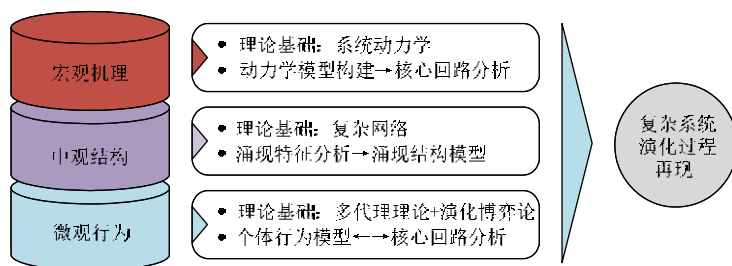


图 3-9 “三层次”复杂系统计算实验结果分析

1) 宏观机理分析：复杂系统当中存在明显的反馈回路，它是一种非线性时变的复杂社会经济系统。由于系统动力学理论是一种处理复杂社会经济系统的有效工具，这里将以系统动力学理论为基础，分析整个系统演化过程的内在机制（如个体热度变化过程、个体退出机制等）和反馈回路（如干预机制影响个体行为等），构建系统动力学模型。基于数据分析研究模型参数的学习方法，分析不同要素对整个系统演化的影响，进而对整个系统的演化过程进行预测。

2) 中观结构分析：复杂系统当中的个体在长期的演化过程中会涌现出复杂的结构特性，包括幂律特性、小世界特性和异配特性等。为了理解这些涌现特性的产生机理，这里将以复杂网络理论为基础，分析个体根据需求选择合作伙伴的过程数据，构建竞争与合作过程的概率模型，从而得出系统复杂结构的涌现机理，在此基础上研究整个系统复杂结构的演化过程。

3) 微观行为分析：复杂系统当中的个体将能够根据其所在小生境的状态调整其行为，如进入系统、更新功能、调整个体行为、退出系统等行为，从而使整个系统具有适应性和自组织性。因此，可以基于多代理理论^[18]和演化博弈论^[19]，从个体微观行为的角度入手，对个体自适应行为的状态转移轨迹^[20]进行分析，进而判断整个系统的自适应行为和系统鲁棒性。

针对每个层面，计算实验可以采取多种分析形式，包括过程分析、假设分析以及情境分析等。过程分析主要关注个体在时间和空间上的相互作用，尤其是关键阈值与关键属性。假设分析可以对使用不同规则集的模型进行比较，如摩尔（E.F.Moore）邻域与标准冯·诺依曼（J.von Neumann）邻域的差异。情境分析可以为实验模拟分析提供更全面的可用方法。例如，研究传染病在特定地区的传播情景，可以分析环境因素对于疾病传播的影响，哪种干预措施能够有效遏制疾病传播，干预措施在什么时间点介入效果最佳等。

3.6 模型评估

分析计算模型是否通过概念模型发挥预期作用的过程被称为评估。这里需要根据模型计算出的结果,结合其具体的应用领域,评估该模型是否适合解决实际问题。如果能够解决,选择最优的一套模型或者多套模型,为实际问题服务;如果都没能解决实际问题,则需要对模型进行调整或者重构。目前,计算实验研究中通用的质量评估标准尚未出现,这里给出一些被广泛接受的原则,以供参考。

(1) 有效性

分析计算模型结果是否与已知系统规律相匹配的过程被称为有效性验证。从本质上讲,有效性验证就是计算实验输出和已知系统模式之间的匹配性问题。未经有效性验证的模型不能用于实际问题分析,所得出的实验结果也无法令人信服。目前,存在各种各样可用于计算模型有效性验证的方法。其中,最重要和常见的有结构有效性与行为有效性。

结构有效性是指模型的内部功能,包括 Agent 相关属性、交互规则和环境的主要假设是否有效。测试计算实验模型中的结构效度时,应考虑以下因素:①目标方法中所使用方程式的规范以及属性是否有效。如果采用结构稳定性的假设,则整个模拟过程中系数是不变的,即场景中 Agent、规则和环境间的规则设置不会随时间变化。②采用不同学科的知识来检查计算实验模拟的假设,这是比结构效度检验更为广泛的角度。

行为有效性与计算实验模型运行取得的实际结果有关,例如成长模式、衰败或振荡等定性和定量特征。确定行为有效性的关键在于计算实验所产生的模拟空间模式是否与参照系统内的已知经验模式相对应。例如,糖域模型^[21]产生的洛伦兹曲线和财富分布直方图。这些模型的实验运行结果与很多已知的社会经验规律非常匹配(帕累托法则)。

通常情况下,验证计算实验模型是利用已有的经验数据进行推论。但有的时候,计算实验的结果可能并未在现实世界中产生过。在这种情况下,需要充分利用领域知识,探寻产生这种现象的原因,是否是现实世界中一种可能的发展趋势,以此判断该模型的有效性。

(2) 简洁性

简洁性用于衡量模型的表示形式是否优雅,这会体现在规则式、句法结构以及相似特征等方面。例如,有些物理方程式就非常符合这一标准,逻辑清晰、明确且体系完整。与有效性不同,简洁性不一定是必备属性,但肯定是一个理想的科学属性。

(3) 实用性

实用性是指分析计算实验场景是否能够对改造现实世界发挥作用。举例来说,计算实验模型可以帮助研究人员处理很多棘手问题,并有助于改进现行政策,例如冲突解决、城市发展、资源均衡化、新技术冲击等。然而,社会仿真模型是通过代码实例化呈现的,对于评估社会领域的模型质量的检验和确认往往是不充分的,所以需要常见的必要实践。虽然 20 世纪 70 年代初所使用的众多模型假设被证明是不正确的,但方法本身是强大的且持续发展至今。另外,在有些情况下,无法针对相关参照系统开展实验,例如理解两个星系碰撞会发生什么的唯一途径是进行计算实验。

3.7 结果应用

计算实验中的仿真实验建立在真实世界的基础上，是真实世界的系统映射。通过对仿真结论的综合分析，可以对一些在现实系统中难以评估的指标进行准确可靠的量化分析，帮助理解宏观现象与各种影响要素之间的联系，进而给出系统演化方向的预测判断，给决策者提供辅助决策。计算实验的目的是得出有利于真实世界发展或解决真实世界中问题的方案和结论。通过对多种可能性的实验环境进行预演，可以检验措施是否有效，对策略进行分析和效果评估，并不断进行优化。

根据分析结果可以对不同的策略进行比较，从而为系统演化干预提供决策支持。实际系统可以根据实验系统中的结果动态调整各种机制和策略，完成对实际系统的管理和控制；实验系统可以根据实际系统的反馈数据重新设计相关实验，并根据评估结果完成对实验系统的优化和改进，进一步与实际系统相吻合。通过实际系统与实验系统的相互反馈，对二者之间的行为进行对比和分析，完成对系统的调整与优化，真正达到有效管理的目的。

对于复杂系统而言，干预策略需要对知识库中的预案或范例进行调整、整合、重造，形成新的干预策略。策略的生成一般有两种方法：①基于规则生成，即根据规则调整或整合合适的预案，生成合适的处置方案；②基于范例生成，即调整或整合合适的案例。两种手段可同时使用，最后需要对结果进行综合处理。

事件风险预测与策略生成的整体结构如图 3-10 所示。事件解析器基于系统的本体知识，对事件进行分类、分级，根据语义抽取对现场事件属性的信息，形成 XML 的事件描述。干预策略所涉及的知识有：规则型知识（临机处置规则、预案修正规则等）、本体知识（推理术语、领域术语及其他语义相关的术语）和范例知识（事件处置的历史范例）。图 3-10

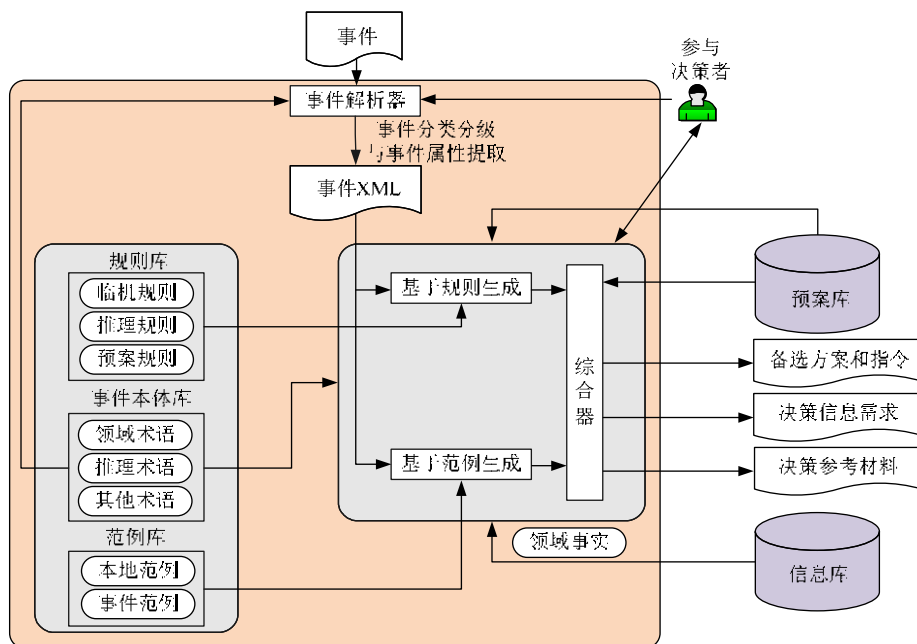


图 3-10 风险预测与决策生成整体结构

中综合器的功能是恰当地选出处置预案或范例的若干片段，对其工作流程片段或某些任务节点进行调整，最后对两个或多个事故处置的片段进行整合。人（决策者和顾问）在其中扮演重要的角色，他们是理解、分析和制定工作流程的最后拍板者，而决策的信息系统是知识与信息的提供者和分析者，基于相关的本体集来“重造知识”。

与预案集、范例集和规则集相关联的本体，对指导风险预测和决策生成有至关重要的价值。通过对这一本体进行归纳，抽取出临机决策中表达知识的基本框架，统一对临机决策知识的描述，从而最终作为异构本体间交流、共享的基础。这里将这种知识框架（schema）称为面向事故的决策知识服务。决策知识服务的形成需要研究三个方面的技术：对本地知识（预案、范例和规则）的语义要素抽取、语义标注，以及所涉及词汇的语义关联。

这里将决策生成按照三种形式进行处理：再生型是指有合适的应对预案，但预案中的资源、动作等由于某种原因无法工作，而找其替代品以达到原有效果的决策方式；调节型是指没有合适的预案，但某几个预案片段调整、整合后能应对事件的决策方式；创新型是指没有合适的预案能应对事件，决策者根据自身的知识制定应对事件的决策方式。

1) 再生型决策的关键在于找出用来生成处置方案的范本预案中的要素（包括组织结构、资源、行为和任务等），并在分析当前意外事件所处现场状况的基础上对要素内容进行调换，形成实际可用的处置流程。因此，需要从现有的本地预案集中提取出这些要素，并和上层的本体知识进行关联，从而支持决策者对相关预案中要素内容的快捷浏览，提高决策者对要素的即时理解程度和要素调换时的准确程度。

2) 调节型临机决策需要依据当前的意外事件情况，寻找紧密相关的一组预案集和范例集，并抽取出符合实际需求的预案或方案片段，重新整合成可用的处置流程。因此，需要分析进行预案或范例重组时的关键词汇，使用这些词汇对预案集和范例集进行语义标注，并和上层的本体知识进行关联。决策者在进行方案重组时，可以通过快速定位和浏览预案或方案中完成目标功能的候选片段，并根据实际情况进行斟酌、选取。

无论是再生型决策中的要素替换，还是调节型决策中的方案重组，都可能在既往的决策实践中总结出有效的规则，这里把这些规则称为决策规则。决策规则往往使用了上层本体中的关键词汇。因此，需要将规则和相关的本体通过这些词汇形成关联，从而有效地提示决策者选择当前合适的决策行为，提高决策者的决策能力。

3.8 小 结

本章详细阐述了计算实验方法包含的六个步骤：①方法的适用性判定，关键在于提醒使用者，计算实验方法并非适用于所有领域；②计算实验系统的开发，包括实验初始设置以及个体的行为规则，通过改变内外部因素的组合方式，可以建立各种各样的实验场景，从而对不同因素在演化过程中的影响进行实验；③计算实验的执行，利用实验可设计和可重复的特点，对系统演化进行各种压力实验以及极限实验，从而解决此类实验无法在现实中去实现的缺陷；④计算实验结果的分析，包括对涌现现象进行观察和解释，以及按不同指标体系对各种演化模式进行全面、准确、及时、量化的分析；⑤计算模型的评估，主要是通过和历史数据的比对判断模型的有效性；⑥实验结果的应用，侧重于将计算实验的运行结果用于理论分析或者决策支持，体现其实际价值。本章能够使读者对计算实验的过程

有所了解，为后续章节的学习奠定基础。

参 考 文 献

- [1] 王飞跃. 计算实验方法与复杂系统行为分析和决策评估[J]. 系统仿真学报, 2004(5): 893-897.
- [2] 王飞跃. 关于复杂系统研究的计算理论与方法[J]. 中国基础科学, 2004(5): 5-12.
- [3] 崔凯楠, 郑晓龙, 文丁, 等. 计算实验研究方法及应用[J]. 自动化学报, 2013, 39(8): 1157-1169.
- [4] ESPTEIN J M, ASTELL R. Growing artificial societies: social science from the bottom-up[M]. Cambridge: MIT Press, 1996.
- [5] WOOLDRIDGE M, JENNINGS N R. Intelligent agents: theory and practice[J]. The Knowledge Engineering Review, 1995(10): 115-152.
- [6] CRAIG R. Boids background and update[EB/OL]. [1995-6-29]. <http://www.red3d.com/cwr/boids/>.
- [7] SCHELLING T C. Dynamic models of segregation[J]. Journal of Mathematical Sociology, 1971, 1(2): 143-186.
- [8] LAURIE A J, JAGGI N K. Role of vision' in neighbourhood racial segregation: a variant of the Schelling segregation model[J]. Urban Studies, 2003, 40(40): 2687-2704.
- [9] XUE X, CHANG J, LIU Z. Context-aware intelligent service system for coal mine industry[J]. Computers in Industry, 2014(65): 291-305.
- [10] 缪青海. 基于 Artience 的人工交通系统研究与设计[D]. 北京: 中国科学院自动化研究所, 2007.
- [11] 赵红霞. 人工人口与交通需求关系计算实验研究[D]. 北京: 中国科学院自动化研究所, 2009.
- [12] 吕宜生, 欧彦, 汤淑明, 等. 基于人工交通系统的路网交通运行状况评估的计算实验[J]. 吉林大学学报(工学版), 2009, 39(S2): 87-90.
- [13] 汤淑明. 人工交通系统的基本方法研究[D]. 北京: 中国科学院自动化研究所, 2005.
- [14] ZHU F, LI G, LI Z, et al. A case study of evaluating traffic signal control systems using computational experiments[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4): 1220-1226.
- [15] 朱凤华. 基于人工交通系统的城市交通信号控制系统评价研究[D]. 北京: 中国科学院自动化研究所, 2008.
- [16] Sample models[CP/OL]. <https://repast.github.io/>.
- [17] 周起钊. 多体系统的动力学[J]. 力学进展, 1982(3): 287-299.
- [18] 蒋云良, 徐从富. 智能 Agent 与多 Agent 系统的研究[J]. 计算机应用研究, 2003(4): 33-36.
- [19] 王文宾. 演化博弈论研究的现状与展望[J]. 统计与决策, 2009(3): 158-161.
- [20] ZHANG J, LI X. Explore complex adaptive systems with digital artificial life model[J]. Complication System and Complexity Science, 2005, 2 (1): 39-48.
- [21] 张江, 李学伟. 人工社会: 基于 Agent 的社会学仿真[J]. 系统工程, 2005, 23(1): 13-20.

第 2 部分

计算模型篇

第4章 基于数学变量的模型

4.1 引言

计算实验是一个交叉学科领域，由一系列的概念、原则、理论和研究方法构成。对社会经济系统的复杂性进行模拟研究时，这些领域知识互相依赖，并且协同工作^[1]。本章主要介绍基于数学方程的计算模型。从历史的观点上讲，基于数学变量的模型是计算实验方法的先驱，主要包括系统动力学（system dynamics, SD）模型、博弈模型、排队模型和马尔可夫模型^[2]。这类模型均聚焦于系统随时间变化的特征，非常适用于解决管理、工业和操作设置中的策略问题，通过理论分析，结合实际经验，使各项应用更具成效。由于篇幅所限，这里只对每种模型的代表性技术进行介绍。

4.2 系统动力学

4.2.1 系统动力学概念

系统动力学用于研究和管理复杂反馈系统（例如，业务系统和其他社会系统）的计算机仿真模型。它以系统论和事件概率理论为基础，吸收了控制论、信息论的精髓，是一门综合了自然科学和社会科学的学科。系统动力学采用自上而下的结构，运用“凡系统必有结构，系统结构决定系统功能”的系统科学思想，根据系统内部组成要素互为因果的反馈特点，从系统的内部结构来寻找问题发生的根源，而不是用外部的干扰或随机事件来说明系统的行为性质。系统动力学模型是一种系统动力学模拟，属于以变量为基础的计算模型，用于分析含有变量和变更率、反馈和前馈相关性的复杂高维系统。正规的 SD 模型包括一个离散时间差分（前差分或后差分）方程系统，可以为确定性模型，也可以含有随机变量定义的随机噪声。表 4-1 展示了 SD 模型发展过程及代表性成果。

表 4-1 SD 模型发展过程及代表性成果

时间	人物	代表性成果
1909 年	丹麦数学家和工程师 A.K.Erlang	最先通过建立哥本哈根电话局模型，对排队系统进行科学研究
1953 年	英国统计学家和数学家 D.G.Kendall	在《数理统计年报》上提出了排队系统中使用的标准标志法 ^[3] ，目前仍然在使用
1958 年	麻省理工学院 R.Bennett	创建了第一套系统动力学的计算机建模语言，即 SIMPLE（用大量方程式来模拟工业管理问题）
1959 年	缅因大学地质科学系 A.Pugh	编写 DYNAMO(动力学模型)1.0 版本，是 SIMPLE 的改进版本。DYNAMO 快速成为管理科学和操作模型的正规通用模型 ^[4]

科学出版社
www.abook.cn

续表

时间	人物	代表性成果
1961年	麻省理工学院斯隆管理学院工程科学家 J.Forrester	出版了开创性的书籍《工业动力学》 ^[5] ，这本书是 SD 的首本经典书籍
	美国国家工程院院士 T.L.Saaty	出版了经典排队理论著作《排队论的要素及应用》，应用了排队理论元素 ^[6]
	麻省理工学院斯隆管理学院终身教授 J.D.C. Little	在《运筹研究期刊》上发表了著名的排队系统定律 ^[7]
	英国数学科学教授 J.F.C. Kingman	在剑桥哲学学会《数学学报》中发表了同样著名的定律 ^[8]
1969年	J.Forrester 和 J.Collins（前波士顿市长）	提出了城市动力学，将系统动力学模拟在真正意义上拓展到了社会复杂性和复杂社会仿真中 ^[9]
1972年	美国环境科学家 D.Meadows	出版《增长的极限》 ^[10] ，这是一本让世界各地关注到 SD 的经典书籍。工程师奥雷利奥·佩西（A.Peccei）的罗马俱乐部为其提供赞助。《增长的极限》出版后，立即被转译成多种语言版本
	美国加利福尼亚大学的文化人类学家 L.S.Cordell	在其博士论文“Settlement Pattern Changes at Wetherill Mesa, Colorado”中首次对 Puebloan（Anasazi）政策进行了社会模拟。Cordell 获得了美国考古学会颁发的终身成就奖和美国人类学会颁发的阿尔弗雷德基德奖，并于 2005 年成为美国国家科学院院士 ^[11]
1975年	苏黎世大学政治科学研究所所长 D.Ruloff	首次应用 SD 模拟政治叛乱和稳定的过程，随后他建立了阿富汗政治瓦解和塔利班叛乱分子的首个动力学模型 ^[12]
	国际研究中心（CIS）高级教授 N.Choucri 和 R.North	发表了《冲突中的国家》，是首个国际关系方面的离散时间模拟，建立了一战发生的模型 ^[13]
1979年	瑞士日内瓦高等国际关系研究生院，政治科学家 U.Luterbacher 和 P.Allan	创建了 SIMPEST，针对冷战期间美国-苏联-中国的战略博弈构建了首个动力学模型，该模型准确预测到了苏联的解体。U.Luterbacher 和 P.Allan 在苏联莫斯科的国际政治科学协会的国际会议上介绍了其文章
	英国考古学家 C.Renfrew 和 K.L. Cooke	开用数学方法模拟文化变迁之先河
1981年	美国考古学家 J.Sabloff	首次实现了考古方面的模拟 ^[14]
	N.Choucri	出版了世界能源市场经济和政策前景方面的首部 SD 建模书籍《国际能源的未来》 ^[15]
1984年	SD 科学期刊	创立《系统动力学评论》
20世纪80年代中期	美国政治科学家 M.Wallace	发表了一篇文章，利用 DYNAMO 演示了在 SD 模型中应用 L.F.Richardson 军备竞赛理论的情况
1985年	ISEE 系统公司	发布了系统动力学建模软件 Stella 1.0
1998年	N.Choucri 和其麻省理工学院的 学生	在《系统动力学评论》中发布了首个状态稳定的 SD 模型 ^[16]
2000年	美国管理科学家 J.D.Sterman	出版了《企业动力学》《复杂世界的系统思考和建模》，是 SD 方面的经典大型教科书 ^[17,18]

SD 模型在处理众多领域的复杂性研究问题，特别是在指定参照系统中出现下列特性。

1) 适用于处理长期性和周期性的问题。例如，自然界的生态平衡、人的生命周期和社会经济危机等都呈现周期性规律，并需通过较长的历史阶段来观察，已有不少系统动力学模型对其机制作出了较为科学的解释。

2) 适用于对数据不足的问题进行研究。建模中常常遇到数据不足或某些数据难以量化的问题，系统动力学依靠各要素间的因果关系、有限的的数据及一定的结构仍可进行推演分析。

3) 适用于处理精度要求不高的、复杂的社会经济问题。上述问题的描述方程通常是高

阶非线性动态方程，应用一般数学方法很难求解。系统动力学则借助计算机仿真技术仍能获得关键信息。

4) 强调有条件预测。本方法强调产生结果的条件，采取“如果……则……”的形式，为预测未来提供了新的手段。

5) 在宏观系统水平上，变更轨迹可以包括稳定、增长、抑制、循环、振动、渐近行为和其他时间定性模式。系统复杂性的突变性质是由最低因果层级的变量相互作用导致的。因果性网络中各个节点的噪声会影响最终结果的轨迹（噪声是后期分离出来的，作为概率分布）。

4.2.2 描述模型

系统动力学的建模过程是一个由粗到精、由浅入深地将思维模型转化成数学模型和计算模型的过程。在这个转化过程中，系统动力学有一整套将模型逐步量化的方法：方框图法、因果关系图法和库存流量图法等。这些方法各自有不同的特点和功能，依次使用这些方法，能够方便而有效地将定性模型转化为定量模型^[19]。

1. 方框图

方框表示系统元素、子系统或功能模块，方框中填上相应的名称、功能或说明。带箭头的实线表示各元素、各子块之间的相互作用关系、因果关系或逻辑关系，也可以表示流量的运动方向（流量写在实线旁）。以公司为例，其模型方框图如图 4-1 所示。

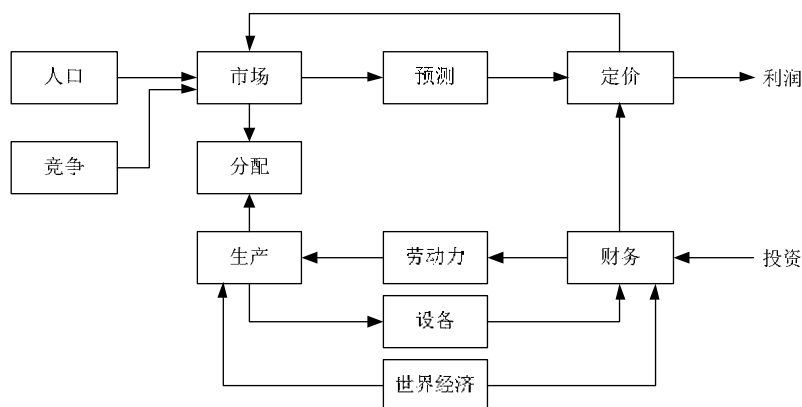


图 4-1 公司的模型方框图

2. 因果关系图

因果关系图在构思模型的初级阶段起着非常重要的作用。在建模过程中，它既可以初步明确系统中诸变量间的因果关系，又可以简化模型的表达，使人们能很快地了解系统模型的结构假设，使实际系统抽象化和概念化。在因果关系图中，用因果链表示各变量彼此之间的因果关系。以下是因果关系图的几个重要概念。

1) 因果链：因果链是一个带箭头的实线（直线或弧线），箭头方向表示因果关系的作用方向，箭头旁边有“+”或“-”号，分别表示积极的或消极的作用关系，具体取决于该

关系是促进还是阻碍指定变量的变化。

2) 正向反馈：一种积极的反馈循环，是增加变量值的一种因果关系。积极反馈属于强化动力学范畴，会产生增长、扩张、增加、放大、增加改进、增补、扩散上升或时间轨迹变量中的其他增长模式。

3) 负向反馈：一种消极的反馈循环，是减少变量值的一种因果关系。消极反馈属于平衡动力学范畴，会产生疲劳、衰退、减少、损失、缩减、降低耗损、收缩、限制、衰减或时间轨迹变量中的其他降低模式。

4) 关系回路：多个因果链以同向封闭的形式连接起来组成因果关系回路。回路的极性取决于回路中负向因果链的数目。若回路中负向因果链的数目为偶数，则回路极性为正(+)；若回路中负向因果链的数目为奇数，则回路极性为负(-)。

因果关系图的一个典型用处就是描述某种社会现象，例如社会成员逐步遵循某种新的社会规范。如图 4-2 所示，从 SD 的角度而言，有助于理解社会成员采用新规范是如何作为一个社会过程运行的。在图 4-2 中的社会规范采用过程中，设定了两个反馈循环关联，且两个循环同步作用。积极反馈循环 R（右侧）表示的是社会从众性如何随着新采用者的数量增多而产生更多的新规范采用者。遵守新规范的人越多，采用该规范的动力就越大，这就是强化动力学。左侧的反馈循环 B 代表负增强或“平衡”，表示随着采用新规范的社会成员人数的增多，潜在的采用者就会减少。采用新规范的比例越高，还没有采用该规范的人数就越少，因此，左侧的循环代表消极反馈。

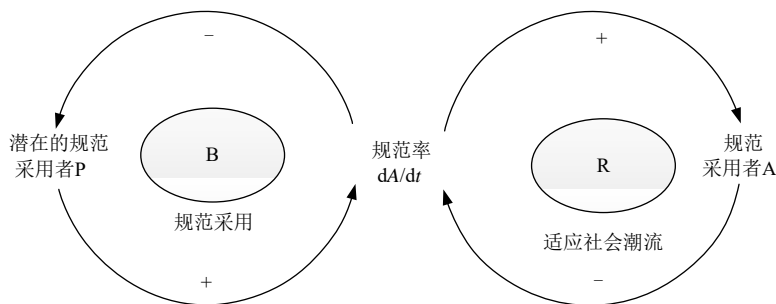




图 4-2 因果关系图表示的社会规范采用过程

总之，规范采用率是潜在规范采用者数量与新规范采用者数量的函数。潜在的规范采用者和实际的规范采用者会分别根据采用率降低和增长；降低和增长的结果是两个关联动力学在不同时间的不同行为。若社会成员中仍存在少量未采用规范者，则该过程早期采用者人群的增长率比该过程后期阶段的增长率大。社会规范的相关示例可用于表示改革、意见、技术革新、态度和行为等模式。因此，社会规范的采用过程可以广泛应用于社会科学领域。

3. 库存流量图

SD 模型开发过程中的第二个阶段利用库存流量图的符号体系描述系统中不同类型的变量以及各变量之间的相互作用关系。库存流量图采用的基本符号及含义如表 4-2 所示。

表 4-2 库存流量图采用的基本符号及含义

符号名称	符号标识	含义
状态变量		状态变量是表征系统状态的内部变量，可以表示系统中物质、人员等的稳定或增减状况
速度变量		速率变量控制着状态变量的变化，规定了这种控制的方式和强度
辅助变量		辅助变量是为了建模方便而人为引入的信息反馈变量，是状态信息变量的函数
外生变量		外生变量呈现固定不变的状态时就退化成常数
表函数		反映了两个变量之间某种特定的非线性关系
物质延迟		指发生在物质流上的延迟
平滑、信息延迟		指发生在信息流上的延迟
物质的源与汇		源抽象为白洞，物质只出不进，所以与它相连的物质流线箭头均朝外；汇表示永远填不满的坑，抽象为黑洞，物质只进不出，所以与它相连的物质流线箭头均向内
物质流		物质通过物质流线运动
信息流		信息通过信息流线传送

在库存流量图中，所有物质都取于源而聚于汇，但是出自不同源的物质只能聚入与其同质的汇中，这就是守恒子系统的概念。守恒子系统是指以一个状态变量为中心组成的局部系统。根据源和汇的特点可以得到以下两个结果。

1) 同质守恒子系统之间既可以用物质流线连接，表示正向的因果关系，也可以用信息流线连接，表示反馈的因果关系。例如，在图 4-3 所示的人口年龄链结构中，各守恒子系统均是同质的，状态变量都表示人口，所以它们既可以用物质流线直接连接，也可以用信息流线反馈连接。

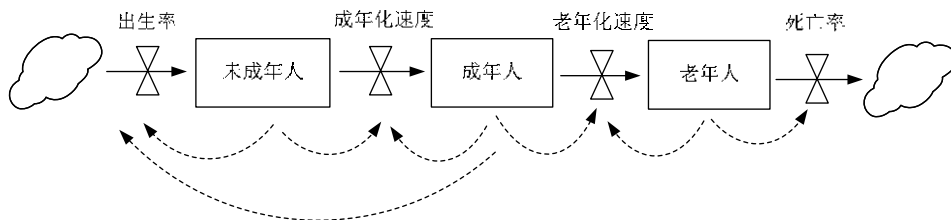


图 4-3 人口年龄链结构

2) 不同质的守恒子系统之间只能用信息流线连接，不能用物质流线连接。例如，图 4-4 所示的物价-收入子系统中，物价和收入分别属于不同质的守恒子系统，它们的量纲也不一样，所以它们之间不能用物质流线连接，只能以信息传送的形式相联系。

库存流量图的特点是将系统中各变量按其不同的特征，以及在系统中所起的不同作用划分成不同的种类，并用物质流线和信息流线按照其特有的作用方式将它们联结起来，组成系统的结构。所以，库存流量图法比因果关系图法更加详细地反映出系统内部的反馈作用机制，使人们对系统的构成有一个更加直观、更加透彻的理解。在系统动力学建模过程

中，库存流量图法是介于思维模型和计算机模型之间的一个十分重要的过渡办法，它为构造系统的数学模型打下良好的基础。对于一个比较熟练的建模者，可以省去用因果关系图法描述系统的步骤，但是通常不能省去库存流量图法的步骤。

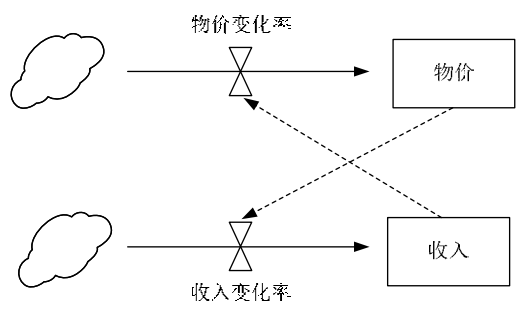


图 4-4 物价-收入子系统结构

4. 混合图法

在实际建模过程中还经常采用一种混合图法，将系统中物质流线上的状态变量和速度变量按库存流量图的方式画出，而将信息流线上的各种反馈变量按因果关系图的方式画出，如图 4-5 所示。混合图法汲取了因果关系图法和库存流量图法的优点，既保持了因果关系图简单明了的特点，又将系统中的重要变量鲜明地突出出来。因此，混合图法得到了比较广泛的应用。

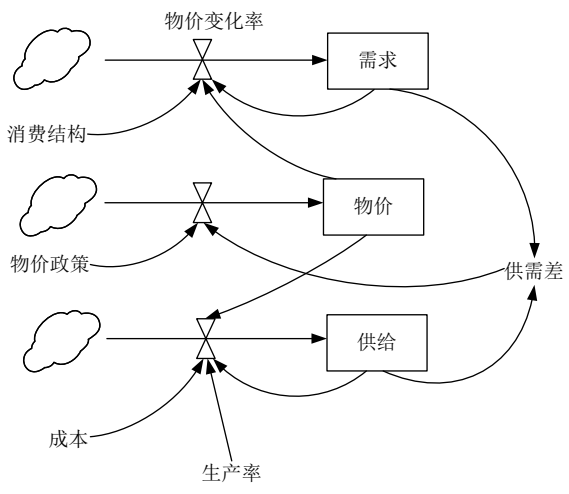


图 4-5 混合图法示例

4.2.3 模型示例

1. 示例描述

图 4-6 给出了基于 Richardson 的群体间军备竞赛动力学竞争模式，群体间冲突中也存在类似的示例。此种情况下，每个群体的武器采购率受两个相反的动力学影响。一方面，

科学出版社
www.abook.cn

由于其武器采购率受到竞争对手武器的水平驱动，因此该循环中存在增长动力学。增长动力学越高，通过增加己方武器采购率来赶上另一方的需求就越大。另一方面，该循环中存在减少动力学，一方受到其已经拥有的保持成本驱动，因此若另一方的军备水平越高，则己方经济负担越大，从而更加难以进一步增长。目前，直接经济成本中必须添加支持高级能力的组织复杂性。Richardson 将该约束力称为“疲劳”^[20]。

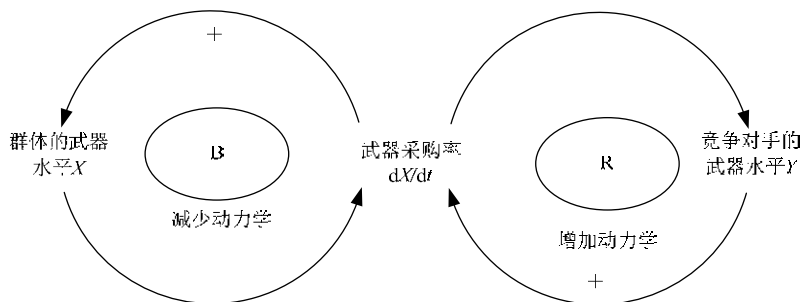


图 4-6 基于 Richardson 的群体间军备竞赛动力学竞争模式

一个整体系统由成对的因果循环表示，该循环表示某个基本因果循环如何与另一个循环关联。如图 4-6 中的竞争过程，假设两个反馈也循环关联，它们会同步作用。武器采购率是竞争对手武器水平和群体（自己）武器水平的一个函数，右侧的增加动力学是一个自行加强的驱动器（积极反馈），左侧减少动力学是一个平衡驱动器（消极反馈）。但是，与上一示例不同，本示例假设是作用于武器采购率的两种不同类关联。

1) 反馈联轴器：如规范出现的示例中所示，关联积极和消极反馈过程。

2) 作用物联轴器：由于每一方武器水平的结果是由己方的决策以及竞争对手的决策而确定的，因此可以用对策论的战略互动来联结两个竞争对手。此种情况下，竞争过程中这两个关联的动力学在不同时间的行为也不同，其取决于哪个动力学驱动占优势。

总体而言，因果关系图可从定性角度通过提取积极和消极反馈循环以辅助构建概念 SD 模型，该积极和消极反馈循环分别与加强/增加和抑制/减少驱动器一致。但是，构建一个足够完整的基准系统概念模型需求更多，该系统能够以计算机代码形式实施。

2. 库存流量图

图 4-7 给出了军备竞赛模型的库存流量图，展示了两个完全关联的冲突系统，说明竞争对手之间是如何作用和反馈循环的，是战略互动概念的形式化^[21]。竞争对手的动力学过程均由三个因素驱动。

1) 竞争对手的当前武器水平（积极反馈，增加力量）。

2) 群体本身的武器水平（消极反馈，减弱力量）。

3) 作为背景的敌对状态，其产生了这样一种想法，不管竞争对手的武器水平如何，为了以防万一，群体都需要获得某些最低的军事能力。

系统动力学设计阶段主要的成果，是设计一个由一组方程规定的基准社会系统概念模型。例如，在冲突模型中，下列连续时间方程系统[式 (4-1) 和式 (4-2)]指定竞争动力学。

$$\frac{dX}{dt} = aY - bX + g \quad (4-1)$$

$$\frac{dY}{dt} = \alpha X - \beta Y + h \quad (4-2)$$

其中, a 和 α 为反应系数; b 和 β 为缓解系数; g 和 h 为敌对状态系数; X 和 Y 为军备水平。以下为离散时间的方程系统。

$$X(t+1) = aY(t) - bX(t) + g \quad (4-3)$$

$$Y(t+1) = \alpha X(t) - \beta Y + h \quad (4-4)$$

此种情况下, 由于系统比较简单, 可对方程系统进行分析, 以获得闭环型解决方案。这些方程系统的解决方案能产生包含指数关系的时间轨迹, 该时间轨迹方便验证。大多数情况下, 这些系统不会产生时间轨迹, 因此需要进行模拟^[22]。

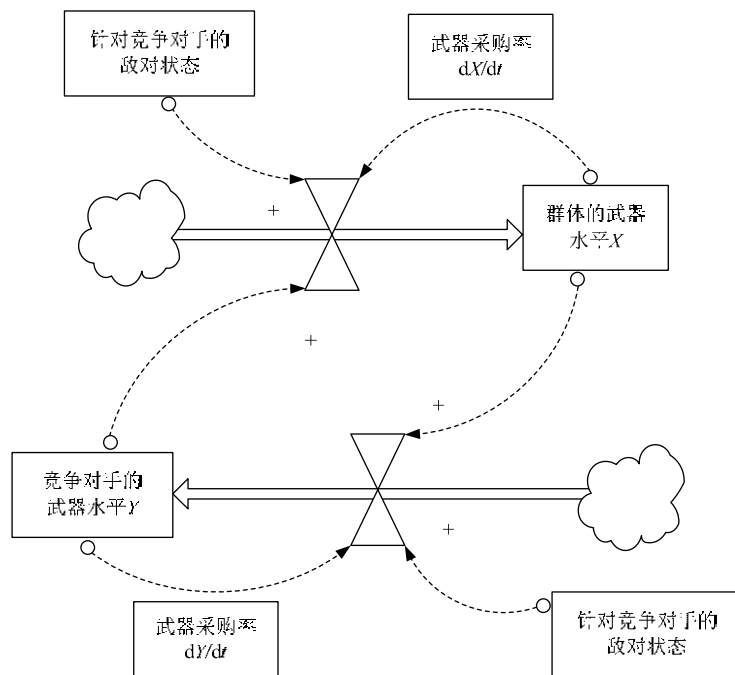


图 4-7 军备竞赛模型的库存流量图

4.2.4 分析

复杂系统模拟研究的主要目标是获得定性定量结果来更好地理解参照系统。显然, 开发一个 SD 系统模拟的主要目的是用来分析检查系统故障的原因。SD 系统模拟的主要形式包括问题假定分析、形式化分析, 以及情景化分析。

SD 模型的形式化分析会产生多种结果, 例如所有变量 (保有量) 的时间轨迹, 参数空间中的相图、灵敏度分析, 比较静态分析, 以及动力系统分析中的类似结果。例如, 形式化分析的冲突模型可以显示出群体军备水平在冲突演变中的时间轨迹, 作为一种参数组合的相图轨迹来表示相应的定性/定量结果。形式化分析的结果能够显示性能, 如轨迹、奇点、渐近线、吸引子、梯度场、周期性、混沌、分岔、各态遍历性 (时间平均值和空间平均值之间的平等)、相变、稳定性及其他在模型结构中显著的动态特征。

问题假定分析是另一种 SD 系统模拟的主要分析方法。例如, 在冲突模式中, 可能问

其中一方对其对手的敌意是另一方敌意的几倍时会发生什么，或者当两个群体间的反馈系数显著不同时会发生什么。采用问题假定分析时，可以用来探索当不同外力控制时，系统会发生什么变化。

SD 系统模拟中一个更全面的分析方法是情景分析，它涉及一系列相关的问题情景定义，而不是一次分析一个问题。例如，在冲突模型中，可能希望检验一个反馈系数相对较小，缓和系数数倍于反馈系数并且敌对系数很弱的情景。直观地讲，这样的一个小情景应该会引发一系列措施来降低冲突，如减小战争范围或强度、裁军等。相反的情景会存在一组位于相反范围的系数，引发螺旋状的失控扩大态势（甚至崩溃）。在这两种极端情景之间，还有许多其他定性或定量特性，仍然是未知的且依然在探索中。

SD 系统模拟的这些分析方法中，可以用于分析基本的系统分析以及政策评估问题。SD 系统也能够和其他模型进行组合，从而提升模型的应用范围。

4.3 博弈模型

博弈论用来研究这样一种情境，即最终的结果不仅取决于个体如何在不同的备选项之间进行选择，还取决于与他们互动的其他人所作出的选择。博弈论的思想出现在许多不同的情境之中。在一些情境中，博弈的意思非常明显，例如可以用博弈论的工具分析如何选择球员罚点球以及如何进行防守。另外一些情境则并不总被人们称为博弈，但也可以用同样的工具加以分析。这样的例子包括：当市场上已经有某一产品时，生产相似新产品的定价问题；在拍卖会上如何确定投标竞价；选择因特网或者交通网络中的一条路径；在国际关系中是选择一种比较强硬的立场还是比较温和的立场；在职业体育比赛中选择是否服用兴奋剂等。在这些例子中，每一个决策者的选择结果都取决于别人的决策。这就为博弈论的分析提供了一种策略要素。

正如本书在随后要讨论的，博弈论思想跟那些看起来似乎没有人作出明显决策的情境也存在一定的关联。也许进化生物学为此提供了最为显著的例子。进化生物学最根本的原则是变异，生物体为提高自身对环境的适应性会发生变异。但是，对这种适应性的评估常常不能单独进行。相反，它依赖于其他所有（非变异）生物体的行为，以及变异行为与非变异行为之间的相互作用。在这样的情境中，对变异成功或失败的分析便加入了博弈论的界定，并且非常接近于对人类决策的分析。类似的分析已经被应用于研究新文化习俗的成功或失败，它取决于所植入的已有行为模式。这就意味着博弈论的思想并不仅限于人们在互动中的行为推理建模，其内涵思想更加宽泛。博弈论一般会强调哪种行为得以推广时，对于维持自身利益更有利。

4.3.1 博弈的概念

作为研究的出发点，首先讨论博弈论背后的一些基本理念。博弈论关注的背景是决策者彼此之间是进行互动的，即行为具有相互连通性。每个参与者对结果的预期不仅取决于他们自身的决策，而且取决于与其互动的其他人所作出的选择。为了更好地理解博弈的定义，下面从案例分析开始。

1. 案例

假设你是一名大学生，在规定的截止日期前一天，你有两项需要准备的工作：一是考试，二是报告。此时，你需要在“为考试而复习”和“为报告而准备”两者之间做取舍。为使例子表达更加清晰，这里将利用一些假设。首先，假设你可以在“为考试而复习”或者“为报告而准备”间进行选择，但只能选择一种。其次，假设在不同决策结果公布之前，你对预期成绩有准确估计，如果进行复习，则预期成绩是 92 分；如果没有复习，则预期成绩是 80 分。

报告需考虑的因素相对复杂。因为报告是你和拍档的联合性互动行为。假设你和拍档都做了充分准备，报告十分完美，因而你们预期的成绩都是 100 分。假设只有一个人做了准备（拍档没有准备），则你们的预期成绩都是 92 分。假设两个人都不做准备，你们的预期成绩都是 84 分。

这个例子在推理时需注意，所有这些结果对你的拍档也是一样的。对于考试，假设他会有同样的预期结果。同样地，拍档也必须在复习考试或是准备报告之间作出抉择。

进一步假设，你们彼此之间不能相互沟通。所以，你们不能共同商讨行为选择。而且，在彼此进行独立决策时，相互都知道对方也在进行决策。假设你们都追求个人平均成绩的最大化，下面分析一下这种平均成绩是如何通过彼此之间投入的努力来决定的。

1) 假设你们都选择准备报告，则都将在报告上得 100 分，考试上得 80 分，每个人的平均成绩是 90 分。

2) 假设你们都选择复习考试内容，则都将在考试上得 92 分，报告上得 84 分，每个人的平均成绩是 88 分。

3) 假设一方复习考试，另一方准备报告，则得分结果如下所示：为报告准备一方在报告上得 92 分，考试上得 80 分，这方的平均成绩是 86 分；选择复习考试一方在报告上得 92 分，因为报告成绩是共有成绩，这方因对方的报告准备行为而获益，通过复习考试得 92 分，所以他获得的平均成绩是 92 分。

下面，通过 2×2 表格对这些得分结果情况进行总结。此处，行代表你的两种选择行为：是准备报告或是复习考试；列代表你拍档的两种选择行为。所以，2×2 表格中的每个单元格都表示你们的一种联合选择行为。在每个单元格中记录你们的平均成绩：左侧是你的成绩，右侧是你拍档的成绩。全部的记录结果，如图 4-8 所示。

		你的拍档	
		报告	考试
你	报告	90, 90	86, 92
	考试	92, 86	88, 88

图 4-8 决定是复习考试还是准备报告的博弈收益矩阵

2×2 表格很巧妙地表现了博弈论背景的设置。现在，需要决定行为选择：是准备报告，还是复习考试。很显然，各自的平均成绩不仅取决于个体在这两个备选项之间的选择，还取决于你拍档的选择，即互动方的选择。因此，作为各自决策的一部分，参与方必须对对方可能的行为进行合理推理。当考虑自己策略的后果时，必须想到他人决策的影响，这正