

诺贝尔经济学奖获奖者著作丛书

# 现代决策理论 的基石

〔美〕赫伯特·西蒙 著

北京经济学院出版社

C93  
88-64 87074

·诺贝尔经济学奖获奖者著作丛书·

# 现代决策理论的基石

## 有限理性说

〔美〕赫伯特·西蒙 著

杨砾 徐立 译



S0271252



北京经济学院出版社

1989年·北京

**现代决策理论的基石**  
yian dai Juece Lilun De Jishi

〔美〕赫伯特·西蒙 著  
杨砾 徐立 译

北京经济学院出版社出版  
(北京市朝阳区红庙)

北京经济学院出版社永乐印刷厂印刷  
新华书店北京发行所发行

850×1168毫米 32开 6.125印张 159千字  
1989年3月第1版 1989年3月第1版第1次印刷

印数：0001—7000

ISBN 7-5638-0051-4/F·40

定价：2.70元

2166161  
《诺贝尔经济学奖获奖者著作丛书》

编 辑 委 员 会

顾 问 厉以宁 宋承先 张培刚 胡代光  
钱荣堃 高鸿业 黄范章 谭崇台

主 编 贾 湛 梁小民

编 委 于同申 马在新 甘华鸣 厉以平  
吕立勤 李玉臣 李盛平 余永定  
张家慈 沈志华 邹 蓝 吴衡康  
杨河清 杨德明 俞品根 施 炜  
彭 元 彭剑锋 薛进军

## 编者献词

近年来，我国经济学界介绍了不少现代西方经济学家的著作，对我国经济学理论的研究起了积极的作用。为了使这种介绍更具有全面性和代表性，我们决定系统地翻译出版获诺贝尔经济学奖的学者的主要著作，作为翻译丛书奉献给广大读者。

诺贝尔经济学奖是当今世界上最有影响的经济学奖。它是1968年瑞典中央银行在其建立300周年时为纪念诺贝尔奖提供者而设立的基金，全称是“纪念诺贝尔经济学奖”。该奖由瑞典皇家科学院委任组成的“经济科学委员会”评定。从1969年起在每年的10月中旬与其他各项诺贝尔奖同时公布，到1987年为止已颁发19届，共有欧美的25位经济学家获奖。获奖的经济学家的经济理论基本上代表了西方经济学的主要成就，对西方经济学的发展有较大的影响。读者通过这些著作能全面、深入、系统地了解和研究西方经济学，从中吸取对我国经济改革和经济学发展有价值的东西。

当代西方经济学最显著的特点之一是数学方法的应用。经济学要反映经济中的数量关系，因此，经济学与数学早就结下了不解之缘。从17世纪起经济学家就用数学公式来表述经济理论。本世纪以后经济学家又把经济理论、数学和统计学结合起来建立了经济计量学，用于解决实际问题，并取得了重要进展。从某种意义上可以说经济学的数学化是经济学精密化、实用化的标志。在获奖的经济学家中，有三分之一以上在这方面作出了开创性的贡献。获得第一届诺贝尔经济学奖的挪威经济学家尼·费瑞希和荷兰经济学家J.丁伯根是经济计量学的创立者。以后获奖的美国经济学家P.萨缪尔森、K.阿罗、W.列昂惕夫、T.库普

曼、L.克莱因、G.德布鲁，英国经济学家J.希克斯、苏联经济学家L.康托洛维奇都在经济学数学化方面作出了重要贡献。

在流派林立的当代西方经济学中，新古典综合派是公认的“主流派”。这一派继承并发展了凯恩斯主义，对西方各国的经济政策和经济理论都有重大的影响。因此，在获奖的经济学家 中属于这一派的有相当比例，如美国经济学家P.萨缪尔森、L.克莱因、J.托宾、F.莫迪利亚尼、R.索洛，都是这一派的代表人物。

70年代后西方经济政策与经济理论的显著变化是自由主义的影响日益增强。代表自由主义的奥地利经济学家F.哈耶克，美国经济学家M.费里德曼、G.斯蒂格勒、J.布坎南先后获得诺贝尔经济学奖，便反映了这种趋势。

经济学是一门实用性很强的科学。因此，获奖者中还有一批对解决各种实际问题作出贡献的经济学家，如：美国经济学家S.库兹涅茨和英国经济学家R.斯通对国民收入统计作出开创性贡献；美国经济学家A.刘易斯和T.舒尔茨对研究发展中国家经济发展问题有独特贡献；瑞典经济学家B.奥林和英国经济学家J.米德对国际贸易和国际经济关系问题作出重要贡献；美国经济学家H.西蒙对管理科学的发展起了重要作用。

本丛书的选编以获奖者的代表作，特别是在西方经济中影响重大的名著以及对我国经济改革和社会主义经济理论有启发和借鉴的著作为主，每位获奖者的著作至少选一本，以保持其系统性与代表性。

还应该说明的是，西方经济学是资本主义的意识形态，获奖的经济学家是资产阶级学者，因而在这些著作必然有这样那样的问题，在阅读时我们应注意这一点。但是，西方经济学作为人类文明发展的成果之一，有其不可忽视的精华，资产阶级学者对经济问题的研究有其可取之处。从这种意义上说，只要我们以分析的

态度认真阅读这些著作，必将得到许多有益的收获。

本丛书的编辑出版得到了老一代学者与中青年学者的关心与帮助，谨向关心与帮助这套丛书编辑出版工作的所有朋友表示感谢。

编 委 会

1988年4月

## 编 选 说 明

有限理性说，是赫伯特·西蒙（司马贺）的全部重要理论的核心学说。鉴于它在经济理论、管理科学、计算机科学、认知心理学、设计科学和科学哲学等领域有着重大影响和显要地位，我国一些著名学者自70年代末以来已对它做了很多精细描述和介绍。西蒙教授本人多次来华讲学，以及《管理决策新科学》、《管理行为》和《关于人为事物的科学》等译本的出版，使得这一学说的概貌对我国读者来说，已经是比较熟悉的了。但是，对于它的深刻而广泛的真正含义、它同其他有关学说的差别和关系，它在科学描述上的意义和在实际应用上的作用及其区别，以及它的近期发展动态，显然还需要做进一步的研究。我们编选、翻译这本集子，就是为了给这方面的研究工作提供重要资料。当然，读者如希望了解有“科学奇才”、“世界著名博学家”之称的西蒙教授是如何驰骋于自然科学和社会科学这两大领域前沿的；或者仅希望一般性地了解有限理性说及其在诸多学科上的意义，无疑也会发现本书的价值。

根据逻辑关系和时间顺序，我们将本书全部内容分成了两部分。第一篇（“有限理性说”）收录了西蒙的5篇重要文章，其中第一篇描述了理性定义，指出了“理性”一词的多层次、多方面、多学科的不同含义及其联系。这篇短文（辞条）是了解西蒙“理性”或“合理性”概念的重要原始资料。①

---

① 西蒙的这一概念有其独到之处，不能被看作是一般的（传统的或某些当代的）理性概念。可参见G.米歇尔：《社会科学新词典》（纽约：阿尔顿出版公司1979年版）第154—155页。

有限理性说最初是西蒙在《管理行为》一书（第4、5章）中提出来的<sup>①</sup>。本书第二、三篇论文就是对这一基本理论的形式化概括。西蒙在谈到这两篇论文时说过：它们“为比较有限理性说和那些体现在古典经济理论及统计决策论中的各种完美理性说（全面理性说），为比较它们两者的基本假定和结构，提供了一个良好的基础。如果让我从自己已经发表的全部经济学文献中仅仅挑选两篇，送给刚刚被我们发现的外星文明人，我就会挑选这两篇。由于这两篇论文对理性行动者所做的全部假定，实际上就是面临极度复杂性时的计算顺序性（串行性）<sup>②</sup>，因此，我相信，收到这两篇论文的外星人会毫无困难地理解它们的意思，并应用于他们自己的事务。当然，我写这两篇论文是要解释地球上的事情，而不是其他世界上的事情。”<sup>③</sup>这里反映了西蒙撰写和评价这两篇论文的科学立场。它们除了在经济学方面使读者感兴趣之外，作为人工智能与认知科学先驱在人工智能和认知科学发展前夜所写的重要论文，也必定会使从事计算机科学和认知科学研究工作的我国读者感兴趣。西蒙在他后来编撰的一部经济学论文集（《有限理性模型》）和一部认知心理学论文集（《思维模型》）中，均收录了这两篇论文，同样说明了这一点。

收入本书第一篇的第四篇论文，详细地介绍了认为理性有限的一些理论（“有限理性论”）的结构，特别突出地论述了以有限理性说为基础的有限理性论。西蒙在撰写和发表该文时，已经在《关于人为事物的科学》（1969年第1版）一书中提出了建立“设计科学”的设想，这也是当今西方“设计研究”界公认的首

---

① 中译本列入“诺贝尔经济学奖获得者著作丛书”，已由北京经济学院出版社出版。

② 即指“一时干一事”、“一时想一事”而不是“好多事”。详见正文。

③ 《有限理性模型》（坎布里奇：麻省理工学院出版社1982年版），下卷，第204页。

次系统设想。④因此，西蒙在该文中还特别谈及“设计领域的有限理性”。此外，该文对管理科学上的分析方法和以有限理性说为依据的启发式方法，也做了中肯的论述和评价。

第五篇文章是西蒙获得1978年诺贝尔经济学奖时发表的演讲。这篇文章集中论述了有限理性说在经济学界的发展经历、当前状况和未来前景，这自然就需要涉及围绕有限理性说而进行的其他学科领域的工作，比如心理学、组织理论等方面成果。西蒙有一个非常著名的观点，即：经济学（特别是厂商理论）、认知心理学和组织理论所研究的课题，实际上是同一事物——人的决策过程和问题求解过程。以有限理性说为根据、从决策角度出发研究经济现象的经济学分支，在“决策理论”的总标题下，有时又叫作“行为经济学”或“行为理论”。西蒙及其同事们所发展的这一理论，在现代西方厂商理论或企业经济学上，占据着强大优势。瑞典皇家科学院1978年诺贝尔经济学奖公告指出：“现代厂商经济理论和管理研究，大多是以西蒙的思想为基础的。”⑤但是，西蒙等人的行为理论不仅对微观经济学很重要，对宏观经济学也是有重要意义的。西蒙说：“…在这篇文章（《企业组织的理性决策》）中，我试图概述程序理性研究和企业行为理论自第二次世界大战以来的整个发展经历…。读者将会看到，我有一个毫不动摇的信念，那就是：人类思维过程与其产物一样，是经济理论和经济知识持续进步的具有基本重要性的课题——对厂商理论来说当然尤其如此，对经济周期理论和宏观经济学其他课题来说，也是这样。”⑥西蒙还说：“关于经济增长和经济周期的正确解释是什么，其现存不确定性，靠新古典体系中的总量分析是无法予以解决的。现在的理论争论，主要在于一些特定的、个

① 详见杨砾、徐立：《人类理性与设计科学》，辽宁人民出版社1987年版。

② 《斯堪地那维亚经济学杂志》，1979年，第1期，第72页。

③ 《有限理性模型》下卷，第403页。

别经验性的假定，在于有关不确定情形下偏离全面理性此一事实的这种假定。这类争论，只有通过针对人类决策和问题求解过程的辛辛苦苦的微观经济学实验研究，才可以得到解决。”<sup>①</sup>

本书第二篇是西蒙近著《论事理》的全部译文。这是西蒙依据有限理性说去探讨新问题的重要作品。其中概述和比较了四种最突出的理性说；也代表西蒙对其有限理性说的最新解释。这一部分中有关进化论与有限理性说的关系，和关于生物社会学观点的述评，在国内介绍西蒙理论的论著和译本中还相当少见。这一部分还特别突出地探讨了有限理性说在社会事务上的意义和应用；社会制度“理性”的限度及其扩展途径；科学知识在社会生活中的重大意义；专家在公共政策制定上的作用等等一系列重要的新课题。读者不难发现，这当中有不少值得探讨的新观点。

同本书第一篇相比，第二篇在论述方式上一般比较通俗，内容更广泛，论题更为新颖。不过，正如西蒙本人所说，这一部分主要是“探索一下迄今一直不在我本人的主要探索道路上。但我感觉有兴趣的和重要的偏远小路。…集中探讨那些尚成问题的或争论不休的课题，以及那些对于理解理性在人类事务中的角色尤为关键的专题。”可见，这一部分有关社会制度等方面的内容，同该部分有关各种理性说的论述及本书第一篇内容相比，具有更大的探索性和猜测性。我们不准备在这里用马克思主义观点同西蒙有关社会制度方面的观点进行分析、对比。但必须指出的是，它们两者在出发点、侧重点和论证方式上，都存在着差异。尽管西蒙的有些观点同马克思主义观点相抵触，必须以批判的眼光对待之，但若从管理技术的角度去看，无疑会发现其中的合理成分。而且只有从这一角度去理解一个科学家和管理学家的论著，才是比较合适的。我们还应注意到，西蒙虽然以“人类事务”、“社会事

<sup>①</sup> 《论经济动态学的行为基础和理性基础》载《经济行为与组织》，1984年，第5期，第35页。

务”为题，但他讲的许多（不是全部）是欧美的事情。在阅读这方面的内容时，不做分析地进行简单类比或外推，是不正确的；不考虑社会背景就加以贬斥，也是不正确的。

最后，我们还想稍许说明一些译法问题。第一，“理性”有好几个替代译法，如经济学界有时采用的“合理性”和其他地方有时采用的“理智”、“有理”、“推理”等等。第二，“有限理性”也可译为“有限制的合理性”、“有界合理性”、“有约束的理性”或其他组合。第三，“有限理性说”同“有限理性论”在很多情况下是可以替代的，但根据西蒙本人所做的区别，我们用前者特指西蒙关于有限理性或合理行为的学说要旨，后者有时则指包括西蒙学说在内的多种合理行为理论——它们都给理性加上了限制和约束条件，但所加条件不同（不过，在某种程度上，它们都符合有限理性说，属于同一个“时代精神”）。第四，本书第二篇的“事理”是一个已被赋予新含义的词，指“人类事务上的理性或道理”。我国科学界早就有人主张建立一门与“物理学”相对应的“事理学”。这一主张在管理科学和运筹学（系统工程学）界常常可以听到，后来甚至有人如法炮制所谓“人理学”。尽管据我们所知“事理”一词还没有一个明确的、系统的、学术上的统一定义，但该词的含义还不至于引起混乱。本书第二篇（特别是最后一章）讲的就是事理。当然，它并不是可能的“事理学”的全部，而只是其中一个较高层次上的部分内容。除了以上说明的译法之外，还有个别译法问题将在适当的脚注里予以说明。

西蒙学术思想在众多领域里的重要地位，是得到了当今世界学术界公认的。经济学上就不必多说了。<sup>①</sup>从心理学上看，出版包括认知心理学著作在内的现代心理学著作，在促进我国学术进步

---

<sup>①</sup> 可参阅《管理行为》中译本序言（范家骥撰），或厉以宁、秦婉顺：《现代西方经济学概论》（北京大学出版社1983年版），第十五章。

上，是有重要意义的。

虽然我们力求译好这本书，但由于能力有限，误译之处在所难免。真诚希望读者提出宝贵意见。

杨砾徐立

1987年12月14日

# 前　　言

为中译本出版而作

为了建立理论，特别是为了建立那些使我们能够据之推理的理论，我们要对现实进行简化。我们不是试图捕捉真实世界的每一个复杂因素，而是仅仅抓住其中最重要的因素，并且小心防止使我们从理论中得出的推论，超越理论本身对现实的近似界限。

古典经济理论对人的智力做了极其苛刻的假定，为的是产生那些非常动人的数学模型，用来表示简化的世界。在这方面，近年来人们已经提出了疑问，怀疑那些假说是否与人类行为的事实相距过远，以至根据那些假说所得出的理论同我们所处的现实状况已经不再有什么关系了。本书各篇文章正是提出了这种疑问。

不过，除了批评正统经济学家关于具有完美理性的经济活动者的观念之外，我还试图更进一步。我试着描述了一种受到较多限制的理性。在人的决策行为方面，这种理性同事实相吻合的程度要好得多。我认为，我所提出的“寻求满意的人”这一模型同那种“追求最优的人”的古典概念相比，在有关经济行为的经验研究中获得了多得多的支持。

我们在理论中采用不同的人类行为模型，会给经济政策造成差异。如果说人们只具备有限的理性，那么搞清楚实际决策过程就变得重要起来了。在这种情况下，我们就不可效仿经济学家们那样单纯地论证说：人们会采取最合理的决策。因此，本书各篇文章讲的是寻求满意的人的决策过程。

杨砾和徐立先生为翻译这些文章而付出了辛勤劳动，我要向他们表示感谢。不过，书中任何缺陷的责任当在作者，不应归咎于他们两人。

赫伯特·西蒙

1988年1月9日于美国宾州　匹兹堡

# 目 录

## 前 言

### 第一篇 有限理性说

<b>一 理性的定义</b> .....	( 3 )
<b>二 理性抉择的行为模型</b> .....	( 6 )
理性抉择的某些普遍特征.....	( 8 )
实质性的简化.....	( 12 )
解的唯一性与存在性.....	( 19 )
关于动态关系的进一步说明.....	( 22 )
结论.....	( 23 )
附录：关于如何合理决定可接受报酬的一个例子	
.....	( 25 )
<b>三 理性抉择与环境结构</b> .....	( 29 )
生物的环境.....	( 31 )
感知力，贮存能力与生存.....	( 33 )
多目标抉择机制.....	( 36 )
关于环境的进一步说明：线索.....	( 39 )
关于多目标的最后评注.....	( 41 )
结论.....	( 43 )
<b>四 有限理性论</b> .....	( 45 )
引言.....	( 45 )
棋艺的理性抉择途径.....	( 49 )
设计领域的有限理性.....	( 57 )
管理科学上的有限理性.....	( 59 )

结论	( 62 )
<b>五 企业组织的理性决策</b>	( 64 )
作为经济学的决策理论	( 65 )
刻画有限理性的特征	( 76 )
新古典理论复兴	( 83 )
行为理论的进展	( 88 )
结论	( 94 )

## 第二篇 论事理——人类事务上的理性

<b>一 几种理性说</b>	( 106 )
推理限度	( 107 )
价值	( 109 )
主观期望效用	( 113 )
行为说	( 117 )
直觉理性	( 122 )
直觉与情感	( 126 )
结语	( 130 )
<b>二 理性与目的论</b>	( 132 )
作为理性适应的进化	( 132 )
达尔文模型	( 135 )
社会进化与文化演变	( 144 )
进化过程中的利他主义	( 147 )
进化的短见	( 153 )
结语	( 157 )
<b>三 社会事务上的理性过程</b>	( 160 )
制度理性的限度	( 163 )
制度理性的加强	( 169 )
公共信息库	( 173 )
结语	( 182 )

# 第一篇

# 有 限 理 性 说



## 一 理性的定义<sup>①</sup>

1. 广义而言，理性指一种行为方式，它（1）适合实现指定目标，（2）而且在给定条件和约束的限度之内。

2. 在某些特殊场合下，这个定义的（1）、（2）两方面可以有更精细的规定。这类特殊用法中的重要者包括：

（1）目标可假定是效用函数期望值在某一时间区域上极大的形式（对策论称此为极小极大化）。效用函数的存在性，可从抉择者偏好的有序性和一致性假定上导出。形式化的经济理论就是这样认为合理消费者谋求最大期望效用而合理企业家谋求最大期望利润的。如果要使这种极其严格的形式区别于更一般的形式，则可将前者称为最优化，称后者为适应性或功能性。

（2）目标可假定是意欲达到的一些准则所构成的，但要么全部达到，要么全都达不到（例如欲望水平的达成）。

（3）条件和约束的一般定义，可指抉择者的外部环境的客观特征，可指该环境被感知到的特征，也可指以固定形式出现的、不受自身支配的抉择者自身特征。我们有时通过区分客观理性和

---

① 本文是西蒙为《社会科学辞典》撰写的条目。吉尔德和科尔伯编，自由出版社1964年版，第573—574页。另载西蒙：《有限理性模型》，下卷，第405—407页。——译注

主观理性及有限理性，来区别上述第一种定义和后两种定义。

(4) 定义中的所谓目标，可以是抉择者的目 标，抉择者所 属社会系统的目 标，或观察者的目 标。

(5) 毫不含糊地使用理性一词，要求使用者讲清楚他对 目标和条件这两者所做的假定。

3. 在社会科学作为独立学科群出现之前，理性及其同义词在 哲学和伦理学语汇中十分重要。理性一词的当代用法，同亚里士 多德的谋略或慎思精神的概念非常接近。在这个意义上，一项行动 的理性，包括该行动从成立的前提出发得以导出的逻辑过程。 理性有时指运用才智进行选择的过程，有时则指抉择本身。前者 多在早期心理学、逻辑学和伦理学中使用，后者则在经济学和社 会学界流行。

(1) 詹姆士在《心理学原理》(纽约：霍尔特出版公司 1890年版，第22章)一书中将理性作为“称作推理的特定思考过 程”的同义语。按照这种观点，抉择的理性取决于做出抉择的过 程。相应地，非理性一词在心理学文献中指的是，主要靠情感机 制(情绪、激励、天性、冲动)而非理智机制所进行的抉择(阿尔波特：《现代社会心理学的历史背景》，载林蔡编：《社会心理学 手册》，坎布里奇：艾迪生—韦斯利出版公司1954年版，第15—18 页)。由于理性同理性主义说教在历史上的联系，近来心理学论 著趋向于采用认知过程(布鲁纳、古德诺和奥斯丁：《思维研 究》，纽约：威利父子出版公司1956年版，第Viii页)或理智过 程的提法。因此，在社会科学文献中，指抉择过程的理性一词， 看来正在消声匿迹。

(2) 经济学家一般用理性一词表示靠抉择过程挑选出来的 行动方案的属性，而不是表示抉择过程的属性。达尔和林德布鲁 姆这样说道：“一项行动是理性的，就是说，对于指定目标及 其真实处境来说，该行动被‘正确地’设计成为一种能谋求最大成

功的行动”（《政治学，经济学和福利》，纽约：哈泼兄弟出版公司1953年版，第38页）。在社会学文献中，我们也可找到与此类似的定义，例如韦伯的《社会经济组织理论》（汉德森、帕森斯译，纽约：牛津大学出版社1947年版，第117页），曼黑姆的《重建时代的人与社会》（伦敦：基冈·保罗出版公司1940年版，第51—57页），帕森斯的《社会系统》（自由出版社1951年版，第549—550页）。（帕累托则更喜欢用“合乎逻辑”一词，而不是“理性”一词。）这个术语在社会学上用得比较随便，但在数理经济学和统计决策论中则有严格的公理描述（见上文2.）。

(3) 在韦伯和其他人关于官僚制的论述中（如《社会经济组织理论》，第329—341页），理性还有另外一种有点特别的用法：对理想型的“合理合法当局”来说，理性指的是组织对目标的自觉适应性，指的是组织的运转不受组织成员个人目标的有害影响，而是靠规章制度的非人格化的实施。近似的同义语是曼黑姆的功能合理性（《重建时代的人与社会》，第53页）。

## 二 理性抉择的行为模型<sup>①②</sup>

传统经济理论假定了一种“经济人”，这种人在行动过程中既具有“经济”特征，同时也具有“理性”。传统经济理论认为，这种人具备关于其所处环境各有关方面的知识，而且，这些知识即使不是绝对完备的，至少也相当丰富、相当透彻。此外，这种人还被设想为具备一个很有条理的、稳定的偏好体系，并拥有很强的计算技能；他靠这类技能就能计算出，在他的备选行动方案当中，哪个方案可以达到其偏好尺度上的最高点。

经济学的近期发展，尤其是厂商理论的进展，对这种系统化的经济人模型，提出了尖锐的挑战；还对这种模型是否为建立厂

---

① 本文原载《经济学季刊》第69卷（1955年），第99—118页。另载西蒙的《思维模型》和《有限理性模型》，译文是参照几个版本译出的。西蒙本人对这篇论文及下一篇论文评价很高。他说，这两篇论文提出了“有限理性的形式模型…如果让我从自己已经发表的全部经济学文献中仅仅挑选两篇，送给刚刚被我们发现的外星文明人，我就会挑选这两篇。”（《有限理性模型》，下卷，第204页。）——译注

② 本文所体现的想法最初是在同赫伯特·鲍纳特、诺曼·达尔基、杰拉尔德·汤普森和罗伯特·沃尔夫森于1952年夏季进行的一系列讨论当中形成的。对理性抉择进行研究的这一途径如有任何好处，其很大一部分理应归功于我的这些同事。本文初稿原是为我成为兰德公司顾问而准备的。后来我在同考尔斯经济学研究基金会合作研究“不确定情形下的决策”时，对本文初稿做了进一步的扩充和改进（包括本文附录）。这项研究工作是根据同海军研究部所签合同进行的。本文的最终完成得到了福特基金会的赞助。

商理论（无论是描述厂商~~真实~~行为的理论，还是论述其“应当”怎样合理行动的理论）提供了妥当的基础，提出了深刻的疑问。本文的目的不是要讨论这些疑问本身，也不是要确定这些疑问是否已得到证实。相反，我将假定经济人这一概念（也许还可以加上经济人的兄弟——管理人）确实需要做重大修正，并就进行这种修正的可能方向，提出一些建议。

广言之，本文的任务是要用一种符合实际的理性行为，来取代经济人那种全智全能的理性行为。我们所提出的符合实际的理性行为，就是符合生物（包括人在内）在其生存环境中所实际具备的信息存取能力和计算能力的一种理性行为。在这方面，人们可以向心理学文献索求答案。心理学家当然一贯关心理性行为，他们对学习现象的兴趣尤其如此。不过，关于学习和抉择过程的现有心理学知识，同经济学和管理理论所需要的心理学知识相距如此之远，那条遥远道路上的界碑会使两个方向上的行路人望而生畏。

由于缺乏一套明确理论所需的关于决策过程的经验知识，现在看来，我们还只能以不太系统的、不太严密的方式，把有关真实世界的确凿事实纳入理论当中。另一方面，大家都了解人类抉择的一些大致特征，了解人类抉择所处环境的主要特点。就建立关于人及其世界的本质的理论这一目的而言，我将随意利用那些常识性的经验，作为所需理论假设的依据。

对这里提出的问题，我们基本上有两条探索途径。一条途径是研究抉择者的性质；另一条途径则是研究抉择环境。本文将采用前一途径。以后我还要进一步探讨环境特征，以及环境与抉择者的相互关系。<sup>①</sup>

---

① 西蒙在本文发表后的第二年，在《理性抉择与环境结构》（见本书第一篇，三）一文中完成了这一课题。——译注

因此，本文试图把抉择者的某些性质作为要素，明确地包含在特定情境下的理性行为的定义之中，并用这一定义来描述合理行为的选择。一方面，本文将研究这方面的一些近期工作尚未明确论述的东西，予以进一步明朗化——信息条件既可视为决策者的环境特征，又同样可以认为是决策者本身的特征。另一方面，本文还包括一些新的观点——特别是本文考虑到了，抉择者会有意简化其情境模型，以使模型适合其计算能力。

### 理性抉择的某些普遍特征

种种理性抉择模型，各有各的“风味”。究其根源，主要是因为在模型中引入了一些特定的假设，认为合理的适应行为必然发生在那些“给定条件”或约束下。各种模型的共同约束包括：

(1) 可供选择的一组替代方案或备选方案；(2) 决定着报酬①(“满足程度”，“目标实现程度”)的关系，作为已选定的方案的函数；(3) 报酬的优劣次序。这些约束本身并不是合理计算的对象。一个理性行为模型究竟选用(或放弃)了哪些具体的约束，要看这样一些隐含的假定：首先，理性的生物“控制”哪些变量——从而能将这些变量作为进行合理调节的手段，进行最优控制；其次，它必须把哪些变量视为固定不变的。此外还有关于固定变量特征的假定。例如，在备选方案与报酬之间的关系方面，生物拥有一定的信息量，关于这一信息量的不同假定，可能会导致选用不同的最优化过程，比如选用某个最大期望值，或极大极小期望值。

刻画给定条件和行为变量的另一种办法，是用行为变量指生

---

① 又译支付、赢得、损益等。下文“报酬函数”可类比于通常所谓“目标函数”、“费用函数”或“指标函数”等。——译注

物体本身，用给定条件指生物的环境。但是，如果采取这种观点，我们就必须承认这样一种可能性，那就是：我们所谓“环境”的一部分，可能存在于生物机体的内部。换言之，在最优化问题中必须取为给定条件的某些约束，可能就是（生物学上定义的）生物自身的生理、心理限度。例如，一个动物所能达到的最高运动速度，便给其可能行为设定了一种界限。与此类似，计算能力的局限性，对于具体场合下的理智抉择的定义来说，可能也是一种重要的约束。我们将探讨理性抉择在特定条件下形成可能方式。这里所说的特定条件，就是我们希望加以明确考虑的“内部”约束和“外部”约束这两者。正是它们，规定了生物的理性问题。

无论我们的兴趣是在理性抉择的描述性方面，还是在规范性方面，这类模型的建立都将对我们有所助益。由于生物本身是受到心理限度的制约的（尤以计算、预见能力的限度为甚），因此，真实的人对理性的追求，至多只能是对策论之类模型那种完美理性的极其粗略的、简化的近似而已。生物实际使用的这种近似理性固然不是最佳的——即使从其有能力处理的计算复杂性水平上看，也不是最佳的——但是，通过考察人类和其他生物所实际使用的近似理性图式，我们或许能对那些可能的机制获得相当多的了解。

为描述我们提出的模型，我首先从完美理性模型的共同要素入手，然后再介绍简化假定和近似程序（这与简化假定是一回事）。

### 原始术语和定义<sup>①</sup>

理性行为模型——无论是通常构造的那种全面理性模型，还

---

<sup>①</sup> 下文采用了集合论等数学语言。由于有关数学知识可在普通数学教本或手册中找到，译文一般不加注。——译注

是这里讨论的有限理性模型——一般均需用下列要素的一部分或全部：

1. 一组备择行为（抉择或决策的备选方案）。这在数学模型中可用一个点集  $A$  来表示。

2. 生物“考虑到的”（或“感知到的”）备择行为子集。也就是说，生物的抉择范围可能要比客观存在的备选方案范围小。

“考虑到的”子集可表示为点集  $\hat{A}$ ， $\hat{A}$  包含于  $A$  ( $\hat{A} \subset A$ )。

3. 未来可能状态，或称抉择结果，以点集  $S$  来表示。（现在暂不必区分实际结果和感知的结果）。

4. 一个“报酬”函数，用来表示生物给抉择的各个可能结果赋以多大的“价值”或“效用”。我们可以用一个实函数  $V(s)$  来表示这种报酬（其中， $s$  为  $S$  的任一元素）。在很多情况下，我们所要知道的，仅仅是  $S$  中各对元素的次序关系——即  $s_1$  优于或劣于  $s_2$  的关系。但为了避免不必要的复杂化，现在我们姑且假定，已经定义了一个基数效用函数  $V(s)$ 。

5. 关于一旦选定了  $A$  中的一个具体方案  $a$ ，将实际出现  $S$  中哪些结果的信息。这类信息可以是不完备的——也就是说，每个备择行为  $a$ ，都可以有一个以上的可能结果  $s$  与之对应。因此，我们用  $A$  中每个元素  $a$  到子集  $S_a$  的映射，来表示这种信息。这里， $S_a$  表示选择  $a$  时所可能导致的结果的集合。

6. 关于一旦选定了一个具体方案，一个特殊结果将会发生概率的信息。这类信息要比上面（5）假定的信息更精确，因为它把集合  $S_a$  中的每个元素  $s$ ，同一个概率函数  $P_a(s)$  联系在一起了。 $P_a(s)$  表示选择  $a$  时， $s$  将出现的概率。它是一个非负实函数，而且

$$\sum_{S_a} P_a(s) = 1$$

我们要注意备择行为集合  $A$ 、结果集合或未来状态集合  $S$  和

报酬函数V之间的定义区别。以简化形式的对策论为例，按照通常采用的报酬矩阵表示法，集合S对应于矩阵的各元素，集合A对应于第一个局中人的策略，函数V对应于元素的价值。这样，集合 $S_a$ 就对应于矩阵第a行的元素。读者如记住这一解释，就可将这里的公式同“经典”对策论做比较了。

### “古典”理性概念

采用这些术语和定义，我们就可以确定一般对策论模型和概率模型所对应的理性抉择程序了。<sup>①</sup>

**甲。极大极小法则。**假定无论选择哪个备选方案，都将发生其相应可能结果中的最糟结果——将出现 $S_a$ 中 $V(s)$ 值最小的 $s$ 。因此，要选取一个使这种最低报酬尽可能高的备选方案 $a$ 。

$$\hat{V}(a) = \min_{s \in S_a} V(s) = \max_{a \in A} \min_{s \in S_a} V(s)$$

我们可以不对客观存在的备选方案集合A求极大，而代之以对“考虑到的”备选方案集合 $\hat{A}$ 求极大。选择结果的概率分布(6)在极大极小法则中不起任何作用。

**乙。概率法则。**对(假定已知的)概率分布 $P_a(s)$ ，求 $V(s)$ 的最大期望值。

$$\hat{V}(a) = \sum_{s \in S_a} V(s) P_a(s) = \max_{a \in A} \sum_{s \in S_a} V(s) P_a(s)$$

**丙。确定性法则。**给定A(或 $\hat{A}$ )中每个a到S中一特定

<sup>①</sup> 参见肯尼斯·阿罗(Kenneth J. Arrow)：《构造风险抉择理论的几条替代途径》，载《经济计量学报》，XIX(1951年10月)，第404—437页。

$S_*$ 上的映射，选取其结果报酬值最大的备择行为。

$$\hat{V}(\hat{a}) = V(S_{\hat{a}}) = \max_{s \in A} V(s)$$

### 实质性的简化

周密考察以上概述的“古典”理性概念，我们就能立即看到，它们给作为抉择者的生物强加了一些严格要求。生物必须有能力给每个可能结果赋以明确的报酬值（或至少是一个明确的报酬范围）。这当然也包括确定那些结果的精确性质的能力——这一图式没有给“出乎意料的后果”留下任何余地。报酬必须完全排好优劣次序——它必须总能始终一致地确定某一结果比另一结果好、一样或欠佳。而且，若使用确定性法则或概率法则，那就要么必须知道具体备选方案的结果，要么至少必须能给每个结果指定一个确定的概率。

我要讲的第一个经验命题是：对于具有任何复杂程度的人类真实抉择来说，完全不存在表明人类能进行（或实际上等价于进行）这些计算的任何证据。内省方面的证据诚然足够清楚，但我们也当然也不能排除无意识决策优于有意识决策的可能性。无论如何，在经典概念所描绘的那种决策制定过程没有证据的情况下，我们完全有理由考察一下真实过程与那些法则所描述的过程颇为不同的可能性。

我们的程序是对上面列举的模型进行某些修正，这些修正（按照草率的经验论）看来对应于我们观察到的人类行为过程，并且能大大简化做抉择所需的计算。这决不是说，人类在任何时候都使用所有这些修正的和简化的程序。本文也决不想担负确定人类在多大程度上、在什么情形下实际使用简化程序这一艰巨的经验考察任务。实际上，我们所要指出的是，这些程序看来

是人类在复杂抉择处境中为找到能够处理的近似模型而经常使用的程序。

### “简单”报酬函数

进行简化的一个办法是假定：对S中的所有s来说， $V(s)$ 只须取(1, 0)两值之一，或取(1, 0, -1)三值之一。根据具体情况，我们要把这些值解释为：(1)满意或不满意；(2)胜、平或负。

作为(2)的一个例子，设S表示一盘国际象棋赛中白方第20步棋的可能状态。那么，(+1)表示白方的位置使其获得了一个优势策略，无论黑方怎么走，该策略总能争得赢局。(0)表示白方能争得平局，但不是赢局。(-1)则表示黑方能争得赢局。

作为(1)的一个例子，设S表示一个人出卖一所房子的可能要价。他可能认为15 000美元是“可以接受的”价钱，大于这个数是“满意的”，小于这个数则是“不满意的”。从心理学理论上讲，我们可以在“欲望水平”<sup>①</sup>上固定这一界线；从经济理论上讲，我们可以把这条界线固定在卖不卖两可（机会成本概念）的价钱上。

有人可能会对此提出异议说，虽然16 000美元和25 000美元都是出卖房子的“非常满意”的价钱，一个理智的人将喜欢以高价出售，因此，简单报酬函数并不是抉择问题的恰当表达。对这种异议，我们可给以多种不同回答。每一种回答对应于一种处境，其中简单函数看来都很恰当。

首先，卖主可能不会同时碰到好几个出价不同的买主，相反，他可能依次收到一系列报价，必须先决定是接受还是拒绝一

① 又译作“抱负水平”、“志向水平”等。——译注

个人的报价，然后才能转而考虑下一个人的报价。（或者更一般地讲，卖主可能收到两个一组、三个一组或n个一组的报价，必须先决定是否接受n个一组报价中的最高者，然后再接收、考虑下一组报价。）因此，如果S中的元素对应于n个一组报价，那么，只要其中的最高出价高于卖主当时决定的“接受价”， $V(s)$ 就等于1。这样一来，我们就可以提出进一步的问题：决定接受价的合理过程是怎样的？①

其次，即使存在着一个更一般的报酬函数 $W(s)$ ，其取值范围不限于两个不同值，简化函数 $V(s)$ 仍可以是 $W(s)$ 的一个满意的近似。例如，假定我们可设法引进一个定义在S上的基数效用函数 $U(s)$ 。再设 $U(W)$ 是一个具有二阶负导数的单调递增函数，其二阶导数是一个绝对值很大的负数（也就是说， $U(W)$ 是一个递减边际函数）。那么， $V(s) = V\{W(s)\}$ 就可以是图1所示的近似函数。

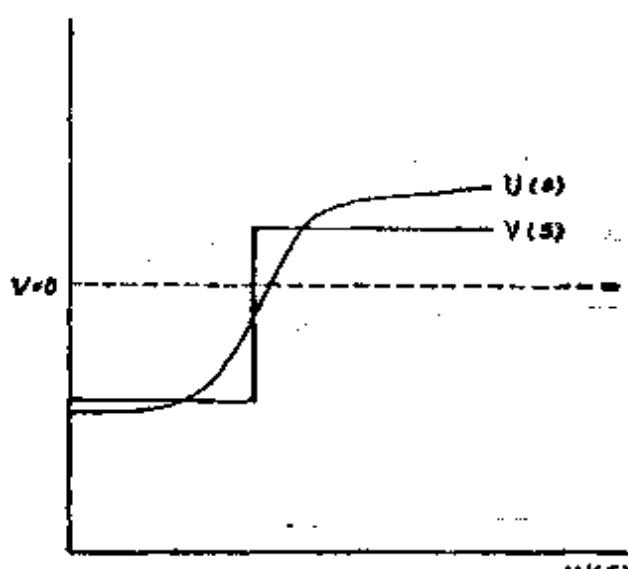


图1 具有简单报酬函数的近似效用

如果说，仅取(+1,0)两值的简单函数 $V(s)$ 在上述情形或其他情形下是可以接受的，我们就可以定义另一种（第四种）理性决策过程：

#### 丁. 满意法则

(1) 寻找一组可能结果(S中的子集 $S'$ )，使所有这些可能结果( $S'$ 中的全部元素s)的报酬值都是满

① 参见本文附录。在此可说明一下，华尔德(Wald)在可计算性方面引进统计决策论问题的简单风险函数，是这里所定义的简单报酬函数的一个例子。

意的 ( $V(s) = 1$ )。

(2) 寻找一个备择行为 ( $\hat{A}$  中的一个元素  $a$ )，其可能结果均在  $S'$  中（使  $a$  映射到一个集合  $S_s$  上，且  $S_s$  包含于  $S'$ ）。

如果我们能按这个程序找到一个备择行为  $a$ ，那就保证了获得一个满意结果。当然，这个程序并不能保证具有希望性质的  $a$  的存在性和唯一性。

### 信息收集

我们可能希望引进这样一个现实因素，即： $V(s)$  固然可以事先知道，但  $A$  到  $S$  的子集上的映射则不然。在极端情况下，每个元素  $a$  起初均可映射到整个集合  $S$  上。因此，我们可在决策制定过程中引进一些信息收集步骤，以此形成  $A$  中各元素到  $S$  中不同子集上的更精确些的映射。信息收集过程如果并非毫不费力，决策中就会产生这样一个问题：究竟把映射的精确性提高到什么水平为宜？

对简单报酬函数 (+1, 0) 来说，信息收集过程的一个重要方面是可以改进的。首先，我们设想一个人最初对  $A$  到  $S$  的映射只有非常粗略的了解。其次，我们设想他寻找  $S$  中的  $S'$ ，使对  $S'$  中的每一个  $s$  来说，都有  $V(s) = 1$ 。再次，我们设想他收集信息，以改进  $A$  到  $S$  的那一部分（包括  $S'$  的元素）的映射。最后，在进一步精确了解了映射之后，他寻找一个映射到  $S'$  的子集上的  $a$ 。

在顺利的情况下，这一程序可能只要求那个人收集不多的信息——同  $A$  中元素到  $S$  中各元素的整个映射相比，需要的信息量少得微不足道。他对具有希望性质的  $a$  的搜索如若获得了成功，他就会认为，要想使自己的抉择更好些，就不得不靠收集更多的信息。<sup>①</sup>

---

① 这个程序还使精确估计获得更多信息的费用成为不必要的了。关于这一点的进一步讨论，参见本文后面有关动态关系的论述。

我们刚刚描述过的决策过程，看来是棋手在棋赛中盘或末盘选择走法时所利用的重要手段之一。设 A 为白方第 20 步的可能走法。设 S 是以后达到的可能状态，比如第 30 步的可能状态。设 S' 为 S 的一个子集，由 S 中的明显“赢”势所组成。白方从他关于 A 到 S 的映射的十分粗浅的了解出发，试着选取一个映射到 S' 上的走法 a（假定黑方采取一定策略）。通过对黑方可能采取的种种对策进行思考，白方就“探索”了 a 的全部映射。他的这一试探可能会达到这样一些状态 s，它们虽然不在 S' 中，但当时却被认为是赢势。它们有可能通向 S'。另一方面，通过一系列的探索，也有可能使黑方达到某一状态，而且白方发现该状态对自己来说显然不是“赢”势。这样，白方就会放弃起初考虑的 a，另寻出路。

这个程序是否能使计算得到根本简化，取决于棋赛本身的经验事实。显然，所有状态可从客观上分为“赢”、“负”、“平”三类。但从棋手的角度看，所有状态可分为“定赢”、“定负”、“定平”、“或赢或平”、“或平或负”等等各类——这取决于映射的充分性。如果“定赢”状态是客观“赢”势集合的一个明显的子集，那么，检查一个状态是否可演变为定赢状态，针对黑方所有反应而采取的一切组合走法，可能并非可望而不可及的。<sup>①</sup> 同通常那种一般状态评价函数（取 -1 到 +1 之间的一切值）的更常见观念（开局时可能适用）相比，这一程序的优点在于，它意味着远非那么复杂、那么繁多的评价准则。它所需要的仅仅是：在检查一个状态（赢、负、平三者之一）何时变成另一

---

① 我曾采用两个中盘棋赛阵势做过实验，大致估计了中盘情况下所可能实现的简化程度。双方各走 8 步的一系列走法（16 个走法）总共可产生大约  $10^{24}$ （1 亿亿亿）种合乎规则的不同路线。但通过刚刚描述过的那种普遍程序，有可能将实际考察的那两种状态下的走棋路线，减少到不足 100 种不同路线——这是对抉择问题所做的令人惊叹的大简化。

状态的过程中，评价函数应相当灵。棋手不是在寻求“最佳”走法。只需寻找“好的”走法。

我们看到，通过引进简单报酬函数，以及对备择行为到可能结果的映射逐渐进行精确化的程序，达到理性决策的过程就可以在计算方面得到极大简化。在线性规划理论和实算当中，一般要区分两种计算，其一为确定方案可行性的计算，其二为发现最优方案的计算。<sup>①</sup> 可行性方面的检验，是要确定一项方案是否符合最初给定的线性不等式。例如，可将经济体系中的劳动力和钢铁生产能力的最大值，取为计划变动中的固定因素。这样，一项可行的计划，就是所需劳动力和钢铁生产能力不超过这一上限的计划。

最优计划则是使指定的报酬函数达到最大值的那些可行计划。如果我们不要求报酬函数最大化，只要求报酬函数值大于某个给定值，那么，我们就可以采用通常那种可行性检验方法，找到一个符合要求的计划方案。对报酬值的要求，无非是用一个必须加以满足的附加线性不等式来表示。只要符合了这一要求，就不必进而确定是否存在一个具有更高报酬值的替代计划了。

对任何实际目的来说，只要“合理地”设定了最低报酬要求，这个程序就可能是达到最优化的一个充分途径。在本文下一节里我要讨论一下，怎样能够做到这一点。另外我还要指出，怎样就能将这一图式推广到多分量的向量报酬函数的情形（最优化当然要求各项不同报酬之间完全可比）。

### 报酬的局部有序

古典理论不允许橙子和苹果不可比。<sup>①</sup> 它要求有一个纯量的报酬函数，即要求各项报酬完全可比，或者说完全有序。但是，

① 指要求不同物品在价值上、效用上是可比的，例如古典效用理论，可参看西方经济学教科书有关部分。——译注

我们可能不是只有一个纯量报酬函数  $V(s)$ ，而是有一个向量报酬函数  $V(s) = (V_1(s), V_2(s), \dots, V_n(s))$ 。引进向量报酬函数，可用以处理如下几种情形：

(1) 当一项决策要由集体制定时，分量可表示集体各成员的报酬函数。一个人的偏好，可能并不是其他人的偏好。

(2) 对一个人来说，他可能试图获得好几种价值，而这些价值彼此之间没有共同衡量尺度——例如，他在比较两种职业时，会考虑到薪金、组织风气、工作快乐感、声誉等等。

(3) 对于每一备择行为  $a$  均映射到  $n$  个可能结果集合  $S_a$  的情形，我们可用这样一个模型来替换之，其中，每一备择行为  $a$  仅映射到一个结果上，但每一结果的报酬则是一个  $n$  维向量，其分量表示  $S_a$  各元素的报酬值。

在这三种重要情形下，向量表达式有着突出的类同性：它们都由于缺乏报酬的完全可比和有序，而破坏了传统的最优化模型。第一种情形从来就没有得到过令人满意的处理—— $n$  人对策理论是研究这种情形的最热切和心怀奢望的努力，经济理论的所谓“弱福利原理”则试图回避这一问题。对第二种情形，通常的处理办法是给向量空间强加一种完全有序性（“无差异曲线”）。对第三种情形，通常的处理方法则是引进概率，作为分量求和的权数，或者是采用极小极大满意原则，或遗憾原则①。

对简化的报酬函数概念做一扩充，我们就能以基本上同样的方式，处理上述三种情形。让我们将  $V_i \geq k_i$  (对一切  $i$  皆成立) 的报酬，视为满意的报酬。我们可以得出如下合理决策法则：

戊. 局部有序法则。寻找  $S$  中的子集合  $S'$ ，使  $V(s)$  对  $S'$  中的所有  $s$  均为满意的（亦即： $V(s) \geq k, s \in S'$ ）。

---

① 这方面的有关细节，请参见普通运筹学教本中关于“决策论”的部分。  
译注

然后在 A 中寻找一个  $a$ , 使  $S_a$  位于  $S'$  当中。

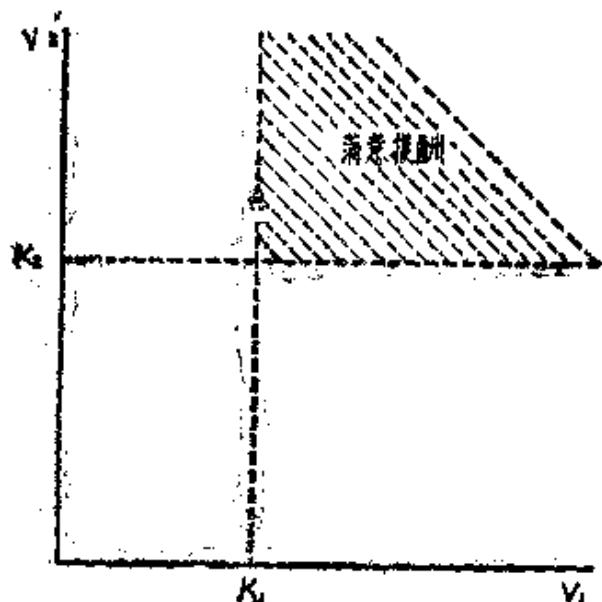


图2 报酬的局部有序性

这个法则仍不能保证解的存在性和唯一性。图2是该法则(法则戊)用于二分量报酬向量的一例。

在上述三种情形的第一种情况下, 满意报酬对应于我过去讲到的可用解(Viable Solutions)。<sup>①</sup>

在第二种情况下,  $V$  的分量定义了报酬各分量的欲望水平。对第三种情形来说(这时假定  $k$  个分量相等是很有些道理的),

$k_1$  可解释为最低报酬保障——这也是一种欲望水平概念。

### 解的唯一性与存在性

在以上通篇讨论中, 我接受了那些不能保证解的存在性和唯一性的决策程序。这样做, 是为了构造一个尽可能接近人类在复杂决策问题面前所实际采用的决策程序的模型。下面我再做一些补充, 以弥补这一空白。

① 西蒙: 《雇佣关系的形式理论》, 载《经济计量学学报》, 19: 第293—305页(1955年7月); 《组织理论的比较》, 载《经济研究评论》, 20: 第40—48页(1952—1953, 第1期)。

## 获得唯一解

全面理性抉择模型，大多是在做出抉择之前，先评价全部备选方案。真实的人类决策，常常是依次顺序检查备选方案。我们可能知道，也可能不知道决定顺序的机制。在顺序检查备选方案的情况下，我们可以认为，评价过程中碰到的第一个满意方案，就是实际选择的方案。

一个棋手在比赛时如果发现了某个能将死对手的走法，那么，一般说来，他来不及考虑是否还有其他能将死对方的走法，便采用了已经想到的走法。在这种情况下，我们发现，要预见棋手究竟选取哪个方案是非常困难的。之所以如此，是因为我们没有一个可以依据的理论，去预见棋手对备选方案进行检验和评价的顺序。不过，对前面讲过的另一个例子（出卖房子）来说，由于卖主所面临的环境是顺序一定的备选方案，因此，所谓第一个满意方案的选取，有着明确的含义。

无论如何，我们应在这点上引进一些具有心理学基础的动态考虑。下面让我们用上述顺序情境，来代替那种一次性的静态抉择情境。在这种系列式的尝试过程中，满意方案的欲望水平，是可以发生变化的。这种变化的一个比较含糊的原则是：一个人如果在探索备选方案时感到容易发现满意方案，其欲望水平便提高；反之，如果感到发现满意方案是比较困难的，其欲望水平就下降。也许我们有可能用进一步了解 A 到 S 的映射的难易程度，或细化这一映射的任务的组合规模，来表示对备选方案进行探索的难易程度。这一过程的形式定义有好几种方法。

欲望水平的上述变化，有可能使我们得到满意解的“近似唯一性”（“准唯一性”）也可能保证满意解的存在性。因为，寻找答案如不成功，欲望水平就会下降，这样就可能找到满意解了。

## 解的存在性、进一步的可能

以上我们讨论了从长期上看，保证解的存在性的一个机制。我们还有另外一种可以表达该过程的办法。到目前为止，我们还没怎么用到  $A$  与  $\dot{A}$  之间的区别，前者是备择行为集合，后者是抉择者想到的备择行为集合。现在我们设后者是前者的真子集。抉择者如果在  $\dot{A}$  中没有找到一个满意方案，他就可能到  $A$  中与  $\dot{A}$  相邻的其他方案当中去寻找。<sup>①</sup> 这个程序无非是精心构成前面讲过的信息收集过程。（我们可将  $A$  中不在  $\dot{A}$  中的元素，看作是起初映射到整个集合  $S$  上的元素。）

对一个生物来说，一系列抉择中的动态适应性变化，可能主要取决于欲望水平的调整。对另一生物来说，其适应和调整可能主要在集合  $\dot{A}$  中：若易于找到满意方案， $\dot{A}$  就变小；反之，若难于发现满意方案， $\dot{A}$  就变大。生物越是固执，欲望水平的调整就越少，相对而言， $\dot{A}$  的调整就越重要。（当然，在向上（大）调整和向下（小）调整之间，可能确实存在着一种不对称性。）

如果报酬可用金钱或效用来衡量，如果发现备选方案的费用也同样是可以这样衡量的，那么，我们就能用完全有序性（以报酬及发现方案的费用的加权和来表示的有序性），去代替图2所示的局部有序性。这样一来，我们就可以讲行为持久性的最优程度——我们可以说，持久性强的生物更有理性，反之亦然。但是，本章

---

① 这里可以提一下，我曾依照朴素经验论的精神，将一个包含多重报酬因素的问题交给一些学生和朋友去考虑。那个问题中的报酬因素大大依赖于一个非常偶然的不确定事件。我发现，考虑该问题的人极不情愿把自己限制在该问题指定的备择行为范围里。他们不喜欢那个有很大好处或坏处，但其概率无法计算的备择行为，而是试图寻找那种报酬较小、但对偶发事件不太敏感的备择行为方案。这里的问题，正是莫地利安尼在其论文中所描写的“热狗摊”问题（见《美国经济评论（文献集）》，1949年，第39卷，第201—208页）。

的中心论点是：作为行为主体的生物，一般并不知道那些费用，它们也没有一组权数，用以权衡多重报酬的各个分量。恰恰因为有这些知识上和能力上的限度，我们所描述的不那么全面理性的模型，才是重要而有用的。在给定了那些限制的条件下，如何“理性地”行动的问题，不同于怎么使能力得到加强，以使行动更“理智”的问题。这里的“理性”、“理智”，是以那种最最全面的理性模型来判断的。<sup>①</sup>

当然，这两种观点并非完全不同，更不能说是对立的。我已指出，生物可以具备一个完整的理性机制等级体系——例如，欲望水平本身可受制于一个某种动态意义上的调整过程。此外，在许多情况下，我们感兴趣的可能是一个决策程序是否比另一程序更富理性这一明确问题；而且，为了回答这个问题，我们通常将不得不构造一个更广泛的、围绕着这两个程序的理性准则来作为近似。我们的完整观点是：讲清楚我们在这样一个模型等级体当中所考虑的是哪个层次，是很重要的；另外，在很多情况下，我们感兴趣的是“有限”理性的模型，而不是更“全面”理性的模型。

### 关于动态关系的进一步说明

至此为止的讨论，只是在一个很特别的意义上涉及到了动态性：T时刻的欲望水平，取决于系统的历史（从前的欲望水平和从前达到欲望的水平）。还有另外一种动态联系，可能也是非常重要的。这种动态联系就是：一次特殊尝试的报酬，可能不仅取决于那次尝试所选择的方案，而且取决于从前尝试所选择的方

<sup>①</sup> 还可加上：“或以其抉择机制的生存价值来判断。”

案。

这种情形的最直接的表述，是把整整一系列尝试中的报酬，看作是向量报酬函数的分量。但这样一来，最优化就要求从这一系列尝试的一开始，就选取一个对整个系列而言的总体策略（参见附录）。这样做，同样会使问题大大复杂化，超出生物能力所及的复杂程度。一个可能的折衷办法是：对每一次尝试，都定义一个二分量报酬函数。其中一个分量表示“直接”报酬（消费），另一分量则表示生物给未来尝试留下的“余地”（储蓄，灵活性）。

让我们考虑这样一盘国际象棋赛，双方每走10步就计算一次成绩。计算方法是：任意给棋盘上尚存的棋子指定价值（比如：王后1分，堡垒（车）10分，等等），然后按比例算出总分。这样，棋手就可以采用一个运筹期限（规划期限），将其在该期限内的走法“好坏”，包括在他的预计报酬当中。在经济学上，与此类似的概念是财产在计划期内的折旧值。要想精确计算这类价值，需要超出期限实际执行其策略才行。报酬若具有随时间久远而下降的性质（时间折现），那就会有一个好处，即：在估算这些折旧值时，可轻视误差。（对于确保报酬之和的收敛性来说，时间折现性质有时是必不可少的。）

我们容易想象出其他一些在实际当中颇为重要的动态复杂性。我们可以提一下其中两三个——但在此不做正式讨论。生物所经历的实际后果，可使报酬函数发生变化——不吃梨子，便说不出喜欢不喜欢梨子的味道。同样，将A到S上的映射予以精细化的方法之一，就是挑选一个特殊方案并体验其后果。在上述两种情况下，报酬同特殊方案相联系的要素之一，就是所收集到的关于映射的信息，或关于报酬函数的信息。

## 结 论

本章的目的是要构造这样一类“理性抉择”的定义，其模型

基础同迄今已经提出的种种定义相比，更接近于生物行为中的真实决策过程。上面我已就静态情形，概述了一个很完整的模型，并且描述了这一模型向动态情形的推广。正如上面最后一节指出的，在我们能够现实地建立一个更完整的动态体系之前，尚有大量工作需要完成。

本章导言部分指出，这类定义既有描述性价值，也有规范性的用处。尤其是它告诉我们，在看来远远超出了现存的或展望中的计算设备的能力的情况下，理性抉择是如何达到的。对计算机和人，不好进行智商比较。我们如果把这两者在综合智力测验中的得分因素逐项分开来看，无疑会发现，在有些方面，他们一个是天才，一个是呆子；而在另一些方面，则恰恰相反。对理性的可能定义所做的考察，可为我们指出设计和利用计算设备的方向：让计算机在目前仍处于低级、愚笨阶段的那些智力方面，也同样能取得相当好的智力分数。

不过，构造上述“近似”理性定义，有一个更加广泛的目的，那就是：它能为在组织环境下做决策的集体或个人的有关行为理论的建立，提供某些材料。现在我们面临着一个明显的佯谬：一方面，厂商经济理论和管理理论试图从行为至少具有“意向”理性的角度，来看待人类行为；另一方面，我们又可以证明，如果采取古典理论那种全面理性，企业组织或其他种类的组织的内部结构问题，很大程度上便消失了。<sup>①</sup>当我们用具有有限知识和能力的生物去取代经济人或管理人时，这个佯谬就不见了，行为理论的轮廓便出现了。生物为进行抉择而对真实世界进行简化，造成了简化模型与现实之间的差距；这些差距又可以为我们解释组织行为中的很多现象。

---

<sup>①</sup> 参阅西蒙：《管理行为》（第1版），第39—41, 60—64, 96—102, 240—244页。（此书有中译本，杨砾等译，北京经济学院出版社1988年版。参见第3、4、5、11章。——译注）

## 附录：关于如何合理决定可接受报酬的一个例子

本章正文部分提出了理性适应可在各种“层次”上起作用的概念。也就是，生物可在模型所假定的限度之内进行理性抉择，但它也可以合理地设置这些限度。前面讲过的出卖房宅的例子，可以说明这一点。

假设一个人要出卖一所房子。他每天（或其他时间单位）定一个接受价，比如，第 $k$ 天房价定为 $d(k)$ 。如果他在那天收到了一个或几个高于该价格的报价，他就接受其中的最高报价。如果他没有收到高于 $d(k)$ 的报价，他就不卖那所房子，等到次日再决定一个新的接受价 $d(k+1)$ 。

他如果对每天报价情况的概率分布有一些了解，那么，他就可以在下述意义上，决定一个最优的接受价，使出卖价的期望值 $V[d(k)]$ 最大。

为讲清楚这一点，我们做如下解释。设 $p_k(y)$ 是 $y$ 成为第 $k$ 日最高出价的概率，那么，房子将在第 $k$ 天卖出（若此前尚未卖出）的概率就是：

$$P_k(d) = \int_{d(k)}^{\infty} p_k(y) dy \quad (1)$$

卖主在第 $k$ 日出卖房子所得收入的期望值是：

$$\varepsilon_k(d) = \int_{d(k)}^{\infty} y p_k(y, k) dy \quad (2)$$

考虑到房子在第 $k$ 日之前卖出的概率，卖主在第 $k$ 日出卖房子所得收入的无条件期望值为：

$$E_k(d) = \varepsilon_k(d) \prod_{t=1}^{k-1} (1 - P_t(d)) \quad (3)$$

出卖价的期望值为：

$$V\{d(k)\} = \sum_{k=1}^{\infty} E_k(d) \quad (4)$$

下面我们要决定所有  $d(k)$ ，要求它们使 (4) 式取到最大值。函数  $d(k)$  的  $k$  个分量是彼此独立的。对每一分量，求  $V$  的偏导数，可得：

$$\frac{\partial V}{\partial d(i)} = \sum_{k=1}^{\infty} \frac{\partial E_k(d)}{\partial d(i)} \quad (i = 1, \dots, n)_k \quad (5)$$

但我们还有以下三式：

$$\frac{\partial E_i(d)}{\partial d(i)} = -\frac{\partial e_i(d)}{\partial d(i)} \prod_{j=1}^{i-1} (1 - P_j(d)) \quad (6)$$

$$\frac{\partial E_k(d)}{\partial d(i)} = e_k(d) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} (1 - P_j(d)) \left( -\frac{\partial P_i(d)}{\partial d(i)} \right) \quad \text{当 } i < k \quad (7)$$

$$\frac{\partial E_k(d)}{\partial d(i)} = 0 \quad \text{当 } i > k \quad (8)$$

因此，对于极大点来说，有

$$\begin{aligned} \frac{\partial V}{\partial d(i)} &= -d(i)p_i(d) \prod_{j=1}^{i-1} (1 - P_j(d)) \\ &\quad + \sum_{k=i+1}^{\infty} e_k(d) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} (1 - P_j(d)) p_i(d) \\ &= 0 \end{aligned} \quad (9)$$

消去  $p_i(d)$ ，我们最终得到：

$$d(i) = \frac{\sum_{k=i+1}^{\infty} e_k(d) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} (1 - P_j(d))}{\prod_{j=1}^{i-1} (1 - P_j(d))}$$

$$= \sum_{k=i+1}^{\infty} \varepsilon_k(d) \prod_{j=i+1}^{k-1} (1 - P_j(d)) \quad (10)$$

要想让这个式子有意义，式中的无穷和必须收敛。从  $\varepsilon_k(d)$  的定义式（2），我们可以看出，如果出价（报价）的概率分布随时间推移而充分迅速地下降，就可以保证和式收敛。这种变化趋势可能对应于：（1）预料价格跌落，或（2）把  $y$  解释为未来价格以足够高的利率予以折现的现值。

另外，我们还可以通过如下假定，避免收敛性问题。这个假定是：第  $n$  天的保留价  $a(n)$  很低，足以使  $P_n(d)$  为一。我们将采用这后一种办法，但在继续讲下去之前，我想解释一下方程（10）。这个式子告诉我们，第  $i$  天的合理接受价  $d(i)$ ，等于第  $i$  天房子尚未卖掉情况下的卖价期望值，加上后续日子里最佳决定的接受价。看一看（10）式右边恰好与（4）式右边相同，但和号起点不是  $k=1$ ，而是  $k=i+1$ ，就会明白这一点。<sup>①</sup>

因此，在求和至第  $n$  日终止的情况下，也就是在确定的第  $n$  天卖掉了房子，而在以前没有卖出，我们就可以采用从终止期倒推的办法算出最优的  $d(i)$ ，不必同时解算方程（10）。

卖主除了需要了解接受价确定前后所需的信息之外，还需要知道为确定合理接受价所需的其他信息，例如，他实际上还需要知道后续时日买主报价的概率分布，知道有关这一分布的完完全全的信息。观察到这一点，是很有意义的。

不具备这类信息的卖主，满足于粗略合理性的卖主，将回避他所不具备的信息，采用近似的模型。首先，他可能假设一个他肯定能在第  $n$  天卖出房子、并且希望到那时能出手的价格，以此限制其计划期。其次，他一开始将接受价定得很高，观察他所收到的报价的初始分布情况，并逐渐地、大致地向上或向下调整其

<sup>①</sup> 方程（10）看来已被达林和金克艾德独立地发现了。参见他们在《美国运筹学学会杂志》（I, 80 (1953年2月) 上发表的文章：《一个库存问题》。

接受价，直到他收到一个可接受的报价为止——他根本不进行概率计算。我认为，这就是人类寻找“足够好”方案的那种理性适应，也是能在广泛实践领域里实际进行的理性适应。

### 参 考 文 献

- 〔1〕阿罗：《构造风险抉择理论的几条替代途径》。
- 〔2〕达林，金克艾德：《一个库存问题》。
- 〔3〕莫地利安尼：《关于流动性和不确定性的讨论》。
- 〔4〕西蒙：《管理行为》。
- 〔5〕西蒙：《组织理论的比较》。
- 〔6〕西蒙：《雇佣关系的形式理论》。  
（文献出处见本文脚注）。

### 三、理性抉择与环境结构<sup>①②</sup>

爱德华在《心理学会刊》上发表的文章(1954年)和圣莫尼卡大会文集《决策过程》(斯拉尔等编, 1954年)等近期出版物, 表明了心理学界对决策制定过程日益增长的兴趣。这方面的研究工作, 对理性抉择的特征描述, 给予了很多注意。由于这一课题一直是经济学的一个主要关切点, 因此, 决策理论已成为心理学理论和经济理论的一个自然的交接地。

对心理学界所采用的适应性行为模型(如学习理论)和经济学界所采用的理性行为模型进行对比考察, 我们就可以看到, 几乎在所有方面, 后者均假设了比前者复杂得多的抉择机制, 后者所假定的生物获取信息和进行计算的能力也比前者大得多。此外, 在对依据这两种理论所做的预测做了比较的有限范围里(参见文末“参考文献”第〔7〕篇的第9、10、18章), 学习理论对所观察到的行为所做的解释, 看来要比理性行为理论的解释更好些。

无论是从这些很少见的资料上看, 还是从对经济模型假设的检验上看, 似乎都可以得出这样的结论, 即生物在学习和抉择情境中的行为尽管是适应性的, 但这种适应性极其缺少经济理论所

---

① 本文原载《心理学评论》63, 第2期(1956); 第129—138页。另载西蒙的《思维模型》和《有限理性模型》(下卷)。本文是上篇论文的姊妹篇。  
——译注

② 阿兰·纽韦尔曾就本文论题同我进行过多次富有启发的交谈; 福特基金会的赞助使我得以完成本文, 在此一并表示感谢。

假定的那种“最大化”理想性质。显而易见，生物有足够的本领进行“寻求满意”的适应；但它们一般并不“寻求最优”。

如果确实如此，我们就可以通过进一步的考察，大大加深和扩展我们对理性决策的理解。这种考察，从一开始就考虑到生物的能力限度和复杂性方面的局限，并且考虑到生物所必须适应的环境有着允许其抉择机制进一步简化的性质。因此，提出这样一个问题是有趣的，即，我们可以假定一组多么简单的抉择机制，而仍能得到所观察到的适应性抉择行为的大致特征？

在上一章里，我已经就拥有有限信息和有限计算能力的生物所可能采用的“近似”理性，提出了一些见解。在关于人类决策过程的明确知识尚属缺乏的情况下，我们现在只能根据日常经验，根据我们的反思，以及心理学关于这些过程是什么样子的很有限的文献，去进行猜测。在这个意义上，我那些见解是“假设性的”。但是，尽管它们具有试探性，却是关于人类抉择和其他生物抉择中所含的某些实际机制的，旨在经验研究的见解。<sup>①</sup>

一个生物如果面临着在某一特殊环境中近似合理地（或适应性地）行动的问题，那么，适当的简化可能不仅取决于该生物的特征——如感觉特征、神经特征等，而且同样取决于环境的结构。因此，我们希望通过认真考察环境的某些基本结构特征来发现某些线索，以揭示决策过程中所采用简化机制的性质。这就是本章的论证路线。

下面我们最先讨论的那种环境，与其说是人的环境，也许不

---

① 自打撰写该文以来，我发现其中若干假设符合德·格鲁特有关棋手思考过程的有趣而重要的研究结果。我打算另行讨论这些实验结果对我的模型有何意义，〔在《思维模型》一书里，此脚注最后一句改为：“我在《求解难题过程中的试误搜索》和《一年将死组合程序》（与G. Taylor合作）两文中，讨论了这些实验结果对我的模型有何意义。”该两文载《思维模型》第175—180页和第181—200页。—译注〕

如说是兔子的环境更合适。环境一词是模糊的。我们对于客观物质世界的总体面貌不感兴趣，只关心生物认为是“生存空间”的那些方面。因此，我们所谓“环境”，将依赖于生物的“需要”、“动机”或“目标”，依赖于生物的感知器官。

## 生物的环境

我们首先考虑一种简单化的生物（也许应称为“头脑简单的”生物），它只有一种需要——吃东西，并能进行三种活动：休息、寻觅和进食。关于这三种活动的确切性质，稍后予以解释。我们可以把该生物的生存空间，描述为它能在其上行走的一张曲面。这张曲面的大部分地方完全是空白，但有很多非常分散的孤立点（食物堆），那里各有少许食物，仅够一餐之用。

该生物的眼力，使它在任何时候都能看清周围一个圆子里的景物。它能以某一固定的最大速度，在曲面上行走、移动。它以一定的平均速度进行新陈代谢，并能贮存一定数量的食物能量。因而，它需要在一定的平均时间间隔内吃东西。它一旦发现了食物，就能以最大行走速度奔过去。这里的理性抉择问题是：要选择行走路线，以保证不遭饥饿之苦。

我认为，这个生物的一种合理行为方式是：（1）随机地探索、觅食；（2）一发现食物，便奔过去吃；（3）如果它在每一餐的觅食和进食所需的平均时间里消耗的总能量，少于一餐所消费食物的能量，它就把剩下的时间用于休息。①

① 熟悉格雷·沃尔特的机械龟的读者会看到，正如我们所讲的，这里对我们这个生物的描述完全可以用作一组设计要求，以保证他那个机械龟在并不总用充电器的环境中生存。由于我在构造这一模型时不熟悉机械龟的结构，两者的行为存在某些差别——但类同性还是很突出的。参阅文末所列第⑧篇文献。

对理性抉择所做的这一描述，除了同经济学家及其他一些人所提出的更深奥的人类理性模型迥异之外，没有任何特别突出之处。让我们看一看，这一个生物和使其抉择如此简单化的环境究竟是怎么一回事。

1. 该生物仅有一个目标：吃东西。它不需要权衡不同目标的相对重要性。它根本不需要用“效用函数”或“无差异曲线”去挑选替代方案。

2. 它没有任何最优化问题。它只需要保持一定的平均进食速度，更多的食物形同泥土，毫无用处。用心理学家的话说，它有一个明确的、固定不变的欲望水平，而且，它的成败不改变其欲望水平。

3. 它的感知和环境的性质，严格限制了它的规划、谋算期限。由于食物堆是随机分布的，故而其搜索活动不需要任何模式。它一旦发现了一堆食物，就能沿一条确定的“最佳”路线奔过去。

4. 它的需要和环境的性质，造成了“手段”和“目的”之间的很自然的分离。除了食物堆之外，曲面上的这一点同那一点没有差别。移动只有成为觅食手段，才具有意义。①

我们将会看到，第1点不是必不可少的。只要欲望水平固定下来，规划期限便受到了限制；而且，手段与目的之间存在鲜明差别，多目标的存在不会给抉择带来任何真正的困难。只有后三个条件的放宽，才会造成真正的困难。为了搞清楚这当中包含着什么，我们必须更精细一点地将这个模式形式化。

---

① 经济合理性模型有一个特征，即：“手段”和“目的”之间的区别在那种模型中毫无实质作用。我们不可将这个区别，同备择行为与效用之间的区别混为一谈；其理由已在我的《管理行为》第4、6章中比较细致地讲过了。

## 感知力，贮存能力与生存

为方便起见，我们可以不把该生物的生存空间说成是一张连续曲线，而是将其描绘为类似迷宫那样的多通路系统，其中每一分支点表示一个抉择点。我们将一条支路的选择和向下一分支点的移动，称作一“程”。全部分支点当中，只有一小部分是有食物的（食物堆）。

设  $p$  是那些分支点当中随机分布的食物堆的百分比 ( $0 < p < 1$ )，设  $d$  为在一个分支点处交汇的（穿过同一分支点的）路径平均数。设  $v$  为生物可以看到的前方“程”数。也就是，如果一个有食物的分支点距该生物当时所处位置不超过  $v$  程，该生物就能正确地选取一条路径奔过去。最后，我们设  $H$  为生物在两餐之间，在不饿着的情况下所能行走的最大程数。

该生物在任何给定时刻所能看到的分支点数目，距其当前位置一程之远的有  $d$  个，距其当前位置两程之远的有  $d^2$  个；一般而言，距其当前位置  $k$  步之远的有  $d^k$  个。合起来看，它眼力所及的分支点共有  $d + d^2 + \dots + d^v = d(d^v - 1) / (d - 1)$  个。当它选择一条支路并行走一程时，就有  $d^v$  个新的分支点进入视野范围之内。因此，在  $m$  程行走过程里，就会出现  $md^v$  个新的分支点。由于它最多可行走  $H$  程，而且只能获取  $v$  程以内的食物，因此，它不能生存下去的概率  $Q = 1 - p$ ，等于它在  $(H - v)$  程中未能找到食物的概率。（ $p$  如果很小，我们可不计在其第一程规划期限内可见到食物的概率。）设  $p$  是在某一程的终点处所看到的  $d^v$  个新分支点均非食物堆之概率。

$$p = (1 - p)^{d^v} \quad (1)$$

因此有：

$$1 - P = Q = p^{(H-v)} = (1 - p)^{(H-v)d^v} \quad (2)$$

我们看到，这个简单动物一餐跟一餐的生存机会取决于 4 个

参数。其中两个描述生物本身，另外两个描述环境： $p$ 指环境中的食物丰富程度； $d$ 为环境中的路径（支路）丰富程度， $H$ 表示生物的能量贮存能力， $v$ 表示生物的眼界远近。

为了使读者对以上各量的大小有些印象，我们设  $p = 1/10\,000$ ,  $(H - v) = 100$ ,  $d = 10$ ,  $v = 3$ 。这样一来，走一程之后看到一个新的食物堆的概率为  $1 - p = (1 - p)^{100} \approx 880/10\,000$ ，生存概率为  $P = 1 - p^{100} \approx 9999/10\,000$ 。因此，在这种情况下，该生物在其生命最后限度之时仍找不到食物的概率的情形，10 000 次中仅有一次而已。现在我们设生存时间  $H - v$  增加三分之一，即从 100 增加到 133。那么，同样的计算表明，饥饿概率将降至十万分之一。如果是  $v$  提高三分之一，当然会有更大的效应，即：将饥饿概率从  $1/10^4$  降至  $1/10^{10}$ 。

仍采用  $p = 0.0001$  和  $(H - v) = 100$ ，我们可计算出该生物在完全随机行动情况下的生存概率。在这种情况下， $P' = [1 - (1 - p)^{100}] = 0.009$ 。由此可见，该生物在较短计划期进行有目的行动的不太强的能力，就能使它容易在随机行为可能导致迅速灭绝的环境中生存下来。简单计算表明，它的感知力乘上 830，是它发现食物的平均速度。

如果给定了  $p$ 、 $d$  和  $v$ ，并且规定生存概率必须不小于某个接近 1 的正数 ( $P \geq 1 - \epsilon$ <sup>①</sup>)，那么，我们可以从 (2) 式算出其相应的最小  $H$  值：

$$\log(1 - P) = (H - v) \log p \quad (3)$$

$$H \geq v + \frac{\log \epsilon}{\log p} \quad (4)$$

例如，若  $p = 0.95$ ,  $\epsilon = 10^{-10}$ ，则  $\log p = -0.022$ ,  $\log \epsilon = -10$ ,  $(H - v) \geq 455$ 。参数  $H$  可解释为该生物的“贮存能力”。也就是说，若该生物以每程  $a$  个单位的速度进行新陈代谢，那

①  $1 > \epsilon > 0$  ——译注

么， $\alpha$ 个单位的食物能量存贮（其中H由（4）式给定）是它在确定的风险水平情况下要获得生存所必需的。

考虑发现食物所需的平均程数M，可使我们更明白H的含义。由（1）式可知，走 $(k-1)$ 程未发现食物、且在第k程发现了食物的概率为：

$$P_k = (1-p) p^{(k-1)} \quad (5)$$

因此，发现食物所需的平均程数M为：

$$\begin{aligned} M &= \sum_{k=1}^{\infty} k (1-p) p^{k-1} \\ &= \frac{(1-p)}{(1-p)^2} = \frac{1}{(1-p)} \end{aligned} \quad (6)$$

由于 $(1-p)$ 是在任意一程发现食物的概率，故而M是此概率的倒数。结合（3）式和（6）式可得：

$$\frac{M}{H-v} = \frac{\log p}{(1-p)} \cdot \frac{1}{\log(1-P)} \quad (7)$$

由于p接近于1， $\log_p p \approx (1-p)$ ，（故7）式可大致简化为①：

$$\frac{M}{H-v} \approx \frac{1}{\log_e(1-P)} \quad (8)$$

例如，我们如果要求 $(1-P) = e \leq 10^{-4}$ （万分之一的饥饿可能），则 $M/(H-v) \leq 0.11$ 。对这一生存水平，我们要求食物能量存贮大致等于 $\alpha(v + 9M)$ ——足以在发现食物所需平均时间的9倍时间，加上取得食物所需时间的总时间之内，维持其生命的食量。②

①  $\log_{e/p} \ln$ ，自然对数。——译注

② 我没有找到有关动物存食能力方面的很满意的资料，但上面提出的平均搜索时间与存食能力之比，数量级上肯定正确。读者可能注意到，至少在有些情况下，“食物”是普遍存在的。因此搜索（觅食）时间可忽略不计，存食能力也很小。于是，对陆地动物来说，氧的存食能量没有多少，在无空气的条件下，生命只能维持几分钟。我这里并不是在主张因果之矢的指向，仅仅是在讲生物在这方面适应其环境，而且没有过多的存食能力。

## 多目标抉择机制

现在我们考虑一种更为复杂的生物，它能寻求两个以上的目  
标，并对之做出反应。我们在作这种考察时，可在抉择过程中任  
意引进复杂性；但是，有趣的问题在于，如何在最低的复杂程度  
上引进多重目标——也就是要构造出这样一种生物，它能以比较  
初等的抉择机制，去处理多目标决策问题。

两个目标的存在，最起码也要引进一致性要求——为达到一个  
目标而花费的时间，限制了为达到另一目标所能用的时间。不过，  
对于我们所考虑的那类环境来说，不需要虑及两个目标之间的  
其他关系了。在我们最初写出的公式当中，不可缺少的规定只是：  
存贮能力H要足以使饥饿风险保持在所规定的(1-P)以下。  
现在我们再做一个新规定：该生物应当仅用其总时间的λ倍  
(0<λ<1)进行觅食活动，剩余部分(1-λ)于其他事。这个  
新规定要求存贮能力增加。

为了防止饥饿风险，该生物每当达到了能维持H期的食物存  
贮水平，就必须开始寻找食物。它如果有了能维持( $\mu + H$ )期的  
食物存贮水平，而食物堆的大小至少是 $\alpha(\mu + H)$ ，那么，  
它需要在最后一餐的μ期之后，立即开始寻找食物。另一方面，  
觅食时间平均为M期。因此，如果已经有了一个饥饿感临界值，  
使该生物每当餐后μ期开始搜索食物，我们就有如下公式：

$$\lambda = \frac{M}{M + \mu} \quad (9)$$

因此，我们如果让μ充分大，就可以使λ任意小。附带说明一  
下，有人可能注意到，我们这里所讲的系统，同工业库存控制中  
十分常见的两箱系统是极其相似的。基本库存H是为了满足收到  
新订货之前的需求而建立的缓冲贮备(缺货风险为1-P)，附加

库存 $\mu$ 确定了“订货点”， $\mu + M$ 是平均订货量。库存量 $\mu$ 的固定，是为了使存贮“费用”同订货次数过多的费用相抵（后者在我们这里相当于时间压力）。

如果食物和另一目标（比如说水源）是随机分布的，那么，两种活动的交叉不会带来任何根本性的复杂化。让我们以下标1表示有关获取食物的变量和参量（如： $\mu_1$ 表示寻找食物的时间临界值），用下标2表示寻找水源的变量和参量。只要 $\lambda_1 + \lambda_2 < 1$ ，该生物就足以进行那两种活动了。

当该生物的饥饿感或口渴感达到了临界值，它就开始寻觅、搜索。我们假定，当它看到了食物和水源两者之一，它将奔向那个目标，去充饥或解渴（这不会使它为达到另一目标所需的平均程数有所增加）；但如果它同时发现了两者，且设 $S_1$ 和 $S_2$ 分别为此时体内食物和水的存贮量，那么，当 $M_1/S_1$ 大于 $M_2/S_2$ 时，它去找食物；当 $M_1/S_1$ 小于 $M_2/S_2$ 时，它去找水。这样的抉择，会使其生存机会最大。因此，所需要的是一个优先选择机制，它能产生正比于 $M_i/S_i$ 的驱动力（动机）。

这种优先选择机制，并不是我们可以假设的唯一机制，也不是最简单的机制。对我们所描绘的那种生物来说，一个更简单的法则是：坚持搜索这样一些目标，它们能满足首先达到临界值并首先启动搜索行为的那种特殊需要。从节省达到目标的时间上看，这种程序固然不是经常有效，但对于存贮能力强的生物来说，这可能是完全合适的。

我们看到，一个生物无需很精致的抉择机制，便可以满足其数种不同需要。特别是，我们不必假定一个效用函数或“边际替代率”。

我们甚至还可以更进一步断言，为满足不同需要而进行的活动的相互依赖性，有可能使生物得到一个重要的好处——节省搜索时间。这种好处如果存在，也可以通过一个初等的抉择机制稳妥地获得。设该生物有 $n$ 种需要，可满足其各种需要的各个地点在

环境中的分布，是独立随机分布，且概率均为 $p$ 。这样一来，某一个程所发现的地点都不满足任何需要的概率为 $p^n$ ，发现第一个满足需要的点所需的平均程数为：

$$m_n = \frac{1}{(1-p^n)} \quad (10)$$

假设该生物开始搜索，先奔向它所发现的第一个满足需要的点；然后再搜索，奔向它所发现的另一个满足需要的点；如此等等。那么，寻找全部n个目标所需的平均总程数（或时间）是：

$$M_n = m_1 + m_{2-1} + \dots = \sum_{i=1}^n \frac{1}{(1-p^i)} \ll \frac{n}{(1-p)} \quad (11)$$

尤其是如果 $p$ 接近1，即满足需要的点很稀少，我们将会有下列式子：

$$\begin{aligned} M_n - M_{n-1} &= \frac{1}{(1-p^n)} \\ &= \frac{1}{1-p} \cdot \frac{1}{\sum_{i=1}^n p^i} \simeq \frac{M_1}{n} \end{aligned} \quad (12)$$

$$M_n \simeq M_1 \sum_{i=1}^n \frac{1}{i} \quad (13)$$

在(13)式中以具体值代换n，可得到： $M_2 = \frac{3}{2} M_1$ ； $M_3 = \frac{11}{6} M_1$ ； $M_4 = \frac{25}{12} M_1$ ，等等。由此可见，如果该生物有两种孤立的需要，则其搜索时间只比仅有一种需要的情形增加50%，而不是增加一倍。四种需要的搜索时间仅为单一需要搜索时间的两倍略强，等等，对上面刚刚讲过的程序稍加检验可知，多种需

要搜索过程不减少该生物所需的基本贮存能力，但确实减少了所需的附加贮存能力。事实上，可能根本不需要附加贮存。

只有在满足需要的点是独立分布的条件下，这个结论才成立。如果那些点的联合分布存在负相关性，那么对该生物来说，分别满足各种需要的做法可能是经济的，从而要有一个包括附加贮存在内的简单指示机制，以启动其多种搜索的动机。这一点将在下一节里阐明。

这里还可以讲一下“躲避需要”的问题。设该生物的行为空间中有某些点被指定为“危险”点。它需要躲开经过那些点的路径。如果危险点占全部点的比例为 $r$ ，随机分布，那么，一程当中可以看到的可行路径便减少为 $(1-r)^d^v$ 条。因此， $\rho' = (1-p)^{(1-r)d^v}$  将小于 $\rho$ （(1)式）， $M$ （(6)式）将相应变大。由此可见，危险点的存在无非是增加了平均搜索时间，从而要求贮存能力提高。

### 关于环境的进一步说明：线索

在以上讨论中，生物预见未来的范围被它每一程的可行备择行为数目( $d$ )和“眼力”( $v$ )限制住了。我们很容易给这个模型引进几种学习效应。备择行为的增多和眼力的加强，可分别以 $d$ 和 $v$ 的变化表示之。

如果有食物的地点（食物堆）不是完全随机分布的，而且存在一些线索（信号），暗示各条路径是否能通向食物多（或少）的地方，那就会出现一种比较有意思的可能。首先让我们设想，在通向任一食物堆的路径上，有 $k$ 个抉择点标有食物线索。该生物一旦学会了发现这种线索同随后发现食物之间的联系，它的搜索便不必一直进行到发现食物为止，发现线索就成为终点。它只要发现了线索，即可从该点出发沿一条确定路线奔过去。这就等于是用 $v' = (v + k)$ 代替了 $v$ 。

另外还有一种不同的线索，可按如下方式起作用：设每一抉择点均各具独特特征，该特征同由该点出发随机挑选一条路径能找到食物的概率相联系，这样，该生物就可以根据这种线索，在每一程都选取概率最大的点。如果仅有部分抉择点上有这类线索，那就可以采用随机搜索与系统搜索的组合方法。这样一来，该生物就可以找到那些达到目标的概率比较高的“区域”，不过，它可能还要在那种区域内随机搜索食物。

对人类来讲，这种行为的一个具体例子，是棋赛第一阶段的“位置竞争”特征。棋手以欲达到位置的某些特征为根据（如：棋子的活动范围），选择走法。有些位置被认为比其他位置更富攻击性或防守性，但最初抉择可能根本不包含这种明确的谋算，不包含达到了“好”位置之后为后续行动而进行的明确谋划。

下面我们考虑另一个抉择问题——当行为空间那些区域中满足第一种需要的点较多( $p_1$ 较大)而满足第二种需要的点较少(同一区域中 $p_2$ 较小)时的抉择问题。在前面讲过的目标冲突情况下(能满足不同需要的两个点或更多的点同时可见)，我们设想了一种优先选择机制：先计算需要的相对强度( $M_i/S_i$ )，然后对较强的需要做出反应。在有线索的环境中，由于存在学习过程，因此可能需要有一种条件机制，一方面用它把优先选择机制同对各种线索的反应联系起来，另一方面又将优先选择机制同各种可知的需要联系起来。

最后需要指出的是，在我们前面所讲的环境中，通向每一点的路径被规定为只有一条。从形式上讲，通过把可经多条路段达到的任何一点，表示为好几个点，我们总可以使这一条件得到满足。有时我们可能希望指定这样一个环境，其中的路径既可以是汇聚式的，也可以是幅射式的。要做到这一点，我们不必废除前面分析所采用的实质性条件。就我们所描述的那种行为而言，我们仅仅要求环境符合如下条件：

1. 如果完全随机地选择路径，生存概率小得可以忽略。

2. 环境中存在线索(或是指满足需要的点实际可见, 或是指预料性的信号), 生物可以靠这些线索, 确定地或从很高概率, 选择能通向满足需要处的特定路径, 获得足够大的生存机会。

### 关于多目标的最后评注

本章的主要问题是构造一种简单的抉择机制, 让我们所描述的那种生物能靠这种机制, 在多目标情形下做出行动。由于该生物如同真实世界中的生物一样, 既没有发现“最优”路径的意识, 也没有做出那种发现的才智(——即便假定最优概念有明确规定), 所以, 我们只关心寻找这么一种抉择机制, 它将能给该生物指出一条“满意”的路径, 一条能使其所有需要达到某一满意水平的路径。

我们为做到这一点而引进的假设, 有些表明了该生物的特征: (1) 它所能谋划的有目的行动系列是很短的(不超过 $v$ ), 不能超越限度。(2) 它的需要不是没边没沿的, 因而不必对边际满意增量进行权衡。如果它的全部需要都满足了, 它无非是闲散无事了。(3) 它拥有足够大的贮存能力, 因此, 任何一种需要得到满足的确切时间, 都不是关键性的。

我们还引进了另外一些有关环境特征的假设, 其中最重要的是: 需要的满足只可能发生在很“稀少”的、随机分布的地点(我们已经指出了这方面的某些数量特征)。

我们所得到的最重要的结论是: 该生物可以将其时间分配于与需要相关的活动(独立的手段-目的链的环节), 根本不产生任何总体配置问题, 或要求使用一般“效用函数”的问题。在这种情况下, 仅有的稀缺资源就是时间, 其稀缺程度用该生物需要进行某项活动的时间占总时间的比例, 予以衡量。通过贮存能力的增强, 可以减轻这种稀缺性。

当然, 这并不意味着我们不能从满足该生物需要所用总时间

的角度，构造一个更有效的程序。我们已经探讨了通过承认各项活动（特别是搜索活动）之间的互补性来提高效率的某些简单可能性。不过，这类复杂化的作法对生物生存来说，不是必不可少的。此外，如果环境中满足一种需要的可能性较大的区域，满足另一种需要的可能性很小（事实上经常如此），那么，这种效率可能无法利用。

有人可能提出异议说，即使是比较简单的生物，其行为看来也遵循效率准则，因此其抉择机制比我们这里所描述的，要精巧得多。例如，鼠能学会在觅食时取近舍远。但是，这种现象不影响我们的中心论点。我们可以引进这样一种机制，它能在存在多条路径均可满足一种需要的情况下，引导生物选择一条省时的路径。我们不必假定一种能在不同需要之间分配时间的机制。前一机制无非是增加了整个系统中的“闲散”时间，甚至更可以在规划该生物总体行为时忽略其活动的互补性。

这里不是详细讨论本模型对人类行为有何意义的地方，但我们可以依次做一点一般说明。首先，我们把生物的需要、感知能力和谋划能力，看成了固定的，因此整个分析都是静态分析。除了少数几处评论之外，我们没有考虑生物如何扩展其需要，或如何学会满足需要。从有关欲望水平的一般观察和实验中可见，对人来说，满足需要所用的时间同可以利用的总时间的平衡，是通过提高或降低欲望水平而达到的。我已经在前一篇论文里更详细地论述了这一点。<sup>①</sup>

其次，我们这个模型丝毫没有说，需要是生理学上的或天生的，而非社会学上的或后天所学会的。只要对各种老练的行为，生物的需要在任意给定的时间都可以用欲望水平来予以说明，这个模型就是适用的。

---

① 参见本书上一篇论文。关于表明了鼠的欲望水平调整的一个实验，及其可观察性，参见弗斯丁格1953年所撰一文。

这个模型的实证意义主要在于：对于那种假定人类或其他生物具备能在多种需要中进行选择的精巧机制的观点，我们应持怀疑态度。多种需要之间，可能根本不存在共同衡量尺度，或只是以很原始的形式存在这种尺度；而且，生物的需要对环境而言的本质，完全允许其不存在。

在关于冲突和挫折的业已完成的实验研究工作中，有一些与此相关的积极证据。在实验室里，造成冲突的常见方法是使动物处于这样一种情境：（1）刺激它同时考虑多种可以替代的有目的的行为，或者是（2）刺激它注意一种有目的行为，但限制它实施那些类似情境中往往很明显的行为。这说明，（至少在一大类情境中）冲突是由于给动物呈现了一些它不“习惯”处理的情境而产生的。人们经常认为，抉择冲突等同于缺乏给定情境下的抉择机制。尽管容易在实验室里构造这种情境，但是，缺乏对付这些情境的机制可能无非意味着，动物在自然环境中很少碰到类似的情境。①

## 结 论

在本章中，我试图辨别生物“心理”环境的典型结构特征。我们已经看到，处在具有这些特征的环境中的生物，只需要有非常简单的感知、抉择机制，便能够满足其多种需要，确保长期生存的高度可能性。尤其是，根本不必给生物假定一个“效用函数”，也无需用于计算不同需要之间边际替代率的任何精致程序。

这里所提出的分析，对当今流行的关于理性行为的经济理论和统计理论作为解释人类和其他生物理性特征的依据的有用性，

① 例如，可参阅尼尔·米勒的《关于冲突的实验研究》，载亨特所编一书（1944年）第14章。

提出了深刻的疑问。我们的分析指出了描述理性行为的一条替代途径，它更接近于有关感知和认知的心理学理论，更接近于人们在实验室和野外观察到的真实行为。

### 参 考 文 献

- (1) W·爱德华：《决策理论》，载《心理学会刊》，1954年，第51期，第380—417页。
- (2) 弗斯丁格：《鼠的有差异食欲的形成》，载《实验心理学杂志》，1953年，第32期，第226—234页。
- (3) 德·格鲁特《下棋过程中的思维与抉择》，荷文版1946年。英文版（海牙：莫顿出版公司）1965年，1978年（第2版）。
- (4) J·亨特：《个性与行为失调》，纽约：罗纳德出版公司，1944年。
- (5) 西蒙：《管理行为》。
- (6) 西蒙：《理性抉择的行为模型》。
- (7) 斯拉尔等编：《决策过程》，纽约：威利父子出版公司，1954年。
- (8) 沃尔特：《活脑》，纽约：诺顿出版公司，1953年。

## 四 有限理性论<sup>①</sup>

### 引言

理性指的是一种行为方式，是指在给定条件和约束的限度内适于达到给定目标的行为方式。关于理性行为的理论，既可以是规范理论，也可以是实证理论——也就是说，它们既可以告诫人们或组织，要想在一定条件下达到一定目标，应当如何去做；也可以描绘人或组织实际上是如何做的。本文研究关于理性行为的理论（合理行为理论）的结构，无论那些理论的本意是规范性的，还是描述性的。

**个人理性与组织理性。**关于理性行为的理论，既可以研究个人理性，也可以研究组织理性。实际上，这两类理论并非完全不同。<sup>②</sup>它们之间有一种貌似有理的区别，那就是：关于组织理性的理论必须研究目标冲突现象，而关于个人理性的理论则不必如此。但是，这样说仅仅对了一半，因为目标冲突现象对个人行为来说，如同它对群体行为一样重要——它是心理学界所谓“心理不和谐理论”的一个重大论题。（参阅N.P.查潘尼斯和J.A.

① 本文是为雅各布·马斯查克纪念文集《决策与组织》（C.B.拉德纳和R.拉德纳编，阿姆斯特丹：北荷出版公司，1972年版）撰写的专稿，载该书161—176页。另载《有限理性模型》下卷，第408—424页。——译注

② 这一点是由马斯查克在其关于组队论方面的第一篇论文《组队论初步》（1955年）中指出的。我将采用他的这一合理倡导。

查潘尼斯1964年。) 个人行为理论，如果微观到了涉及中枢神经系统内部组织(神经组织或功能组织)的地步，那它必定会有突出论述组织现象的部分。组织也是一种关于“个体行为理论，如果宏观到了把组织视为整块石料的地步，那它必体”的理论。尽管本章主要目的在于刻画和理解个人理性，但我将毫不犹豫地利用厂商理论——用古典语言说，就是关于单个企业家的理论，作为一种方便的和富有启发意义的例子。

因此，从本章的观点看，个人与组织之间的区别不很重要。就本章论证而言，对合理行为理论进行分类的一个更重要的判据，在于它们对“给定条件”——既定的目标和既定的状况——所做的假定。一类理论把所有条件和约束都归于理性活动者体外的环境；另一类理论则认定，作为信息处理者的活动者本身的局限性，就是非常重要的约束。这两类理论的区别尤其重要。我们可以把那类考虑到活动者信息处理能力限度的理论，称为**有限理性论**。

**古典厂商理论中的理性。**最简单形式下的古典厂商理论，为我们比较和区别有关理性的理论，提供了一个有益的标准。在厂商理论中，给定的目标是谋取最大利润。这里的利润定义为销售总收入与生产成本之差额。给定条件有两个：

1. **需求函数：**需求量是价格的函数：

$$q_d = D(p), \text{ 或 } p = D^{-1}(q_d) \quad (1)$$

由于总收入等于价格乘以销售量，因此需求函数确定了总收入：

$$R = p q_d \quad (2)$$

2. **成本函数：**生产成本是生产量的函数：

$$C = C(q)$$
 (3)

如果产量等于需求量，

$$q_s = q_d$$
 (4)

则欲求最大的利润无非是总收入与生产成本之差：

$$\text{利润} = R - C \approx pq - C(q) \quad (5)$$

在考慮到可微分性的适当假定下，最大利润将满足如下条件：

$$\frac{d(R - C)}{dq} = p + q \frac{d(D^{-1}(q))}{dq} - \frac{dC(q)}{dq} = 0 \quad (6)$$

这个理论中的约束——需求函数D和成本函数C，均被放在活动者的环境当中了。活动者被认为要解算方程(6)。为此他必须具备有关那些约束的完备知识，必须有能力进行必需的计算——将利润对产量的导数置零，并求解所得到的代数方程。

理性限度。以多种方式改变这些假设，就可以构造出有限理性论。我们可以在需求函数、成本函数或它们两者当中，引进风险和不确定性。例如，我们可以假定它们两者之一或它们两者的某些参数是具有已知分布的随机变量。这样一来，关于活动者完全了解这些函数的假定，就代之以活动者完全了解其分布的假定了。假定方面的这种变化，会使求解最优量的计算变得容易些或困难些——但通常情况下，要比确定性情形下的相应计算困难得多。

给理性加上限制的另一种方式是，假定活动者仅仅具备有关备选方案的不完全信息。针对这种情况所构造的模型，少于假定活动者不完全了解后果的那类模型。不过，某些搜索模型 (search models)<sup>①</sup> 假定，活动者知道利润在很多可能的备选方案当中的概率分布。活动者有具体方案可以采用——比如说可从很多可

① 搜索指寻找和挑选备选方案。搜索模型多见诸“信息经济学”文献。

——译注

能方案中随机抽取其一；这种方案是他用之于搜索的资源数量的函数。他的任务是要找到这样一个备选方案，它能使扣除搜索费用后的预期利润最大化。对这类模型来说，从已经发现了的备选方案中挑选最佳方案，是不成问题的；于是，决策问题变成了活动者应花多大力量进行搜索的问题。<sup>①</sup>

最后，还有一种给理性加上限制的办法，那就是：认为成本函数或其他环境约束极其复杂，致使活动者无力计算最佳行动方案。从这个角度考虑理性限度的观点，在古典的合理行为理论中不曾占据过优势。但在数值分析中，近似计算理论与此类似，因为它是把近似值看作计算量的函数，并研究近似值可以得到改进的速度问题。把计算费用明确引进这种理论，就可以将它变成一种最优近似理论。

替代古典目标的办法。我们不仅可以通过改变条件和约束的性质，而且可以通过变更给定目标的性质，去修正古典理论。有些现代厂商理论，并没有在上面提到的那几个方面偏离古典理论，而是假定了新的目标，不同于古典理论那种谋求最大利润的目标。

例如，鲍莫尔（1959年，第45—53页）提出了一个模型，其中假定厂商谋求最大销售量，约束条件是：利润不应低于某一确定的“满意”水平。按照鲍莫尔的这种理论，古典模型中的（6）式应换成：

$$dR/dq = p + q d(D^{-1}(q)) / dq = 0 \quad (6')$$

① 关于这方面的例子，可参见斯蒂格勒（Stigler, 1961年）的文章。关于搜索的资源配置理论，也可以通过考虑不完全了解后果的情形，加以构造。系列采样理论就属于此类，因为它回答的是这样的问题：我是现在做决策呢，还是等到收集了更多的信息再做决策？回答这个问题的办法是：对增加样本所导致的费用增量，和由于决策的平均改进程度而导致的预期收益，进行比较。

并满足约束条件

$$P = R - C \geq P^* \quad (7)$$

我们看得出，若采用鲍莫尔理论解决具体问题，其信息需要和计算要求同古典模型的要求没有太多的差别。

本文不打算讲那些假定目标不是谋求最大利润或最大效用的理性论变种，除非其中有关目标的假定同有关条件和约束的假定之间，存在着明显的相互联系。不过，我们将会看到，这是一种非常重要的例外。事实上，对有关条件和约束的假定做了明显改变的那些理论变种，大多也要求改变有关目标的假定，使之不同于谋求最大利润或最大效用的古典假定。关于这种相互联系的理由，后面将会讲到。

### 棋艺的理性抉择途径<sup>①</sup>

有些从事理性决策研究的人，一直把棋赛作为反映真实决策情景的一个有趣的微观样板。对棋赛中的理性抉择所做的研究，为其他可能的理性途径，提供了一些有益的例子。

对于轮番走棋的棋手所面临的问题，我们有两种解释。第一种解释是：棋手要寻找一个好的（或最好的）策略。这里所谓“策略”，指的是一系列有条件的走法。它确定了棋手在对方采取了每一个可能对策之后的各接续阶段里将采用什么走法。

第二种解释是：棋手要对自己目前面临的备选走法进行精确的评价。

从古典观点上看，这两个问题不分彼此。棋手如果具备无限的计算能力，那么，他或是为自己在棋赛中的未来行为选取一个完整的策略，或是逐次在轮到他走棋时选择走法，是无关紧要的。

① “棋赛”指国际象棋赛。译文为简便起见，均略称“棋赛”或“下棋”。  
——译注

因为他评价下一步走法的程序是先针对以后全部赛程情况，构造各种完整策略，然后从中选择一个能产生最佳报偿（也就是，在认为对手也将力求获胜的条件下所能得到的最佳报偿）的策略。这就是冯·诺伊曼—莫根斯特恩对策论<sup>①</sup>所采用的方法（冯·诺伊曼和莫根斯特恩，1953年）。

**对策论关于棋赛中的理性的定义。**冯·诺伊曼和莫根斯特恩认为，下棋是一种微不足道的游戏。“……如果国际象棋的理论（即完整的可能对策树图）真已全部为人所知，那么，这种游戏就没有什么值得玩的地方了”（出处同上，第125页）。<sup>②</sup>设可能对策树图中的每一个终点表示白方胜、败或平。棋手在该树图某一分枝上考虑其走法时，沿该分枝逆向检查，就可以看出他所选择的走法会导致哪个终点，因而可以选取一个能导向满意结局的走法，这样一来，那个终点的值就是导向该终点的分枝的值。照这样倒推下去，就可以给白方的每一种棋势（状态）指定一个值——白方胜、败或平；最终则能给白方的每一种初始合理走法指定这样的值。于是，每一位棋手都能通过确定他在该树图每一分枝点上所欲选取的是哪个走法，进而确定一个最优策略——能保证其结局至少不次于对方的策略。

不幸的是，正如冯·诺伊曼和莫根斯特恩也认识到了的，棋赛的这种平易性，根本不能给棋手实际选择走法带来任何切实的帮助。“我们的证明只保证了，在三种可能情形里〔棋赛必有白方胜、败或平三种值〕有一种（而且只有一种）情形成立，它并没有给出在实际上可用的方法，来确定究竟是哪一种情形成立。由于人的因素方面存在着困难，必须运用那些不完备的、启发式的下棋方法。正是那些方法构成了所谓‘高’招，如果没有

① 原著正式译名为“竞赛论”，但目前多译作“对策论”。见王建华、顾玮琳的译本《竞赛论与经济行为》，科学出版社，1963年版。——译注

② 中译本第113页。这段译文引用了王、顾译本，但译本中没有括号里的文字。——译注

有它们，棋赛就毫无‘斗智’和‘计谋’可言了。”（出处同上）。①

我们如果实际计算一下，棋手在比赛时若按对策论去寻找策略，需要做多大的搜索量，那么，上述“不实际”的意思就显得更生动了。平均来看，一盘棋赛的任一给定状态，约有30种合乎规则的走法——以整数概略计之，一步走法及对方的反应，平均有 $10^3$ 种不同的接续走法②。一盘棋赛的平均总步数，可以合理地估计为40步左右。因此，可能出现的不同下棋方法，也许有 $10^{120}$ 种之多。这里的数字是否精确，显然无关紧要； $10^{40}$ 虽则不如 $10^{120}$ 那么引人注目，但其数目之大，已足以支持我们现在 的论断。

对棋手决策过程所做的研究，有理有据地告诉我们：下棋高手在选择一步走法或一个策略时，所考虑的可能走法或策略（从指定状态出发的接续走法或策略）不会达到100种之多。以某种标准看，100种可采纳方法已经算得上够多的了，但它离 $10^{120}$ 还远着哪！实际上，棋手不会去考虑全部可能策略并从中选取最佳者，而只不过构造和检验相当少的可能策略；他一旦发现了一个满意策略，就立即做出抉择（参阅德·格鲁特，1965年）。

在周密考察棋手的真实作法之前，让我们先回顾一下古典模型。我们要看一看，在国际象棋赛之类的竞赛当中，如果考虑到问题空间的规模，那种模型能否对实际抉择有什么用处。似乎存在着这样一种可能的途径：用一个小得多的问题空间去替代真实问题空间，使前者在某种适当意义上近似于实际，然后把古典模型应用于那个近似空间。

为国际象棋赛和西洋跳棋赛而编制的一些早期计算机程序，就采用了这一途径。例如在洛斯·阿拉莫斯的程序中，计算机逐

① [ ]内的文字是西蒙加的解释语。此段译文仍参照王、顾译本（第113页），但做了必要修正。——译注

②  $30 \times 30 = 900 \sim 10^3$ 。——译注

次构造所有合乎规则的走法，及所有合乎规则的对方反应等等，每次纵深考虑两步。然后，对这样生成的每一种终局状态（对两步纵深分析法来说，总共有100万种），应用极小极大方法予以倒推评价，寻找第一步最佳走法。这样一来，就用一个包含大约 $10^6$ 个元素的问题空间，取代了表示“真实”棋赛的那个含有约 $10^{120}$ 个元素的空间。

这个图式是近似的，因为那100万个终局的实际棋赛价值是未知的，而且，除非再转回去考虑那个包含 $10^{120}$ 个元素的空间，除非再去用对策论分析全部可能的棋赛，否则便无法确切知道那些值。该程序使用了一个取代未知值的办法，即：以棋手们普遍使用的经验法则计算近似值——给棋子指定约定数值和灵活性指标。这样看来，上述近似图式并不能保证选出客观上最佳的走法；他所选出的走法，只是在进行了两步分析之后，照上述启发式经验准则看来，可导向表现最佳状态的走法而已。经验告诉我们：采用这种近似评价法的程序，不可能精确到足以走出高招的地步。近似世界里的最优决策，不一定是在真实世界里的满意决策。

对奕思考中的寻求满意的过程。现有的下棋程序所采取的是另一条途径。它们竭力效仿棋手，只考虑非常少量的接续走法。这一图式的效力，大大依赖于三个部分：（1）走法生成器，对所要考虑的接续走法进行选择；（2）评价器，决定每一接续走法的优劣程度；（3）终止法则，用以决定何时应终止搜索、选定走法。

在棋赛当中，棋手只要看一看棋势特征，就能想到一些合适的走法。举一个极端的例子，某棋手在轮到他走棋时发现，他的一个卒处在攻击对方王后的位置上。显然，用那个卒吃对方王后，是值得考虑的一步棋。但这也可能是一步糟糕的走法——比如，他若吃了对方的王后，对方另一棋子就会将死他——不过，

这一步法有着明显的表观优点，而其缺点则只有通过对它进行动态考虑和评价，方能检查出来。要产生这类走法，有一个简单过程就可办到，即：确定对手的哪个棋子已受到一个价值较低的棋子的威胁，或受到任一棋子的威胁，或在受攻击范围里没有设防。这样，一组适当的走法生成过程，就有可能为进一步的分析，提供全部或大部分值得认真思考的走法。生成器的设置如果恰当，通常能首先构造出最有希望的走法，然后生成稍差的走法，如此等等。

对于走法生成器所构造的可能走法，我们可通过静态准则和动态准则的组合指标，予以评价。静态准则，讲的是状态的特征，或相邻状态之间差别的特征。例如，一切棋手所采用的重要的静态评价器之一，就是棋子计分：给每个棋子设一个约定值（比如：卒 = 1 分，马、象 = 3 分，车 = 5 分，王后 = 9 分），比较双方的总分。一般而言，如果一个棋手的总分比另一棋手多一些（许多情形下则可设仅多1分），那么，分数高的棋手就可以找到一个占优势的接续走法，除非对手有一系列高招能使局势迅速逆转。（因此，如果你能在下一步吃掉对方的王后，同时又能免遭报复，那么，少5分也不要紧。）

进行短期战术考虑的办法是：对看来很可能发生的接续情形做分析，直到出现了足以用静态评价器予以可靠评价的静态或“僵局”。然后用常见的极小极大方法，将静态评价过程倒推至所考虑的走法。

具有这种结构的程序，需要两种终止法则：（1）一旦出现了可作静态评价的僵局便终止搜索的法则；（2）一旦找到了一个满意走法便终止全部搜索过程并选定走法的法则。我们已对前者做过讨论了；后者则需要更细致地考察一下。抉择情境中的备选方案如果不是给定的，而是必须发现或创见出来，而且，可能方案又非常之多，那么，我们不及考虑全部或大部分方案，就必须做出抉择。正是由于对全部备选方案进行比较这一古典要求的

难以实现，使我们采取了现在所讲的途径。但是，如果我们不把全部备选方案都检查一番，那就必须用某些准则去确定，是否找到了一个稳妥的或令人满意的方案。在决策过程中起这种作用的准则，心理学文献称之为欲望水平。我们已经发掘出苏格兰方言“寻求满意”，用以表示设置欲望水平的问题求解过程和决策过程——搜索备选方案，直到发现一个符合欲望水平准则的满意方案，并选择该方案（西蒙，1957年，第4部分）。

对寻求满意式的程序来说，满意方案的存在性，多半取决于动态机制。它们以有关环境的信息为依据，按现实情况调整欲望水平。这样，下棋程序就可以根据状态的静态评价，设置最初的欲望水平（最好稍微偏高）。随着对各个备选方案进行动态的和静态的考察和分析，状态评价可逐渐减少，直到所发现的最好走法达到了或超过了欲望水平。

**棋艺理性的限度**。在本文导言部分，我们列举了人们给完美理性所加的三种限制：每一备选方案所导致后果的不确定性；不完全了解备选方案；以及使必要计算无法进行的复杂性。棋赛的情形向我们表明，这三类限制在真实世界的问题求解过程中是如何混在一起的。

我们如果把棋手说成是在选择策略，那么，他的理性行为所面临的困难——以及他照对策论原理去行动的不可能性——在于这样一个事实：他不完全了解自己面前的备选方案（或策略）。他有时间发现的，只是那些方案中的很少一部分；他所仔细考虑的，只是他所发现的那些不完备的方案。

另一方面，我们如果把棋手说成是在选择走法，他的理性行为所面临的困难则在于：他仅仅是非常粗浅地了解采取每一可能方案（走法）的后果。他要想出一系列合乎规则的走法，并不是不可能的，因为它们一般不过30种左右。但是，他对那些走法的评价，即便是近似的，也只有在巨大的分枝走法树图中做进一步分

析才行。由于进行评价的处理时间是有限的，他必须在各种备选走法当中，分配他的评价时间。事实上，仅仅构造全部合法走法中的少数几个并比较彻底地予以评价，往往比构造全部走法并肤浅地予以评价要更好一些。因此，高明的棋手不会去检查他所面临的全部走法，而只检查其中一小部分。（德·格鲁特（1965年）所报告的数据表明，棋手一般只构造和考察30种合理走法中的6~12种。）

仍从第三者立场上看，棋手理性行为的困难，同不确定性毫不相干（无论是后果的不确定性，还是各选方案的不确定性）；其困难在于这种行为的复杂性。因为，棋赛过程中没有任何经济学意义上的，或统计决策论意义上的风险或不确定性。按照冯·诺伊曼和莫根斯特恩的观点，棋赛是一种具有完备信息的竞赛。<sup>①</sup>在它的计算中不需要涉及任何未来事件的概率，也没有任何统计学意义上的意外事件发生。

依照对策论观点，对手的存在不会带来任何偶然性。对手总可以被认为是最强、最凶狠的敌人。如果我们把下棋换成证明定理，这一点就很清楚了。对后一种任务来说，根本不存在什么对手。也没有任何偶然因素：真理和可以推导出来的定理，永远存在于柏拉图理念乐园之中。定理证明过程中的理性之所以成为一个问题，仅仅是因为可能的证明路线所构成的迷宫，巨大而复杂。

所以，我们在讲到下棋或证明定理时所说的“不确定性”指的是：由于无力断定完全确定的环境结构（计算上的低能），即使那种环境带有不确定性色彩。但是，无论不确定性的起因如何，其结果是一样的，那就是：不得不在决策过程中以近似代替精确。特别是，当不确定性是以一个不便入手研究的问题空间的形式出现时，问题求解过程必须包含这样一些机制，它们能用以

---

① “完备信息”在王、顾译本及其他有些书中被译作“完全情报”。

——译注

决定何时终止搜索，选定一个备选方案。

寻求满意与寻求最优。前文介绍的“寻求满意”和“寻求最优”，是一定情形下理性行为的两类广泛途径的标签。这里所说的一定情形，指的是复杂性和不确定性致使全面理性不可能。对这些情形来说，寻求最优的途径变成了近似的优化途径——大刀阔斧地简化真实世界的情景，使之达到决策者能够处理的地步。

“寻求满意”的途径，则是在不同方向上简化真实情景，它保留了较多的真实情景细节，但试图做出满意的决策，而不是最优的决策。一般而言，我们无法预料它们当中哪个能导致更好的决策（以实际结果判断）。但至少对下棋来说，高明的棋手显然认为，“寻求满意”要比“近似求优”更有用。

通过引进最优搜索量法则，或最优设置欲望水平以决定搜索量的法则，我们常常可将“寻求满意”的程序，转变为“寻求最优”的程序。例如，棋赛方面的欲望水平可以动态调整到这样一种地步，使每多搜索一分钟所能带来的走法改进预期值，刚好抵得上这种附加搜索所需的费用。

尽管这种转变在形式上是可能进行的，但是，实际进行这种转变，则要求超出寻求满意程序所需的附加信息和附加假定。首先，必须使衡量备选方案价值的单位，与衡量搜索费用的单位相同，以便对它们进行边际比较。其次，搜索的边际生产率——每单位搜索时间的预期价值增量——必须以这样或那样的根据估计出来。一个人如果正在考虑设计一个下棋程序，会在如下问题上产生疑问：费力将程序纳入这样一种动态优化框架，其价值是否比得上以同样的努力，去改善程序中用以构造和评价走法的启发式法则的选择性。

“最优化”和“寻求满意”这两种图式之间的转变，还有一种与上述方式颇为不同的办法。这种转变方式也是一些人常常指出的。“古典”式的下棋程序，是在一个近似世界里做出最优决

策的。对这种程序，我们可以视之为特殊的寻求满意程序，其中“满意”是以所采用的近似方法来界定的。因此，我们很难在寻求满意和寻求最优这两种程序之间，划出一条泾渭特别分明的正式界线，使其中之一无法在另一框架内得到重新解释。不过，它们两者实际差别——采取不同观点所导致的侧重点上的差别，往往很大。

在有关优化程序的研究中，相当多的注意力被放到了评价函数的形式性质，求优方法的效力和最优解的存在性，以及降低不确定性的方法（如预测方法）上。至于将真实问题纳入适于最优化的形式所必需的近似作法的性质，以及选择各种近似作法的方法，研究得不够充分，系统性也差。例如，人们在求解线性规划问题的有效算法上，花了很多的力量。而如何将具体的实际决策问题整理成为线性规划问题，在很大程度上仍是一种技巧。<sup>①</sup>

对“寻求满意”的程序所做的研究，主要集中在搜索效率上——也就是说，这种研究的主要对象，是那些能使我们以比较有限的搜索量，在巨大的可能空间中找到稀少答案的启发式方法。由于启发式方法的较小变动经常能带来搜索效率上的很大变化（数量级上的变化），因此，对解的质量或其计算量进行评价的高度精确的方法，相对而言，恐怕就不大重要了。要把一个低能棋手的棋术变成一个相当有效的走法选择程序，也许并不需要有十分精致的评价函数或终止法则。

### 设计领域的有限理性

通常被冠以“设计”之称的工程活动，还没有在理性决策的标题下得到过比较多的讨论。其原因从以上论述中应是容易看得

---

<sup>①</sup> 查尼斯和库珀的论著（1961年）有好多关于这种技巧的高深例子。例如可参见第1卷的“附录B”和第11章。

出的：古典决策理论所关心的，是在给定备选方案当中做出抉择；“设计”所关心的，则是发现和构造备选方案。不过，我们对棋赛这一微观样板的探讨已经表明了，设计理论何以类同于寻求满意式的理性抉择理论。下面让我更细致地讲讲这一点。

我们先考虑一下把下棋看作策略选择任务（而不仅仅是选择一步走法）的那种解释。明确地讲，那种解释就是：棋手在寻找一种确能将死对手的组合（策略），即使这要求牺牲一些棋子。正常情况下，一位棋手除非能一步一步地展望到结局——除非能设计一条行动路线，一个万无一失的将死组合，否则他就采用不了这样的行动路线。

正如我们已经看到的，这种设计过程中的评价和比较，一般并非各个完整设计方案之间的比较。首先，搞评价是要指导搜索——设计方案本身的构造。评价为搜索方向的决定提供了依据。它告诉设计者，应朝这个方向努力，而不是向另一方向努力。完整的设计方案（对我们现在的例子来说就是将死组合）如果终于构造出来了，并不是通过不同设计方案之间的比较而得到一般评价的，实际上是通过欲望水平所定义的标准去比较它们，从而予以评价的。对下棋来说，棋手一旦发现了一个保证能将死对手的策略，就予以采用。他不去寻找一切可能的将死策略，并采用其中的最佳者（参阅H.A.西蒙和P.A.西蒙，1962年）。

在设计一个复杂对象时——比如说，设计一座桥梁或一架飞机——设计过程含有更加复杂的搜索结构。在这种情况下，初期搜索是在高度简化的空间中进行的；这种简化空间已把真实问题的大多数细节抽掉了，仅仅以概要形式保留了最重要的因素。我们一旦在设计空间里构造好了一个总计划，一个纲领性的总体设计方案，就可以另行引进细节问题，并将该总体方案用以指导完整设计方案的搜索。

当然，所采用的设计空间可以不止两个，可以有整整一系列设计空间，构成不同的设计层次，从高度抽象的总体设计逐渐发

展到细节的确定。在每一个抽象层次上，设计过程的结构也可以不同。由于比较抽象的空间趋向于“比较平滑”，所以，往往有可能将优化模型用于较高层次的设计工作，等到详细设计时，再重新采用寻求满意式的搜索模型。在这种情况下，可以用线性规划或动态规划搞企业厂级作业总计划，而用经验成份较多的启发式方法搞各项作业的日程安排。在其他情况下，总体设计过程可采用寻求满意式的搜索程序，总体方案确定下来之后，则可用优化技术选定参数。<sup>①</sup>

### 管理科学上的有限理性

作为管理科学与运筹学的技术主体，大多数形式化的技术方法，都是以某种准则去寻找一组备选方案中的最佳者的程序——也就是说，它们属于我们前面所讲的那类“古典”程序。线性规划和动态规划就是这些技术方法中的最强者。这一领域有关问题的研究途径，占优势地位的大多是要简化真实世界中的问题，使之达到最优化形式模型可用作近似的地步。

这种研究途径不易处理有些组合型的工业问题。具有典型意义的困难包括整数解问题，或通常归结为同样麻烦的难题——考虑大量要素的可能的排列组合。仓库选址就是这类问题。解决这个问题，是要“确定仓库位置的区域模式，让运输费用节约额和交货速度提高所导致的利润增加额，同库存业务的边际费用相等，从而使公司获得更多的好处”（库恩和汉姆伯格，1963年，第643页）。

库恩和汉姆伯格为解决仓库选址问题而设计的启发式程序，包括两个部分：“（1）主程序，用来逐一地安排仓库地点，直

---

① 用于化工厂设计的有些现代半自动程序，首先采用启发式方法选择工艺流程，然后再用线性规划方法确定该系统的参数。

到再在批发网上设置仓库就必须增加总费用为止；（2）删减与挪位程序，……通过评价取消仓库或位置挪动对盈利有何影响……对解进行改进”（出处同上，第645页）。

这个程序符合我们在前面对设计过程的刻画。首先通过搜索过程，一步一步地建立一个可能计划，然后在确定最终方案之前，考虑可能的局部改进，在构造初始计划时，尽量使各仓库位置靠近各个需求中心。构造过程的每一部都加上一个仓库，要求使整个系统的费用得到最大节约。在每一阶段，得到详细评价的仓库地点，只是所有可能地点中的一部分，并且是初看起来“不错”的地点。最后，用所谓“删减—挪位”程序，试探性地修改计划。修改办法是：（1）程序执行过程中后加的新地点，会使以前安排的有些地点变得不再合算，因而取消那些地点；（2）在各区域内考虑仓库位置的变更。这个仓库选址程序的流程，示于图3。它可用来说说明公式化启发式程序的典型结构。

库恩和汉姆伯格对这个启发式程序的某些方面，同优化程序做了细致比较。他们的结论是：“从理论上讲，线性规划方法是可以用来解决这类问题的。但从实践上看，许多问题的规模和非线性特征，往往使这种应用难以实行”（出处同上，第658页）。他们把启发式程序的优越性归结为两大理由：“首先是计算简便性，它大大节省了计算时间，并且能处理大规模复杂问题；其次是，对基本费用函数来说，这种方法很灵活，不必做限制性假定”（出处同上，第656页）。

对那些难于应用现有最优化方法的极其复杂的问题，管理科学上使用得最为广泛的技术，就是模拟。大系统模拟所可以考虑的细节数量，仅仅受到组合可行性的限制。但另一方面，模拟没有借助其他形式化的分析工具，因而没有给备选方案的发现和评价，提供任何直接的手段。对模拟来说，进行试误搜索的是研究人，而不是分析技术本身（参见福雷斯特（1961年））。

## 用于仓库选址的启发式程序

**1. 读入:**

- a. 工厂位置。
- b. M个潜在(可能)的仓库地点。
- c. 每次循环时予以详细评价的仓库地点数目(N), 即缓冲规模。
- d. 工厂、仓库、顾客三者之间的货物运送费用。
- e. 对每一顾客的销量期望值。
- f. 与每个仓库的业务相对应的费用函数。
- g. 与运货延误或其对需求的影响相对应的机会成本。

↓

**2. 在缓冲区确定N个可能的仓库地点。只考虑其局部需求, 这些地点如靠当地仓库供货, 而不是靠现已考虑过的仓库供货, 则可产生最大的费用节省。**



↓

**3. 逐次考虑缓冲区内的N个地点, 对每一地点, 考虑添加下一个仓库所造成每一种批发模式, 评价其可能给全系统带来的费用节省。**

↓

**4. 对N个位置中不再提供高于限额的费用节省的地点, 就不做进一步考察了。**

↓

**5. N个位置是否均能提供高于固定限额的费用节省?**

是



**6. 将仓库设在能提供最大节省额的地点**

否



**7. M个可能地点是否均成立, 或均已取消?**

否

↓ 是

**8. 删减-挪位程序**

- a. 删除那些由于后来安排的地点更好而显得不经济的地点。原来考虑让设在这些地点的仓库提供服务的顾客, 将由其余仓库中能以最低费用运转的仓库向他们提供服务。
- b. 就以上设置的每一仓库若移到现在服务地区的需求中心去, 其经济性如何, 进行评估。

↓

**9. 停止**

图3 一个启发式程序的流程图

## 结 论

过去30年来，理性决策论极其迅速地成长起来。这一成长发端于第二次世界大战期间及以后。其动力的一个主要源泉是，人们试图对相当复杂的真实情形，应用形式化的决策程序。为了对付事物的复杂性，形式化模型的效力和高深程度，都得到了提高。但另一方面，复杂性还刺激了新型的理性决策模型的发展。这种新模型特别考虑到人类及其计算机的很有限的信息收集能力和计算本领。

对不确定性的关注，对存在于备选方案设计及发现中的困难的关注，以及对计算复杂性的关注，赢得了这样一个反响，即：把搜索过程和信息传递过程明确地引进了模型。那些关注所赢得的另一个（与前者不相抵触）反响则是：将最优化准则换成了满意性能准则。在借助“启发式”方法或试误方法来寻找似乎合理的备选方案的那些模型中，寻求满意的途径得到了最为经常的使用。

作为所有这些发展的一个结果，企业、政府机构和大学等组织的决策者，今天已经获得了前所未有的种种模型和计算工具，当作他们制定决策的帮手。无论他们为理解和处理现实情况而必须同现实做怎样的妥协，那些工具都有利于让人类的有限能力得以对付困难问题，也都会使这一任务的完成变得容易得多。

## 参 考 文 献

- 〔1〕 鲍莫尔：《企业行为，价值与增长》，纽约：麦克米伦出版公司，1959年，第45—53页。
- 〔2〕 查潘尼斯等：《认知不谐论：五年以后》，载《心理学会刊》1961年。第61期，第1023页，
- 〔3〕 查尼斯等：《线性规划的管理模型与工业应用》，纽约：威利父子

出版公司(两卷集)，1961年。

- [4] 德·格鲁特：《下棋过程中的思考与抉择》，海牙：莫顿出版公司，1965年。
- [5] 福雷斯特：《工业动力学》，坎布里奇：麻省理工学院出版社，1961年（中译本，科学出版社，1985年）。
- [6] 库恩等：《用于仓库选址的启发式程序》，载《管理科学》，1963年，第9期，第643—666页。
- [7] 马斯查克：《组队论初步》，载《管理科学》，1955年，第1期，第127—137页。
- [8] 西蒙：《人的模型》，纽约：威利父子出版公司，1957年，第4部分，第196—279页。
- [9] 西蒙等：《求解难题过程中的试误搜索》，载《行为科学》，1962年，第7期，第425—429页。
- [10] 斯蒂格勒：《信息经济学》，载《政治经济学杂志》，1961年，第69期，第213—225页。
- [11] 冯·诺伊曼等：《对策论与经济行为》，普林斯顿：普林斯顿大学出版社，1953年（第3版），第125页。

## 五 企业组织的理性决策<sup>①②</sup>

阿尔弗雷德·马歇尔在其《经济学原理》一书的开篇绪论中宣称，经济学将成为一门心理学：

“政治经济学，或称经济学，是对一般经济生活中的人类进行研究的一门学问；它所考察的是个人行动和社会行动的一部分，是同安居乐业所必需的物质条件的获得和利用最紧密相关的那一部分。

如此看来，经济学一方面是研究财富的学问，另一方面，也是更为重要的方面，它又是人类学的一部分。因为，除了宗教信仰之外，对人的特征产生最多影响的，就是人的日常活动，以及由此获得的物质资源。”

然而，经济学在其实际发展过程中，却一直仅仅关注人类特征之一——人类理性，而且特别局限于人类理性在稀缺资源配置问题上的应用。不过，当代经济科学的种种定义，无论是用稀缺资源的配置来表述的，还是用理性决策的制定来表达的，均向我们展示了开拓者们的一块极其广大的领地。近年来，经济学家们甚至对这一领地中过去一直被认为属于政治学、社会学和心理学领域的方面，也进行了相当大量的探索。

① 本文译自《美国经济评论》，1979年（9月），第69卷，第4期，第493—513页。另载《有限理性模型》下卷，第474—494页。——译注

② 本文是作者被授以诺贝尔经济学奖时，于1978年12月8日在瑞典斯德哥尔摩发表的演讲。本文已由诺贝尔基金会于1978年正式出版。

作者感谢安多、戴维斯和B·弗里德曼，他们给本文初稿提出了宝贵意见。

## 作为经济学的决策理论

经济学家们居住着的整个经济学大帝国，其人口分布密度很不均匀，大部分人都居住在少数几个不大的区域里。经济学的心脏地带，是对国际经济、国民经济及其市场所做的规范性研究；其主要关切点有三个，一是资源的充分利用；二是资源的有效配置，三是经济产品的公平分配。我想用“政治经济学”一词指这个心脏地带，而用“经济科学”表示包括最偏远的殖民地在内的整个帝国，以此代替“经济学”这一含糊不清且过于宽泛的术语。本文的主要论题，就是被称为决策理论的那个重要的殖民地。下面我要讲一讲这个理论的规范性方面和描述性方面，并特别讨论一下它在厂商理论上的应用。正是这后一个论题，把我们的讨论与叫作“政治经济学”的心脏地带，联系在一起了。

具有政策倾向的全部规范经济学，当然有着一整套令人难忘的描述性理论，或称“实证”理论。这套理论在数学上的精巧漂亮，堪与自然科学的某些最精致的理论相媲美。在此，我只须举例提醒诸位，回想一下瓦尔拉一般均衡论，以及亨利·舒尔茨、萨缪尔逊、希克斯和其他一些人在其著作中论述的一般均衡论的现代子孙；或者回想一下由阿罗、赫维茨、戴布鲁、玛林沃德等人及其同事们所创立的、给人留下难忘印象的深奥理论，它证明了竞争均衡和帕累托最优在一定条件下的等价性。

至于这类理论中的某些高深部分是否与现实世界有什么关系，我们可以、而且已经提出了疑问。或许，有人之所以要攀登某些智慧的高峰，无非是因为那些高峰存在——无非是因为攀登高峰能使人感到刺激和振奋。<sup>①</sup> 人类的任何科学活动或艺术活

---

<sup>①</sup> 西蒙在《管理行为》（新版）第13章中讲过类似的例子。这里的寓意还可参见《关于人为事物的科学》第六章和本书第二篇（第三章）。——译注

动，都应当本着这种精神。但是，无论登山者的动机如何，也不管其符合现实的程度怎样，实证性的政治经济学的形成，无疑是受到了针对基本公共问题提出经济政策建议的要求的强烈影响。

事实也是理当如此。如果有人以为科学的研究要是预示着实用价值，或是以日常生活中的问题为研究对象，则不可能成为基础性的科学的研究，那就完全搞错了。这种观点不过是一种世俗偏见。实际上，在需要进行基础性研究的优秀科研课题的一切来源当中，现实世界恐怕是最丰富的了。

### 服务于政治经济学的决策理论

不过，还有一种与上述谬误相反的错误观点，也是理应受到同样谴责的。这种观点认为，基础研究只有同政策问题直接相关或显然相关，才值得进行。在当今世界上，这一谬见大概没有太多的市场，至少在自然科学领域是这样。我们对自然法则进行生气勃勃的和富有成果的研究，已经有三个世纪或更长时间了。很多研究工作的动力所在，无非是热切希望理解大自然，寻找那种隐匿在复杂性当中的秩序之美。我们一次又一次地发现，研究过程所产生的“没有用处的”真理，实际上对人类实践活动具有最为重要的意义。我不必在此论证这一点，因为，科学家们懂得这一点，工程师和医师们懂得这一点，众议员和国会议员们懂得这一点，老百姓也懂得这一点。

但是，我不敢肯定这一真理在经济学界也得到了应有的广泛认识。要不然，我就无法解释如下问题了：包括厂商理论在内的描述性决策理论，其发展为何颇为微弱，而且相当迟缓？在这个领地上生活的居民，何以既稀少又散乱？这方面的研究工作者，为什么许多（如果不是大多数的话）都来自经济学界之外——来自社会学界、心理学界和政治学界？一些在经济学界受到尊敬的杰出人物——例如梅森、麦克卢和M.弗里德曼，把这一决策研究

领域，划在经济学的管辖区之外了（更确切地说，他们是把经济学划到决策理论的管辖区之外了）。他们给了决策论领地以充分的自治权，但有一个条件，那就是：不得宣称自己同真正的经济研究有近亲关系。

例如，梅森在评论帕潘德罗于1952年发表的有关企业行为理论研究情况的综述时，自言自语地说道：

“……这类文献果真为经济分析做出大贡献了吗？……作为批评家的笔者必须承认，他对这种新的厂商概念较之旧的企业家概念，就经济分析而言显然更有利这一点，本人是缺乏信心的。”（参见文末所列文献，第221—222页。）

M. 弗里德曼也按照同样的态度，总结了他同理论上的现实主义进行论战的著名辩辞：

“彻底的‘现实主义’显然无法实现。看一个理论是否‘足够’现实，只能看它就眼下意图而言，是否能做出充分好的预见，或做出比其他理论更好的预见”。（参见文末所列文献，第41页；重点符是引者自加的。）

隐含在这两段引文中的“眼下意图”，给实证性的政治经济学，从而给规范性的政治经济学，提供了决策论基础。照梅森和M. 弗里德曼的观点来看，对企业组织环境中的人类理性行为进行基础研究，（按定义）根本不是经济学研究——也就是说，不属于政治经济学——除非这种研究主要是为那个意图而进行的。有时，这个观点甚至被解释成为如下意思，即：关于决策过程的经济理论，即使对微观现象的经验预见同观察数据不大相符，我们也不能说它们在任何有关系的或感兴趣的意义上出了毛病。据说，只要这样的理论同政治经济学上的总量观察没有矛盾，那么，它们仍旧是“足够”现实的。这样一来，极力主张经济活动者追求最优的那些经济学家，在评价他们自己的理论时，反倒变成了寻求满意者了。他们明白，经济理论家是寻求满意者，可是他们却认为商人追求最优。

一些人为了庇护这种将寻求满意的原则用于理论问题的作法，有时辩解说，这种作法等于是采用了奥克姆剃刀<sup>①</sup>：接受那个行得通的最简单的理论。<sup>②</sup>不过，奥克姆剃刀是双刃的。词句简明，并不是理论简洁性的唯一标准。奥克姆本人对他那个原则的理解是：要接受的理论，其所做假设仅为解释现象所必需者。利润或效用最大化的理论，同我后面讨论的那种寻求满意的理论相比，叙述上可“严而简明”。但是，对人类认知系统的假设，前者比后者苛刻得多。因此，就这一问题而言，奥克姆剃刀的双刃是朝向不同方向的。

无论我们怎样解释奥克姆原则，在对若干种理论进行选择时，简洁性的考虑只能是第二位的；除非那些理论的预见性不分高下。因此，我们必须回过头来，考察一下被认为是由实证性决策理论去处理的那些现象。它们既可以是微观层次上的有关决策者的现象，也可以是政治经济学上的有关总量现象。

### 寻求其本质重要性的决策理论

当然，重要的并不是“经济学”一词的定义。就象儿歌谜语

---

① 指“解释要简洁”的原则，以14世纪英国哲学家William of Occam命名。

——编译注

② “行得通”一语，最终否定了M·弗里德曼为理论假设可缺乏现实性而高唱的有名赞歌。让我们考察一下他那个有关落体的例子（见第16—19页）。他的成立论点是：忽略了空气阻力的简单定律，如果能提供“足够好”的近似，那么，采用这种定律就是有利的。然而，这种定律所赖以提供良好近似的条件，当然根本不是它变得偏离现实或成为“严重歪曲现实的表述”的条件。我们可以用这种定律预见物体在真空中下落的轨迹。但不能用它预见物体在大气中下落的轨迹。由于篇幅所限，我无法完全提到在弗里德曼那篇长达40页的文章中可以见到的许多逻辑错误，当然更说不上讨论它们了。关于对此所做的更多批评，请参阅本文后面所列出的文献：西蒙（1963年）和萨缪尔逊（1963年）。

“圆圆胖胖，坐地就碎”那样<sup>①</sup>，我们想让谜底指什么东西，就可以让它指什么东西。但是，经济学家的研究范围和专业训练内容，的的确确是重要的。如果我们接受了经济学只关心政治经济学总量现象这一狭隘观点，那就等于排除了把人类理性行为这一富饶领地纳入经济研究范围的适宜性，也就是，不把人类理性行为当作经济研究的适当对象。

有人说，企业行为理论一直没有与政治经济学的建立发生过关系。对这种说法，我也想讲一讲相反的见解。稍后我还要更多地讲述这种相关性。现在则以虚设形式说明我的论点：即使说目前根本没有反映这种相关性的证据，也不能不承认，企业中的人类行为是一系列极有意义的经验现象，它们同所有现象一样，也要求人们予以解释。我们如果可以用其他学科的发展史作类比的话，那就完全可以预料到，随着解释的出现，要它们同实际应用的重要领域发生关系，是不会耽延得太久的。

某些人在说到这类问题时，有时暗指这样一层意思，即：古典厂商理论有关理性行为的基本假定正确与否，不仅无关紧要，而且也无法予以直接的实验验证；唯一行得通的检验，就是看看这些假定是否能在宏观层次上，做出大致正确的预见（弗里德曼，第14页）。假如我们没有显微镜，不能直接观察到微观层次上的行为，那么，上述观点当然没错。但事实上，我们确实有显微镜。我们有许多用以观察决策行为的技术方法；如果需要的话，我们甚至能逐秒逐秒地观察那类微观行为。在检验我们的经济理论时，我们大可不必依靠那种粗糙的总量时间序列——经济计量磨坊所用的主要原料，我们甚至不必依靠公司财务报表。

关于企业和经济决策问题的古典理论，对决策代理者的具体行为，做了非常特定的可验证的预见。行为理论则进行了与前者颇

---

① Humpty-Dumpty，著名英语儿歌中的一个英雄，词义可任意改变，但通常答案是卵或鸡蛋。——译注

为不同的预见。由于这些预见都可以通过观察予以检验，因此，当这种预见出了毛病时，同对总量现象的预见发生错误时一样不难证明，它们两者之一（或两者）是错误的。

### 决策理论的总量检验：边际主义

如果说，有些经济学家错误地认为，对微观经济理论的检验，只能看它们对总量现象的预见是否准确，那么，我们也要防止犯一个相反的错误，那就是：认为总量现象同决策理论的检验没有关系。尤其是要看一看：究竟有没有一种得到经验验证的重要的总量预见，不是由行为理性论，而是由全面理性论做出来的？

古典理论讲的是全知全能的理性，其简洁和漂亮，令人赞叹。此外，那种理论使我们不用站起身来实际观察人类行为，便能坐在椅子上预测它们（正确或错误地预见）。所有这些预见效力都来自一个源泉，即：对人类行为所处的环境形态所做的刻画。体现着完美理性的那个环境，完全决定了行为。关于理性抉择的行为理论——有限理性论——不具备这种简单性。但是，作为一种回报，它们对人的能力所做的假定，要比古典理论宽松得多。因此，它们对人的知识和计算能力所提出的要求，是比较适中的和现实的。不过，它们并没有预见说，决策者将让成本和收益的边际值相等。

### 边际主义的预见得到验证了吗？

有人引举了一些经验现象，以此为古典厂商理论提供多多少少带有结论性的支持，反击古典理论的行为论对手们（参阅乔根森和西伯特的文章）。但是，没有任何直接的观察结果表明，个人或企业果真让边际成本等于边际收益。古典理论在经验上得到了证实的结果，几乎总是不如上述结论那么强。让我们看一看其中最重要的四个：其一，需求曲线一般具有负斜率；其二，拟合

的柯布-道格拉斯函数具有近似的一阶齐次性；其三，递减的规模收益；其四，经理薪金随公司规模的对数而变化。这些“事实”果真是事实吗？那些证据是支持寻求最优论、驳斥寻求满意论的吗？

1. 具有负斜率的需求曲线。根本不存在这样的证据能真正表明：消费者在实际分配其购物金钱时，试图使其效用最大化，从而使边际效用相等。经验数据予以确凿证明的是，需求曲线一般具有负斜率。（正如亨利·舒尔茨早就指出的，即使是这一“明显”事实，证明起来也颇费周折。）但是，具有负斜率的需求曲线，不从基于效用最大化假定的行为上得出，也可以从符合有限理性假定的广泛行为上，推导出来。加里·贝克虽然很难被认为是古典理论的敌对证人，但他在这个问题上讲得很不错：

“长久以来，经济学家们就知道，不管采用什么决策法则，消费者①可行集合（或机会集合）的某些变化，都会带来同样的反应。例如，实际收入的降低，必定会减少花费在商品上的金钱，至少是减少花费在一种商品上的金钱……然而，人们还很少认识到，无论决策法则如何，相对价格变动所导致的机会上的变化，仍会产生有条理的反应。尤其是，传统理论的基本定理——需求曲线向下倾斜的定理，很大程度上是由机会变动单独引出的定理，同决策法则没有多少关系”（第4页）。

说了这段话之后，贝克讲得更明白了，他说：“不光是效用最大化法则，体现多种非理性行为的其他许多决策法则，照样能得出因价格变动影响机会而造成的需求曲线向下倾斜性”（第5页）。②

2. 生产函数的一阶齐次性。据说由古典假设为之提供了充分

---

① 原意为“家庭”，在西方经济学著作中，此词与“消费者”通用。——译注

② 贝克在一个脚注中写明，他所谓非理性，指“对效用最大化的任何偏离”。因此，我讲的“有限理性”，用贝克的术语来说，就是“非理性”。

条件（但不是必要条件）的已观察到的现象，还有一个例子，即：在拟合的柯布-道格拉斯生产函数中，劳动力所占的产品份额，同劳动力要素的指数相等（参见西蒙和利维所撰一文）。拟合型的柯布-道格拉斯函数是齐次的，其阶数一般大致为1，劳动力指数的大小也大体合适。不过，我们不能把这些结果，当作支持古典理论的有力证据。因为，如果我们用实际上是由一个线性会计等式（货物价值等于劳动力成本加上资本成本）所产生的数据，去错误地拟合柯布-道格拉斯生产函数，那么，也不难得出同样的结果（参见菲尔普斯-布朗所撰一文）。对SMAC生产函数①，我们也可以做同样的评论（参见赛叶特和西蒙的论文）。因此，这些经验结果并不能使我们得出任何特别的结论，说古典理论和行为理论两者哪个更可取。它们两者同样与数据相符。

3. 长期成本曲线。长期成本曲线的情形，与上述情况有些不同。在古典理论看来，竞争均衡要稳定，就要有U形的长期成本曲线。有限理性论则没有预见这一点——幸而如此，因为现已观察到的数据表明，成本曲线究竟是不是一般具有那种形状，很值得怀疑。许多产业的证据告诉我们，那些曲线的较高一端基本上持平，甚至下降（参见沃尔特斯所撰一文）。这个结果，同有关企业增长和规模的随机模型相符（参见井尻雄士和西蒙的书），但不符合古典理论的静态均衡模型。

4. 经理薪金。公司高级经理人员的平均薪金，随公司规模的对数而增长（参见罗伯茨所撰一书）。这个结果是从古典理论的利润最大化假定中得出来的，仅仅利用了关于管理能力分布情况的某些很特别的假定（参见卢卡斯，1978年）。但是，用一个简单的行为理论，就可以解释这种观察到的关系。这里所说的简单行为理论仅仅假定，存在着唯一的、由文化环境决定的参数，表示管理人员的薪金和直接下属人员的薪金之间，有一个固定的平

---

① 经济计量学上的更一般的生产函数。——译注

均比值（参见西蒙，1957年）。就经理薪金数据而言，行为模型对现象的解释，比古典模型对同一现象的解释，要简洁得多（以关于外生变量的假定来看）。

5. 小结：没有找到的现象。要想达到结论性的评价，这里的论述还远远不够。但是，我以为，那些要求只用效用最大化或利润最大化理论（而不是用有限理性论）予以解释的特定现象，的的确确还没有在总量数据中被观察到。事实上，如同上述后两个例子所表明的，正是古典理论，而不是行为理论，在解释某些观察结果时，面临着真正的困难。

6. 古典理论的失败。古典理论当然可以被修补得足以处理相当广泛的问题——不确定性和猜度现象不起主要作用的问题，也就是说，比较稳定的、离竞争均衡不太远的经济行为问题。不过，当我们考察包含不确定性和不完全竞争的决策问题时，以有限理性模型取代古典理论，就显得很有道理了。古典理论从来不是为处理那类问题而设计的，也从来没有令人满意地处理过那类问题。采用了主观期望效用概念的统计决策论，以及对策论，为那类问题的概念明朗化，做了很大贡献；但它们并没有给人类真实行为提供令人满意的描述，在大多数情况下，它们甚至没有提供令人满意的规范性理论，也就是，在人和计算机的计算能力有限的条件下实际可用的理论。

后面我还要从正面更多地讲一讲描述性的有限理性论，但目前我想先谈谈经济学领域里的另一领地——规范性决策理论。自从第二次世界大战以来，活跃在这一领地的人数，得到了迅速增长。

### 规范性决策理论

研究决策理论，其目的不仅可以是为政治经济学构筑根基，或是为具有本质重要性的现象寻求解释，而且可以是为了直接向企业和政府的决策者提出建议。我搞不清楚，这一领地在第二次

世界大战之前何以如此人烟稀少。那时，在那里的居民主要是工业工程师、公共管理学者和企业职能方面的专家，他们都没有把自己特别看作是经济学科工作者。他们当中的杰出先驱者包括：数字计算机发明者、数学家查尔斯·巴贝奇，工程师弗雷德里克·泰勒，还有行政领导人亨利·法约尔。

第二次世界大战期间，这块差不多被废弃了的领地，又被从事军事管理与后勤研究的科学家、数学家和统计学家们，予以重新发现，并被重新命名为“运筹学”或“作业分析”。那时，运筹学家们同社会科学界相距如此之遥远，致使希望进入这一领地的经济学家们不得不着手建立自己的殖民地，给它取名“管理科学”。于是，在这一领地上就形成了两个专业社团；尽管它们两者目前在许多共同工作方面，结成了友好联盟，但仍旧各持自己的身份证件。

在这些工作中，最优化技术从经济学界传到了管理科学领域，还创造发展了新的最优化技术；其中包括有名的线性规划，同这一工具的前期开发相联系的杰出倡导者有丹泽、康托罗维奇和库普曼斯。

管理科学所采用的决策工具，有一个突出特征，那就是：它们必须以真实世界中可以获得的经验数据为输入，仅仅进行已有的台式计算器或稍后的电子计算机所可能进行的计算，去实际制定出决策，或提出决策建议。对这些方面来说，那种认为企业家一心求优的理想化的模型，是没有多大用处的。在那种模型中，寻求最优的企业家完全确定地了解一切——或者，至少也完全知道不确定事件的概率分布。而实际上，则必须以实际计算可行性的眼光去构造模型，无论这在近似和简化方面对模型提出了多么苛刻的要求。

在这些严格条件下，模型构造工作出现了两个方向。第一个方向是保留最优化，但对模型进行很大简化，使之足以保证能算出最优解（简化世界上的最优解！）第二个方向则是构造寻求满

意的模型，使之能以不过份的计算费用，提供充分好的决策方案。这种模型由于放弃了最优化要求，故而可保留更多的现实成份。因此我们说，决策者或者寻求简化世界中的最优解，或者寻求更加现实的世界中的满意解。一般来说，这两条途径在相对优势上不分高下，在管理科学界持续共存。

可见，在管理科学领域发展起来的规范性决策理论，如同描述性决策理论一样，都是主要研究决策制定方式和方法，而不是仅仅关心决策结果。正如我在另一篇论文中指出的（西蒙，1978年，b），这些理论是关于如何做决策的理论，而不是关于做什么决策的理论。

在此，我想从本人参与过的工作当中举一个例子，说明规范经济学的模型构造是如何受到计算可行性考虑的影响的（参见霍尔特、莫地利安尼、默茨和西蒙合撰一书）。面对不确定性和产品需求的波动，企业可保持库存缓冲，以此为代价，使其产量和雇员人数平稳化。那么，采用什么样的决策法则能保证合理的费用平衡呢？从形式上讲，我们所面临的是一個动态规划问题，但其求解计算量之大，一般说来可谓令人吃惊，并且常常无法容忍。

寻找某种可以计算的特殊问题，是摆脱这个困难的一种办法。如果我们假定该公司的所有费用函数都具备二次型的形式，那么，最优决策法则就变成了决策变量的一个线性函数，不难以费用参数算出来。同样重要的一点是，在未来销售量存在不确定性的情况下，决策法则中只包含期望值，而不包括高阶矩（西蒙，1956年，b）。因此，二次型费用函数这一假定，把原问题简化成为易于解算的问题了。当然，这样算出的解，尽管给我们所假设的那个简化世界提供了最优决策，但它至多给二次型函数所近似表达的现实决策问题，提供了满意解。这就是为了实际可达的满意，而牺牲了无法达到的理论上的最优。

决策人的理性，如果是其有限计算能力和不完备信息所允许的那种理性，那么，在规范性决策理论和描述性决策理论之间，

就存在着一个紧密关系。它们两者都以程序理性，而不是本质理性<sup>①</sup>，作为基本研究对象（西蒙，1978年，a）。随着用以计算最优决策和满意决策的新工具的发现，随着计算机效力的日益增长，规范性决策理论所提供的决策建议将会发生改变。随着新建议的传播和扩散，人们观察到的企业决策惯例亦将发生变革。这些变化可以产生一些宏观经济效应。例如，由于引进了计算订货点和订货量的形式化程序，美国企业的平均库存水平已有显著下降。对这一点，大家有着某种一致看法。

### 刻画有限理性的特征

通常被称作制度主义的传统学派，是企业行为理论的主要先驱。我们不清楚，那些往往被归到制度主义这一大标题之下所有欧美论著，是否全都有很多共同之处，或者其作者是否见解一致。我认为他们至多有一种共同信念，那就是：经济理论必须得到重新整理，以考虑市场交易所处之社会、法律结构。今天，我们甚至能发现经济学界内部的一种生气勃勃的发展，它试图在新古典价格理论的框架当中去实现制度主义者的目。稍后我再对此多谈一些。

美国的制度学派有一个著名人物——或许是最著名的人物，他就是约翰·康门斯。他的艰深著作（如《制度经济学》）从法学方面借用了很多术语，并企图用事务(transaction)作为其著作中的基本行为单位。在此我不想回顾康门斯的思想，只想简单说一下，在我关于组织决策问题的早期研究工作中，那些著作给了

① “本质理性”系指决策课题业务内容上的理性，程序理性则指决策程序本身的理性。注意切勿将“本质理性”理解为“实质性的合理”（与“表面性的合理”相对）。对决策问题来说，所谓“本质”即指问题本身，而“程序”则指解决问题的过程。——译注

我不少深刻启示（参见《管理行为》（中文本），第132页）。

康门斯对切斯特·巴纳德的思想也产生了巨大影响。后者是一位乐于探索知识领域的企业经理，他把自己在新泽西贝尔电话公司当总裁和其他企业、政府部门及非盈利组织当行政领导的经验，总结成为一本有关决策过程的精湛著作，取名《经理的职能》。巴纳德所提出的原始理论已得到实验研究的验证；这些理论包括关于组织权威机制的本质的理论，以及雇员接受组织目标的动机基础的性质的理论（后者即所谓“诱因—贡献”理论）。此外，他还对组织决策过程做了现实的描述，他将这种过程刻画为“机会主义的”。我在《管理行为》一书中屡屡引用巴纳德的著作，虽然不能充分证明，但确实也表明了他对我本人关于组织问题的思考的影响。

### 对描述性理论的探索

1934——1935年间，我在密尔沃基对市教育委员会和市公共设施处两机构共同负责的公共娱乐设施管理，进行了实际调研。其间我碰到了一种让人迷惑不解的现象。尽管那两个机构的头头看来均赞同娱乐计划的目标，并且不象是争夺权力，但他们在设施保养和游乐监管两方面的资金分配问题上，总是意见不一致，关系紧张。<sup>①</sup>他们不是象我学的经济学教科书上所讲的那样让一种活动的边际费用同另一种活动的边际费用相等，这是为什么？

对这个问题所做的进一步的探讨表明，他们之所以不让两种活动的边际费用相等，是由于他们的智力和知识不足以使他们做到这一点。根本没有可以度量的生产函数，能让他们从中得出有关边际生产率的数量推断；而该两机构的管理者所具备的定性

---

① 这里所说的“娱乐计划”主要是为儿童课外活动而设置的，故有地方教育委员会参加管理。西蒙在《管理行为》（第1版）中稍微更详细地讲了这件事。——译注

的生产函数观念又是彼此不一致的。对公共设施管理者来说，一个游乐场所就是一种实体设施，是灰蒙蒙的拥挤城市中的一块绿洲。而对娱乐活动管理者来说，一个游乐场乃是一种社会设施，在那里，儿童们可在大人的照看和引导下一起玩耍。

在这些情况下，人究竟怎样才能做出理性决策？他们怎样利用边际计算？如果他们不用边际计算，又是用什么去代替它的呢？我在密尔沃基观察到的现象，是人类决策的一种普遍存在的现象。对这种现象，组织理论常谓之子目标认同。如果一个组织的目标无法操作性地与行动相联系，（如果生产函数不能具体的公式化），那就将根据可以有这种联系的子目标，对决策做出判断。对这些子目标，没有什么唯一的决定办法。它们的决定，取决于决策者的知识、经验和组织环境。在这种模糊性面前，子目标的决定还会受到决策者的私心和权力欲的影响；影响方式相当微妙，却并非不可思议。

这种现象在个人决策过程中发生的频繁性，同它在社会决策和问题求解过程中发生的频繁性相同。今天，它在问题表达的大标题下，已成为认知心理学的一项重要的研究兴趣。在具体的刺激环境和先前经验背景下，一个人究竟如何将其复杂的信息集合体，组织到一个便于求解的问题模式当中去？牛顿如果确实看到了苹果落地，他是如何把这个现象表达成为地球吸引苹果这一事实的一个特例的？

这类现象为《管理行为》一书提供了中心课题。该书是“我在公共管理研究工作中试图建立一些有益于我本人在公共管理领域中研究工作的研究工具的探索结果。”实际上，其结果与其说是一个理论，不如说是一个理论的绪篇。它来源于这样一个信念：“决策是管理的心脏，必须从关于人类抉择的逻辑学和心理学上导出管理理论的语汇。”诸位如果愿意，可将其视为问题表达方面的一个练习。

通过考察可知，子目标认同这一现象，是一座非常大的冰山

的看得到的顶端。要看出整座冰山的形状，最好是把它同理性抉择的古典模型对照来看。古典模型要求决策者了解抉择面前的一切备选方案。它要求完全了解每一方案所导致的后果，或具备计算这些后果的本领。它要求决策者现在和将来对这些后果的评价是确定的。它要求决策者有能力用某种一致的效用尺度，去比较那些后果，无论那些后果种类多么繁杂，性质多么不同。因此，我们的任务是用另一种模型去代替古典模型。这另一种模型所要刻画的是：当人们不得不搜索备选方案时；当人们由于计算能力有限，以及外部世界的不确定性，而对选取具体方案的后果仅仅有着非常不完全的了解时；当决策者不具备用以比较不同方案的普遍而一致的效用函数时；他们可以怎样做出决策（也许还能刻画他们实际上是如何做出决策的）。

现已发现了好几种具有普遍应用性和广泛用途的程序，他们能把无法处理的决策问题转变成为可以处理的问题。其中之一就是前面提到的：以满意抉择取代最优抉择。另一种程序是用可以观察和测量出其实现程度的有形的子目标，来取代抽象的全面目标。第三种程序是按照信息沟通与权威关系的结构，给许多专家分配决策制定任务，协调他们的工作。所有这些程序及其它一些方法，均符合“有限理性”这一总标题。我们现在搞清楚了，人类为在现代世界上进行生产和政务活动而建立的种种精致组织，只能被理解为对付个人本领局限性的机器；它们对付的，是个人在复杂性和不确定性面前的理解力和计算力的限度。

有限理性观念最初形成时的这一颇为模糊而一般的形式，要求在两个方向上进行细致工作：其一是理论的进一步形式化；其二是有关主张的经验检验。《管理行为》发表后的十年间，我们在这两个方向上均取得了很大进步。其中有些是通过我的同事们和我本人的努力而取得的，有很大一部分则是通过遵奉同一时代精神的其它研究团体的努力而取得的。

## 经验研究

组织决策方面的实验资料，其主要来源一直是不折不扣的“人类学”实际调研，从中看出决策制定程序的描述，观察特定决策片断的进程。这方面的例子包括，我和盖茨寇、科斯梅茨基及庭达尔关于大公司决策过程中的会计数据利用方式的研究工作（1954年），以及赛叶特、马奇和其他一些人关于数个不同公司的特定的非程序化决策的一系列研究工作（参阅赛叶特西蒙和特罗所撰一文）。赛叶特和马奇大大发展和扩充了后一项研究工作，他们在其重要著作《企业行为理论》中探讨了该项研究对经济学的理论意义。

大约与此同时，迪尔邦和我偶然碰到了可以获得一些数据的机会，那些数据表明了工商管理人员对一本企业政策案例集所描述的某一问题情境的感知状况，这一机会使我们从经验角度，证明了那个在密尔沃基娱乐管理研究中使我难以忘怀的现象——子目标认同现象的认知基础。那些管理人员对案例所描述公司现有主要问题的感知，大多取决于他们自己的工作经历——销售、会计经理人员看出了销售问题，生产管理人员则看出了内部组织问题。

当然，关于组织决策问题需要了解和检验的东西，要比少数几项研究所能获得的东西多得多。尽管后来又有很多这方面的研究工作在欧美国家进行了，这一领域的开拓仍旧相当不够（参见马奇，1965年；约翰逊，1968年；伊利亚逊，1976年）。这类研究遭到相对轻视（比如同社会心理学上的实验室实验相比）的各种原因包括，它们花钱很多，耗用的时间也很长，播种量对收获量的比率较高。进行研究的方法比较原始，不易获得了解决策行为的满意途径。这一部分经济学研究工作，尚未养成其他方面工作中的那种刻意求实的耐心和坚持不懈的习惯，譬如西蒙·库兹涅茨的范例，或是MIT—SSRC—Penn经济计量模型<sup>①</sup>构造者们

① 麻省理工学院—社会科学研究会—宾夕法尼亚大学计量经济模型——译注。

的范例。

### 理论探索

理论上，看来有三个特别需要澄清的问题：第一，在什么样的情形下，一种雇佣关系可成为比其他形式的契约更有利于保证工作绩效的安排；第二，古典厂商理论同巴纳德首先提出的组织平衡论之间有何关系？第三，在复杂性排除了全智全能的情况下，人类理性抉择的主要特征是什么？

1. 雇佣关系，现代工业社会有一个基本特征，那就是：大部分工作不是由制造商品的个人干的，也不是由各个承包人干的，而是由企业中的人们干的。这些人接受了雇佣关系，并接受了因而导致的雇员—雇主权威关系。接受权威，意味着愿意把自己的行动交给雇主支配，至少在一定的无差异范围或接受范围里是这样。这种安排比之针对特定产品或劳务的契约，有何长处？世界上的工作，为什么有如此之多都是在层级制大组织中进行的呢？

分析表明，雇佣合同优于其他契约形式，可以用两个因素合起来予以解释。一个因素是，什么样的未来行为对雇主有利，确定不了；另一个因素是，同雇主相比，雇员（在其接受范围内）对自己究竟做什么活儿不够关心（参见西蒙，1951年）。例如，雇主在雇用一位秘书时，并不知道自己会让秘书打印什么信件，而秘书则对于究竟打什么内容的信件感到无所谓。雇佣合同可将抉择延续到消除不确定性之时，这对雇员没有什么不好，对雇主则大有好处。这一解释同马斯查克对流动性偏好所提出的解释十分相似。在不确定性条件下，以流动的、灵活的方式占有资源，是有利可图的。

2. 组织平衡。巴纳德曾从组织参加者（雇员、投资者、顾客、原料供应者）愿意留在组织系统中的动机出发，描述了组织生存问题。我在《管理行为》一书中，进一步将这一观念发展成

为一种动机理论，①即：组织向其参加提供的诱因与参加者向组织资源所做的贡献两者之间相平衡的动机理论。

这一理论的公式化（西蒙，1952年，1953年），表明了它同古典厂商理论的接近性，不过，它们之间有一个富于启发意义的重要差别。在比较这两种理论时，每一种诱因-贡献关系都变成了向厂商供应资源的计划。生存条件变成了盈利条件。古典厂商理论假定，全部利润都归一类特殊的参加者——所有者，组织理论则更为对称地对待剩余，而且没有预见它会如何得到分配。因此，在垄断和不完全竞争条件下，后一理论为参加者们（如劳工与所有者）就剩余问题进行谈判，留下了解释的余地。此外，生存条件——不是最大盈利，而是盈利——还允许背离那种全面理性假定。

3. **有限理性机制**。在《管理行为》一书中，有限理性主要被刻画为一个余类——缺乏全智全能的理性，就是受限制的理性。全智全能的缺乏，主要指的是：不能知道全部备选方案，有关外生事件具有不确定性，以及无力计算后果。那时，还需要对有限理性条件下的抉择机制，进行更加直接的和更加形式化的刻画。我的两篇论文（1955年，1956年，b）在这方面迈出了第一步。

这种刻画有两个主要概念：其一是**搜索**，其二是**寻求满意**。如果决策者一开始并没有现成的备择方案，他就必须去搜寻。因此，关于有限理性的学说，必须包含搜索理论。后来，斯蒂格勒在一篇很有影响的论文中独立地提出了这个想法；他在该文中以购买二手轿车的决策问题为例做了分析。他把搜索理论又倒回古典的效用最大化的旧瓶子里去了，认为搜索费用要同边际收益相等。在我1956年所写的一篇论文（b）中，我采用房宅出售过程的一个动态规划公式作为例子，证明了同样的形式等价性。

但正如我指出的，效用最大化对搜索图式来说并非不可缺少

---

① 亦译激励理论或工作动机（动力）理论。——译注

——幸而如此，否则会要求决策者在对于全面理性的运用来说本来已经过于复杂的决策问题面前，还必须有能力估算搜索的边际费用和边际收益。另一种办法是，假设决策者对于应当寻找一个好到什么程度的方案，已经形成了某一欲望水平。他一旦发现了符合其欲望水平的备选方案，便结束搜索，选定该方案。我把这样一种选择模式叫作寻求满意。这一模式以莱温等人关于欲望水平的心理学理论；为其经验基础。正如心理学研究已经表明了的，欲望水平不是固定不变的，它可随体验的变化而升降。在好方案多的良性环境下，欲望提高；而在恶劣环境下，欲望则下降。

从长期均势上看，如将搜索费用考虑在内，动态调整欲望水平而做出的抉择，甚至可以等价于最优抉择。不过，对关于搜索和寻求满意的理论来讲，重要的是它告诉我们，怎样就能以不过份的计算量和很不完全的信息，去实际做出决策，而不必进行那种不可能成功的事情——不必执行寻求最优的程序。

## 总结

由以上论述可见，到本世纪50年代中期，（1）已经提出了一种有限理性学说，成为古典的全面理性论的替代理论；（2）进行了相当多的经验研究，它们表明，企业真实决策过程同有限理性说的假定非常一致，但不符合全面理性的假定；（3）对关键性的部分理论进行了形式化的整理，这些部分是：权威关系和雇佣关系的性质，组织平衡，搜索和寻求满意的机制。下面我想首先追踪一下决策理论后来的发展，包括与有限理性论相竞争的其他方面的发展；然后评论一下，新兴的描述性决策理论对政治经济学有何意义（及潜在意义）。

## 新古典理论复兴

在50年代后期，如果有人预见有限理性论将会很快在经济

思想主流中占据很大一块地盘，那也没有什么不可理解的。那时，有限理性论的某些形式化结构已有很大进步，此外，越来越多的经验证据表明，有限理性论对企业组织决策的描述，要比古典的全面理性观念，真实得多。

但是，尽管时代精神的许多方面有利于这一发展方向，历史却并没有沿着这种笔直轨道前进。第二次世界大战期间及其后，有大量从事学术研究的经济学家直接接触到工商界，他们或多或少地获得了观察企业组织实际决策过程的种种机会。此外，积极投身于新兴管理科学发展的人们，也面临着提出实用决策程序的迫切需要。这些趋势，对于有关经济理性的基本假定变得更加现实，肯定起到了促进作用。

不过，这些趋势并不是战后经济学界所出现的仅有趋势。首先，经济学界出现了一股强烈的反行为主义潮流，试图在方法论上维护古典理论。在本文第一部分里，我已经评论了这类方法论主张。虽然我们可以对这些主张表示深刻异议，但它们是巧言善辩之词，并且在从事学术研究的经济学家们中间仍有影响。

其次，数学知识和技能在经济学界迅速传播，使古典理论发展到了深奥而漂亮的新高度；尤其是当古典理论同统计决策论，以及冯·诺伊曼和莫根斯特恩的对策论相结合时，更是如此。这一发展还扩展到了处理某些不确定性和不完备信息现象的地步，尽管其处理手法极其独特。数理经济学和计量经济学之花的开放，给两代经济理论家提供了形式化问题和技术问题的大花园。他们的才能被用之于追求优美的公式，推迟了他们对现实世界粗俗一面的认识。

如果说我的这些话对上述发展似乎有点批评色彩，那么，我必须承认，我本人也处在这些发展的行列当中，我赞赏这些发展，并且肯定不愿意返回到数学化以前的那个已被这些发展所取代了的时代。我这里所提到的是，经济学界人士的活动，显示了人类理性那种一时干一事的分阶段行动的特点，而且，这些职业经济

学工作者有时似乎未能适当权衡其注意力的分配，使新古典理论、宏观经济计量学和描述性决策理论，均得到适当的研究。于是，后两个领域，尤其是第三个领域，并没有得到我们所期望的那样充分的专业研究。在经济学帝国的心脏地带，人口更加拥挤，而其他地方的肥沃土地则仍旧有待耕耘。

### 搜索与信息传递

在以下三小节中，我讲一下古典理论试图克服其传统局限性的三种方式，这方面的努力甚至想证明一种具体采纳心理学假定的行为理论的发展是毫无必要的。第一种方式是引进搜索和信息传递，把它们明确地当作与成本和产出相联系的、可代入古典生产函数的经济活动。在这方面，我在前面已提到斯蒂格勒于1961年发表的信息经济学论文，以及本人于1956年发表的那篇论文中在同一方向上的探讨。

在这种理论中，决策者仍是个人。马斯查克和拉德纳所提出的组队经济理论则是一个重要的新方向。在他们的理论中，制定决策的是组队或组织中的群体。我们看到，真正的组织现象——由信息传递的费用而导致的决策专业化，乃是产生于理性计算。由于数学处理上的极大困难，该理论很大程度上是示例说明性的，而且仅限于微型组织的非常简单的问题。不过，它大大拓宽了我们对信息经济学的理解。

与统计决策论和对策论一样，这些理论都没有废弃全面最优化的假定。它们所引进的信息不完备性和费用，并不是被当作决策者的心理特征，而是被当作决策者技术环境的一部分。因此，这些新理论丝毫没有减轻决策者所面临的计算复杂性——它们不是设法帮助决策者，以大胆近似、简化和寻求满意去对付计算复杂性；相反，它们倒是大大增加了计算复杂性。按照这些理论，决策者不仅要计算其供应曲线和需求曲线的形状，而且要更精确地计算那些计算工作本身的费用和收益。因此，在某种程度上

讲，这些新理论虽然给人一种印象，好象它们处理了从前被人忽视了的不确定性及信息传递现象，但实际上，这种印象只是一种假象。不过，对许多经济学家来说，这一假象一直具有诱惑力。

### 合理预期理论

我要讲的第二个新古典理论发展，是所谓“合理预期”理论。这个理论的起源，有一点历史嘲讽的味道。前面我曾讲到霍尔特、莫地利安尼、默茨和我本人在管理科学上的研究，我们提出了二次型费用函数这一特殊（且易算）情形的动态规划算法。对这种特例来说，决策法则是线性的，未来事件的概率分布可用期望值代替，当作确定性等价物（参见西蒙，1956年；泰尔，1957年）。

约翰·默茨凭借他的想象，认为自己看到了不确定情形下的一种理性行为范式。在我们这个四人研究组的某些人看来是近似的、寻求满意的简化的那些东西，在默茨看来，却是全面理性重大防线。他于1961年在《计量经济学学报》上发表的萌芽性文章中讲：“有时听得到这样一种主张，认为经济学的理性假定造成了理论与观察不符，或是不足以解释观察到的现象，尤其是那些随时间变化的现象……我们的假定则恰恰基于一个相反的观点：动态经济模型的理性假定不够充分”（第316页）。

默茨所提出的理性新增长在于，“预期由于是对未来事件所做的有信息依据的预测，故而它同有关经济理论的预见基本上一样”（第316页）。看来，他是想快刀斩乱麻。他不是依靠精心构造决策过程模型去对付不确定性，相反，他想一劳永逸——让过程变得毫无意义（假如他的假设正确）。合理预期理论后来在萨金特、卢卡斯、普莱斯科特及另外一些人的努力之下得到了蓬勃发展，是大多数读者都了解的（例如可参见卢卡斯1975年所撰一文）。

要对合理预期理论做出最后判定，目前为时过早。正如一切

科学争论的解决那样，这个问题最终将取决于经验证据所做的逐渐甄别，这种甄别过程还刚刚开始。与此同时，一些严重的理论困难也已经为人们注意到了。正象默茨本人指出的，有关的费用方程如果实际上的的确确是二次型方程，采用“合理预期”决策法则就是合理的（即：谋求最大利润的）。因而我曾在另一篇文章中指出（1978年；a），将这种法则叫作“一致预期”法则，似可减少误解。

也许，安多和B·弗里德曼所做的证明更加重要（1978年；1979年）。他们指出，在系统不断吸收新信息、发生结构变化、决策者有学习特性的条件下，同稳态条件相比，合理预期法则的政策含义有很大差别。例如，在更多的动态条件下，对任何有限的时间期限来说，都不能再保证货币中性。但货币中性对于静态的一致预期模型来讲，一般是成立的。

此外，近期有些“修正主义的”一致预期理论，考虑了信息环境的变化，并且重新出现了种种行为假定，用以解释预期是怎样形成的一——决策者将考虑哪些信息，忽视哪些信息。但是，除非这些假定建立在极其独特而专有的基础之上，否则无非再次表明了，需要有一种明确的和正确的决策过程理论（参见西蒙，1958年，a；B·弗里德曼，1979年）。

### 统计决策论与对策论

新古典理论复兴的另外两个重要部分，是统计决策论和对策论。前者所深切关注的，是把不确定性（或更恰当地称之为风险）纳入决策模型的问题。它要求对决策者所具备的、有关变量概率分布的信息，做出英雄般的大胆假设。它无非是把决策者所面临的计算问题，又提高了几个数量级。

对策论所研究的是“猜度”问题，即一个经济活动者考虑其他活动者对自己的决策做何反应，而产生的猜测问题。在我看来，对策论这架非常精巧的机器的主要产品，已经十分清楚地告

诉我们，要给这类猜度问题定义一个毫不含糊的理性准则（或者说一场竞赛的“解”的精确定义，指的是同一个东西），实际上是不可能的。因此，对策论并没有象人们起初对它怀抱的希望那样，使有关寡头垄断和不完全竞争的理论，摆脱其矛盾和复杂的境况。相反，它证明了，这些困难是无法消除的。对于某个理性准则是否适合某种竞赛这一问题，我们可能有本事达成一致意见，但如果有人对这种一致意见提出挑战，偏喜欢另一个准则，我们将没有任何逻辑根据让他承认，他是不对的。

## 结论

以上关于新古典理论复兴所做的论述，恐怕已经足以表明，为什么那些发展在与行为理论的竞争当中，一直是很有吸引力的一种商品。至少对有些经济学家来说，它展示了这样一种希望和可能，那就是：不必牺牲全面理性这一重大假定，便可以对付那些曾使古典经济学感到棘手的重要问题了，因而还可最大限度地做先验推断，最小限度地做辛辛苦苦的经验数据研究。另一方面，我关于这些新体系局限性的论述，大概也足以说明，为什么我不认为它们解决了促使其发展的那些问题。

## 行为理论的进展

在过去二十年里，有限理性论和企业行为理论在整个经济研究活动中一直扮演的是默默无闻的角色，但它们在这段时期内还是取得了稳健的进展。由于综述其全部进展还需要做大量工作，因此在这里不得不满足于引述其中种种重要研究工作的少许例子，以此展示全部工作的风格。对于已有综述的专题，我将仅限于考虑那些综述。

首先，目前已经进行了一些心理学实验和实际调研，以检验

处在简单抉择情境中的人类行为是否象统计决策论所讲的那样（期望效用最大化）。其二，为发现人类决策和问题求解的微观真实过程，现已进行了很广泛的心理学研究；阿兰·纽韦尔和我一直对这类研究工作深深着迷。其三，关于组织和企业环境中的实际决策过程，现已做了许多经验观察——其中大多是以“案例研究”的形式出现的。其四，已经对厂商理论进行了重新构造和扩充，以行为决策假设代替了古典的最优化假设。

### 效用理论与人类抉择

效用和概率理论在第二次世界大战后的公理化，以及贝叶斯统计学的复兴，为从实验上检验统计决策论，检验抉择情境中的人类行为是否谋求最大的主观期望效用，打通了一条道路。在早期研究所采用的极其简单的抉择情境下，人的行为似乎符合主观期望效用理论。但是，即使是在抉择情境中引进少量的复杂因素，实际行为与主观期望效用理论的预见之间，就立即明显出现了种种背离。卡尼曼和特沃斯基提出了从经验角度否定这种理论的一些最鲜明、最令人信服的证据；他们指出，在有些情况下，决策者很不重视先前知识，差不多完全根据新的事实做出抉择，而在另外一些情况下，新的证据对成见没有多少影响。库恩鲁瑟及其同事在研究水灾保险方面的个人投保决策时发现，实际行为同主观期望效用理论所预见的行为之间的偏离，同样很大，很显著。根据这些证据及其他一些证据，我们必然会得出结论说，主观期望效用理论没有对真实行为做出好的预见，甚至也没有做出好的近似的预见。

应当注意到，对这种理论的否定是在决策的本质上做出的，而不仅仅是从制定决策的过程上做出的。问题并不在于，人们确实没有做出主观期望效用决策所需的计算——新古典主义思想从来不认为人们做过这样的计算，现已表明的是，人们甚至没有摆出一种仿佛是做了这样计算的姿态，这显然是对新古典主义假定的

直接否定。

### 问题求解心理学

关于理性决策过程的证据，很大程度上是反面证据，也就是表明人们不这样做的证据。不过，以往20年关于人们用来制定困难决策和求解复杂问题的过程所进行的研究，也积累了大量正面证据。根据这类证据而建立的一种理论，叫作信息处理心理学。这种理论，在形式上通常以计算机程序语言来表述。纽韦尔和我在《人的问题求解过程》<sup>①</sup>一书中，总结了我们自己的这种理论。该书只是这方面迅速增长的大量文献之一。那些文献均假定了一个信息处理框架，而且把计算机模拟用作表述和检验这种理论的主要工具。

信息处理理论认为，问题求解就是在往往非常大的问题空间中，进行具有高度选择性的搜索。由于以经验规则或“启发式法则”为依据的选择性，可以把搜索引导到有希望获得答案的区域去进行，因此，一般只搜索全部问题空间的一小部分，就能找到问题的答案。一旦找到了满意答案，寻求满意式的准则就起了作用，从而结束搜索。可见，问题求解理论显然符合我在这里所讲的有限理性说。

对问题求解过程的这个一般图景，我们现已获得了广泛的经验证据。大部分这样的证据适用于相对简单的、象猜谜似的抉择情境，这类情境可在心理学实验室中用有控制的研究来得到。但是，我们对人类所做出的具有专业水平的各种判断过程，也有了相当多的了解。例如医疗诊断、股票和债券的证券投资、国际象棋比赛等等。对这类事务来说，一般搜索机制是在人类长时期记忆的信息极多的环境中起作用的，但作用过程的一般组织，同解决

<sup>①</sup> 有时被译为《人类问题解决》、《人类问题求解》等的著名著作。——译注

简单、具体问题时的情形基本一样。

目前，信息处理心理学研究正在朝着数个方向前进。对专业技能的探索仍在继续。另外，还有大量工作试图决定，人们在考虑一个新问题时，是如何获得其初始表达形式的。即使是对简单问题来说，问题求解者对其将要搜索的问题空间的表述方式，也有很大的伸缩性。这一发现再次表明，解决问题的真实过程，同那种搜寻唯一确定的最优解的假设，相距何等之遥（参见哈耶斯和西蒙所撰一文）。

信息处理心理学研究对经济理论的主要贡献在于，它提供了比较具体的经验证据，说明了问题情境中的决策过程很符合前面所讲的有限理性模型。而这一结果又意味着，抉择并不是唯一取决于问题情境的客观特征，它还取决于制定抉择所用的具体的启发式程序。由此可见，过程模型是旨在描述真实世界的任何实证性决策理论的不可缺少的一部分，至于试图避免这种模型必要性的新古典雄心，是无法实现的（参见西蒙，1978年，a）。

### 组织决策过程

除了需要上述心理学研究的证据之外，我们还要对组织环境中的决策过程，进行经验研究。个人问题求解和决策制定过程的研究，不涉及组织决策过程中的许多社会心理因素。过去20年在组织决策过程方面虽然做了大量考察，但对它们不易进行归纳。这里的困难在于，这方面的研究方式，大多是针对个别组织的特定决策或具体类型的决策，进行案例研究。据我所知，这方面的文献还没有用一篇好的综述加以概括，因此，就连确认已经进行的研究，也是困难的。<sup>①</sup>至于从各个案例研究当中抽出它们对决策过程

<sup>①</sup> 关于进一步的参考文献，可查阅文末：马奇，西蒙；约翰森；杜顿，斯塔布克。但是，还有大量具体案例研究，从来没有在这些文献上记载过，其中有些案例研究是学位论文课题，有些以管理应用的特殊领域为考察对象（可参见伊利亚森，1976年）

普遍理论的意义，也没有任何系统的方法被提出来并加以检验。

因此，组织决策方面的案例研究，代表着一种科学探索的自然发展阶段。那些案例研究向我们提供了关于决策过程的大批事实——它们差不多完全符合我们所讲的行为模型。但是，我们还不懂得怎样利用那些事实，去正式检验我们的模型。<sup>①</sup>我们也不大懂得如何来处理不同的组织分别采用的不同决策程序这样一种现象，并且即使是在一个组织的内部，所采用的决策程序也随着情况的变化而发生变动。我们千万不可以认为，对这些数据、资料的概括，竟能象体现在新古典理论中的概括那样简洁而精密。

怎么从案例研究中抽象出与理论有关的信息？解决这一问题的方法也许有一条最为接近的途径，那就是计算机模拟。把决策过程方面的经验证据转化为计算机程序，就开辟了一条渠道，它既可检验用以解释数据的程序机制是否恰当，又可以使我们发现能对实际行为的有趣而重要的性质予以定性解释的程序有何关键特征。采用了模拟技术的例子包括：克拉克森对投资信托员决策过程的模拟，赛叶特、费根鲍姆和马奇对一种双头垄断过程的模拟，以及波尼尼关于会计情报和管理压力对改变企业职工工作动力的效果的模型。杜顿和斯塔布克从各种不同角度，讨论了模拟方法论问题。<sup>①</sup>

### 企业理论

有限理性的普遍特征——选择性搜索、寻求满意等等，已被好几种试图构造包含行为假定的企业理论的工作，当作了自己的出发点。这类理论大致包括上文提到的赛叶特和马奇的理论，鲍

① 除了对企业所做的模拟研究之外，还有一些非常有趣、具有潜在重要性的研究工作，即利用模拟方法去建造决策理论直接通向政治经济学的桥梁。参见奥卡特(G. Orcutt)和卡德威尔—沃特亥默的著作，以及伊利亚森(1978年)的著作。

莫尔的满足最低利润约束条件的销量最大化理论，马里斯以增长速度表述目标的企业模型，雷本斯坦将产量控制在理论可达量之下的“X-低效”理论，科奈关于供应驱动型管理和需求驱动型管理的二分论，威廉森的交易费用论，纳尔逊和温特的进化论模型（1973年），赛叶特和德格鲁特考虑了适应性学习特性的模型（1974年），拉德纳的寻求满意显式模型（1975年，a，b），等等。

按这种方式来描述，好象所有这些理论和模型彼此之间没有多少共同点，只不过它们皆以某种方式偏离了厂商决策全面理性的古典假定。但是，更仔细地考察它们，更抽象的描述其假定，我们就可看出，它们都有几个基本特征。它们大多背离了短期利润最大化的假定，而且代之以一个用定额确定目标的假定——也就是说，它们在不同程度上成为了寻求满意式的理论。它们即便保留了最大化假定，也同时包含了某种机制，靠着这些机制，至少能保证不去追求短期最优。在赛叶特-马奇理论和雷本斯坦理论中，我们可以认为，产生“组织余资”<sup>①</sup>的，就是这种机制。组织余资的大小可以是激励变量和环境变量的函数。

最后还要讲一点，在这些理论当中，有几种理论假定了可发生组织学习行为，因此，如果环境在足够长的时间里稳定不变，那么，系统的平衡势态将会越来越接近古典理论那种寻求最优的均衡。当然，这些理论一般还假定，环境扰动一般将充分大，足以避免古典理论解变成真实行为的恰当近似。

组织余资之类的概念在企业模型中的存在，给企业短期行为带来了复杂性。由于企业可以在距离任何最优解都非常远的状态下运转，因此，余资是作为企业决策与环境之间的缓冲而存在的。这样，对于外部事件所引起的反应，就不再能简单地靠分析“形势需要”来预测了。那些反应还取决于企业所采用的特定的

<sup>①</sup> 直译可作“组织松弛”。按照赛叶特-马奇的企业行为理论，这个概念是刻画组织剩余资源的抽象概念，其直接来源之一是巴纳德-西蒙的组织平衡论。

——译注。

决策过程。然而无论企业模型的这种特征多么符合现实，它毕竟减少了那种模型对许多经济学家的吸引力。我所讲的这些经济学家不愿意放弃古典理论那种不靠过程做预见的作法，他们不熟悉揭示真实决策过程所需的经验考察。

不过，这个问题还有另外一面。如果说，在相同的环境下，不同的决策制定机制能导致不同的企业行为，那么，结果对过程的这种敏感性，对市场和经济层次上的分析来说，可以具有重要意义。无论是描述性的还是规范性的政治经济学，都不可对这种反应多变性的根源漠然视之。它至少要求我们——在由理论导出政策结论之前，尤其是在依照那些政策去行动之前——要用敏感性分析去检验一下，若对微观层次的决策制定机制做出不同假定，会使结论发生多大变动。

假如我们的结论很稳健①——如果我们用这种或那种行为模型去代替古典模型，并不能使结论发生大的变化，那么，我们对自己的预见和建议就能增加信心。另一方面，假如结论对这样的代替很敏感，那么，我们在使用那些结论时就要加倍小心，直到我们能确定哪个微观理论是正确的。

正如本节援引的文献所能证明的，我们关于市场和经济运行的预测，对决策过程层次上的机制假定，是敏感的。此外，各种行为理论的假定，几乎肯定要比古典理论假定更接近现实。这两点合起来，就直接否定了那种认为古典理论的不现实假定没有害处的论点。我们不可用真空落体定律去预见一个重物在糖浆里下沉的过程。我们必须以百倍的小心去看待古典理论和新古典理论所做的预见，以及由此提出的政策建议。

## 结 论

政治上有这么一句话：“没有武器，何以出击”。的确，仅

① 指结论对条件（或结果对参数）不敏感。此词似尚无正规的学术意译，但科技界有音译作“鲁棒”。——译注

仅靠指责人家的缺陷和不足，是无法推翻一项措施，或击败一个竞选人的。你必须提得出替代措施或另一人选。

同样的原则也可用于科学理论。一种理论一旦有了稳固的根基，将能经得住想推翻它的人以不利证据对它进行的大批攻击，从而在逆境中生存下来，除非有一种符合证据的理论已经准备好取代它了。既成信念的这种保守的防卫性，的确并非毫无道理。首先，对经济科学来说，我们只是想接近真理；我们并不想找到一个能囊括全部真理、且不含任何杂质的公式；纵然是一个相当复杂的公式，我们也不奢求。我们满足于一种逐步逼近的策略，当发现理论与数据不符时，我们的第一个想法是进行修补，而不是推倒重来、从零开始。

其次，当一个理论出现了缺陷时，出毛病的地方很少能立即明朗化。麻烦可能出在理论的基本假定当中，但也很可能仅仅出在辅助假定或计量方面的假设上。那些辅助假定和计量上的假设，只是为了使理论同观察到的现象相联系而不得不设置的。修正理论结构的这一部分，有可能足以挽救其余部分。

那么，古典厂商理论的现状如何呢？现已毫无疑问，这种理论的微观假定——全面理性假定——是背离事实的。这不是近似的问题，它就连不近似地描写人类面临复杂问题时所采用的决策过程，也说不上。

此外，现在存在着一种替代理论。或许，替代理论竟能多到令人眼花缭乱的地步。今天，我们已经从实验室研究和实际调研当中获得了大量描述性的数据，它们表明了人类在多种多样的处境下是如何求解问题和制定决策的。另外，目前已构造了一些用以解释实际数据的理论，尽管它们尚未形成一套和谐一致的整体，但它们有着很多共同之处。它们均以这种或那种方式，体现了有限理性观念：要求对决策方案进行搜索，用定额和寻求满意的目标取代最优化，以及学习及适应的种种机制。如果我们的兴趣所在是描述性决策理论（甚至规范性决策理论），那么，我们

现已完全清楚，古典理论和新古典理论已被另一种优势理论所取代了，后者在对实际过程的描述方面，向我们提供了一种远为贴切的近似。

但是，如果我们的兴趣所在不是经济科学广阔领域的遥远地带，而是规范性的政治经济学，那又将如何？有没有理由说明，为何应当放弃那些更熟悉的理论？“就经济分析而言”，关于企业和决策的新概念是否显示了它们的优点？

如果真象人们有时争辩的那样，古典理论和新古典理论无非是推导总量效应的有力工具，而那种总量效应对全面理性和有限理性两者来说又都是成立的。那么，就此而言，我们完全有理由保留那些理论。但是，我们所看到的却与此相反，在总量现象和政策的层次上，新古典理论的论证结果，并不总是与某种有限理性论的论证结果相同。因此，我们不能以假定的真实性不重要为由，来辩护那种不加批判地使用违背事实的假定的作法。事实上，在许多情况下，对于达到政治经济学中心问题的正确结论来说，假定的真实与否可能很关键。至于某种情况是否属于这许多情况之一，我们只有对那种情况下的理论预见进行比较，才能知道。

社会科学习惯于在自然科学最显赫的成就当中寻找榜样。这没有什么不好，只要不是象奴隶那样亦步亦趋就行。在经济学界，赞美牛顿力学（或是前面讲过的赞美落体定律）并寻找经济学上的物体运动定律，一直是很常见的作法。但对一门科学来说，这并不是唯一的模型，并且就我们的意图而言，看来这也确实不是合适的模型。

对人类行为，即使是人类理性行为，我们不能用少许几个不变量加以解释。采用完全适应环境这一假设，肯定是解释不了人类行为的。不过，人类行为的基本机制可能比较简单，~我就是这样认为的。但这种简单性处在极其复杂的边界条件之中，与之发生相互作用。那些边界条件，则是由环境和人的长时记忆，以及个体和集体的人类学习所设置的。

如果我们想用自然科学上的比喻来作指南，我愿意从生物学上，而不是从物理学上，提出这样的比喻（参见纽韦尔和西蒙，1976年）。从生物进化论上，我们可以找到一个明显的启迪；但我们也完全可以从分子生物学方面得到教益，尽管这不是那么明显。特别是，我们可以在分子生物学领域看见这样一幅图景：少许几个基本机制——如双螺旋的脱氧核糖核酸，或是彼得·米歇尔所精巧刻画的能量迁移机制——竟能用以解释五花八门的复杂现象。我们能了解这是怎么做到的。我们还可以看到定性结构定律在科学上的作用，看到定性解释如同定量解释一样的效力。

我总是不愿意用将来时态，来结束一次关于人的科学的讲话。因为，如果用将来时的话，会给人留下这样一种印象，好象那些科学学科是以后某一天有可能实现的潜科学，但如今它们还不存在。实际情况自然决非如此。无论我们关于人类行为的知识离需要还相差多远，我们现在已然拥有很多这类知识。我们之所以有时出现一种贬低它们的倾向，那只是因为，很多现象都是我们作为人，在人群当中的日常生活里容易看到的，故而有司空见惯之感。此外，那些知识并不总能回答我们要知道的事情。我们不能很准确预见经济周期过程，不能很好地控制就业率（还可以补充一句，我们不能很准确地预报斯德哥尔摩的下一个雷雨天，也不能控制世界气候）。

尽管有这些限制条件和保留，今天我们确实懂得了人类理性抉择的许多机制。我们确实知道了，那种被叫作“人”的信息处理系统，在其知识范围所不能匹敌的复杂性面前，是如何利用自己的信息处理能力去找出备选方案，并计算方案后果，消除不确定性，从而（有时而不是总能）找到那些在目前看来是充分的和令人满意的行动方案。

## 参考文献

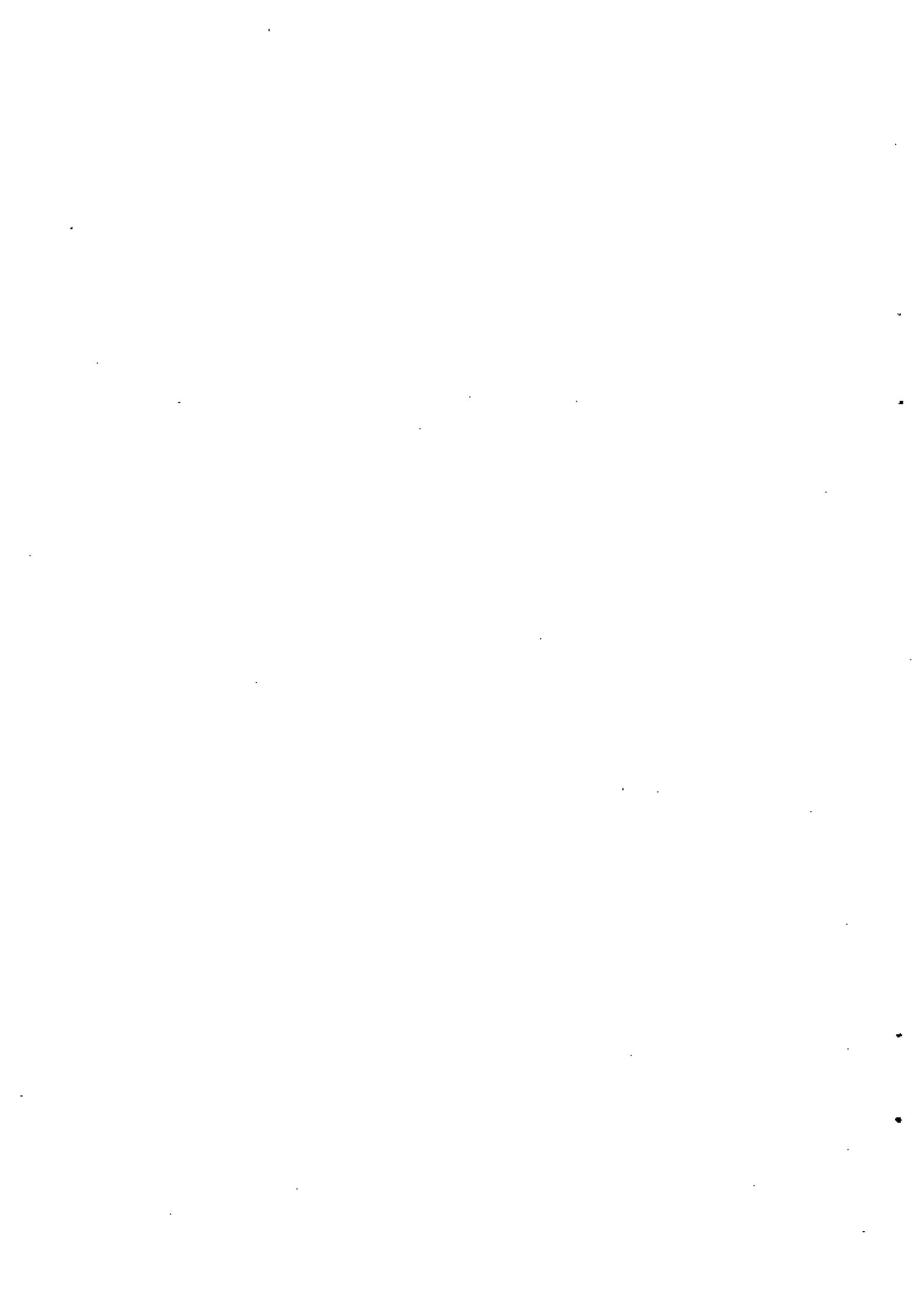
- 【1】阿尔奇安：《不确定性，进化和经济理论》，载《政治经济学

- 杂志》，1950年6月，第211—221页。
- [2] 安多：《论宏观经济模型的理论基础与经验基础》，宏观经济模型会议论文，1978年10月。
- [3] 巴纳德：《经理的职能》，坎布里奇，1938年。
- [4] 鲍莫尔：《企业行为、价值与增长》，纽约，1959年。
- [5] 贝克：《非理性行为与经济理论》，载《政治经济学杂志》，1962年2月，第1—13页。
- [6] 波尼尼：《企业信息决策系统的模拟》，英格尔伍德，1963年。
- [7] 钱德勒：《策略与结构》，坎布里奇，1962年。
- [8] 邱吉尔等：《审计行为效应的实验室研究和实际调研》，载波尼尼编：《管理控制》，纽约，1964年。
- [9] 克拉克森：《信托投资过程模型》，载费根鲍姆与费尔德曼编：《计算机与思维》，纽约，1963年。
- [10] 康门斯：《制度经济学》，麦迪逊，1934年。
- [11] 赛叶特，费根鲍姆和马奇：《企业行为理论模型》，载《行为科学》，1959年4月，第81—95页。
- [12] 赛叶特和德格鲁特：《合理预期与贝叶斯分析》，载《政治经济学杂志》，1974年5—6月，第521—536页。
- [13] 赛叶特和德格鲁特：《适应性效用》，载R·Day等编：《适应性经济模型》，纽约，1975年。
- [14] 赛叶特和马奇：《企业行为理论》，英格尔伍德，1963年。
- [15] 赛叶特和西蒙：《厂商理论：行为主义和边际主义》，未发表，1971年。
- [16] 赛叶特，西蒙和特罗：《企业决策的观察》，载《企业管理》，芝加哥大学，1956年10月，第237—248页。
- [17] 迪尔邦和西蒙：《选择性感知：经理的认同》，载《社会计量学》，1958年，另载《管理行为》，第15章。
- [18] 杜顿和斯塔布克：《人类行为的计算机模拟》，纽约，1971年。
- [19] 伊利亚森：《企业经济计划》，纽约，1976年。
- [20] 伊利亚森：《瑞典经济的微-宏模型》，斯德哥尔摩，1978年。
- [21] B·弗里德曼：《最优预期和“合理预期”宏观模型的极端信息假定》，载《金融经济杂志》，1979年1月，第23—41页。

- 〔22〕B.弗里德曼：《关于卢卡斯和萨金特方法论前提的讨论》，载《菲利普曲线以后》，波士顿，1978年。
- 〔23〕M.弗里德曼：《实证经济学论文集》，芝加哥，1953年。
- 〔24〕哈耶斯和西蒙：《理解书面问题指导语》，载W.Gregg编：《知识与认知》，1974年，第167—200页。
- 〔25〕赫施曼：《退场，声音，忠心》，坎布里奇，1970年。
- 〔26〕霍尔特、莫地哥利安尼、默茨和西蒙：《生产、库存和劳动力的规划》，英格尔伍德，1960年。
- 〔27〕井尻雄士和西蒙：《偏斜分布与企业规模》，阿姆斯特丹，1977年。
- 〔28〕约翰森：《多目标决策模型研究》，伦德，1968年。
- 〔29〕乔根森等：《公司投资行为理论的比较》，载《美国经济评论》，1968年，58：第681—712页。
- 〔30〕卡尼曼和特沃斯基：《论预测心理学》，载《心理学评论》，1973年，7月，第237—251页。
- 〔31〕科奈：《反均衡》，阿姆斯特丹，1971年。
- 〔32〕库恩鲁瑟等：《防灾保险：公共政策教训》，纽约，1978年。
- 〔33〕雷本斯坦：《经济人之上》，坎布里奇，1976年。
- 〔34〕利索尼：《经济分析的个人论》，第1卷，阿姆斯特丹，1977年。
- 〔35〕卢卡斯：《经济周期的均衡模型》，载《政治经济学杂志》，1975年12月，第1113—1144页。
- 〔36〕卢卡斯：《论企业规模分布》，载《贝尔经济学杂志》1978年秋季号，第508—533页。
- 〔37〕马奇：《组织手册》，芝加哥，1965年。
- 〔38〕马奇和西蒙：《组织》，纽约，1958年。
- 〔39〕毛里斯：《“管理”资本主义的经济理论》，伦敦，1964年。
- 〔40〕马斯查克：《流动性在完全及不完全信息情形下的作用》，载《美国经济评论文献集》，1949年5月。第182—195页。
- 〔41〕马斯查克和拉德纳：《组队经济理论》，纽黑文，1972年。
- 〔42〕马歇尔：《经济学原理》，第8版，纽约，1920年。
- 〔43〕梅森：“评注”，载B.T.Haley编：《当代经济学概论》，第2卷，霍姆伍德，1952年，第221—122页。

- 〔44〕蒙夏斯：《经济系统结构》，纽黑文，1976年。
- 〔45〕默茨：《合理预期与价格运动论》，载《计量经济学》1961年7月，第315—353页。
- 〔46〕默茨：《指数加权预测的最优性》，载《美国统计学学会杂志》，1960年6月，第299—306页。
- 〔47〕纳尔逊和温特：《走向经济能力进化论》，载《美国经济评论文献集》，1973年5月，第440—449页。
- 〔48〕纳尔逊和温特：《经济增长问题的新古典理论与进化论》，载《经济杂志》，1974年12月，第886—905页。
- 〔49〕纽韦尔和西蒙：《人的问题求解过程》，英格尔伍德，1972年。
- 〔50〕纽韦尔和西蒙，《作为经验研究的计算机科学》，载《美国计算机协会通讯》，1976年3月，第113—126页。
- 〔51〕奥卡特等：《通过微观分析模拟探讨政策》，华盛顿，1976年。
- 〔52〕帕潘德罗：《厂商理论基本问题》，出处见〔43〕。
- 〔53〕费尔普斯-布朗：《拟合柯布-道格拉斯函数的意义》，载《经济学季刊》，1957年11月，第540—560页。
- 〔54〕拉德纳（1975a）：《费用减少的行为模型》，载《贝尔经济学杂志》，1975年春，第196—215页。
- 〔55〕拉德纳（1975b）：《寻求满意》，载《数理经济学杂志》，1975年6—9月，第253—262页。
- 〔56〕罗伯茨：《经理的报酬》，格兰科，1959年。
- 〔57〕萨缪尔逊：《方法论问题的讨论》，载《美国经济评论文献集》，1963年5月，第231—236页。
- 〔58〕舒尔茨：《需求理论与需求衡量》，芝加哥，1938年。
- 〔59〕西蒙：《管理行为》。
- 〔60〕西蒙：《雇佣关系的形式理论》，1951年。
- 〔61〕西蒙：《组织理论的比较》，1952年。
- 〔62〕西蒙：《理性抉择的行为模型》，1955年。
- 〔63〕西蒙：《理性抉择与环境结构》，1956年。
- 〔64〕西蒙：《不确定情形下带二次型指标函数的动态规划》，载《经济计量学学报》，1954年1月，第74—81页。
- 〔65〕西蒙：《人的模型》，纽约，1957年。

- [66] 西蒙：《经济学与行为科学中的决策理论》，载《美国经济评论》，1959年6月，第273—283页。
- [67] 西蒙：《方法论探讨的问题讨论》，见[65]，第229—231页。
- [68] 西蒙：《从本质合理性走向程序理性》，载拉特西斯编，《经济学中的方法论评价》，剑桥，1976年。
- [69] 西蒙(1978a)，《作为过程的理性和作为思想产物的理性》，载《美国经济评论文献集》，1978年5月，第1—16页。
- [70] 西蒙(1978b)：《论如何决定做什么》，载《贝尔经济学杂志》，1978年8月，第494—507页。
- [71] 西蒙和科斯梅茨基等：《会计部门组织中的集权与分权》，纽约，1954年。
- [72] 西蒙和利维：《柯布-道格拉斯函数评注》，载《经济研究评论》1963年6月，第93—94页。
- [73] 斯蒂格勒：《信息经济学》。
- [74] 戴尔：《动态计划中确定性等价物评注》，载《经济计量学学报》，1957年4月，第346—349页。
- [75] 冯·诺伊曼和莫根斯特恩：《对策论与经济行为》。
- [76] 沃尔特斯：《生产与费厄泼兰函数：计量经济学综述》，载《经济计量学学报》，1963年1—4月，第16—6页。
- [77] 威廉森：《市场与等级制：分析及反托拉斯意义》，纽约，1975年。
- [78] 温特：《寻求满意，选择与创新之余》，载《经济学季刊》，1971年5月，第237—261页。



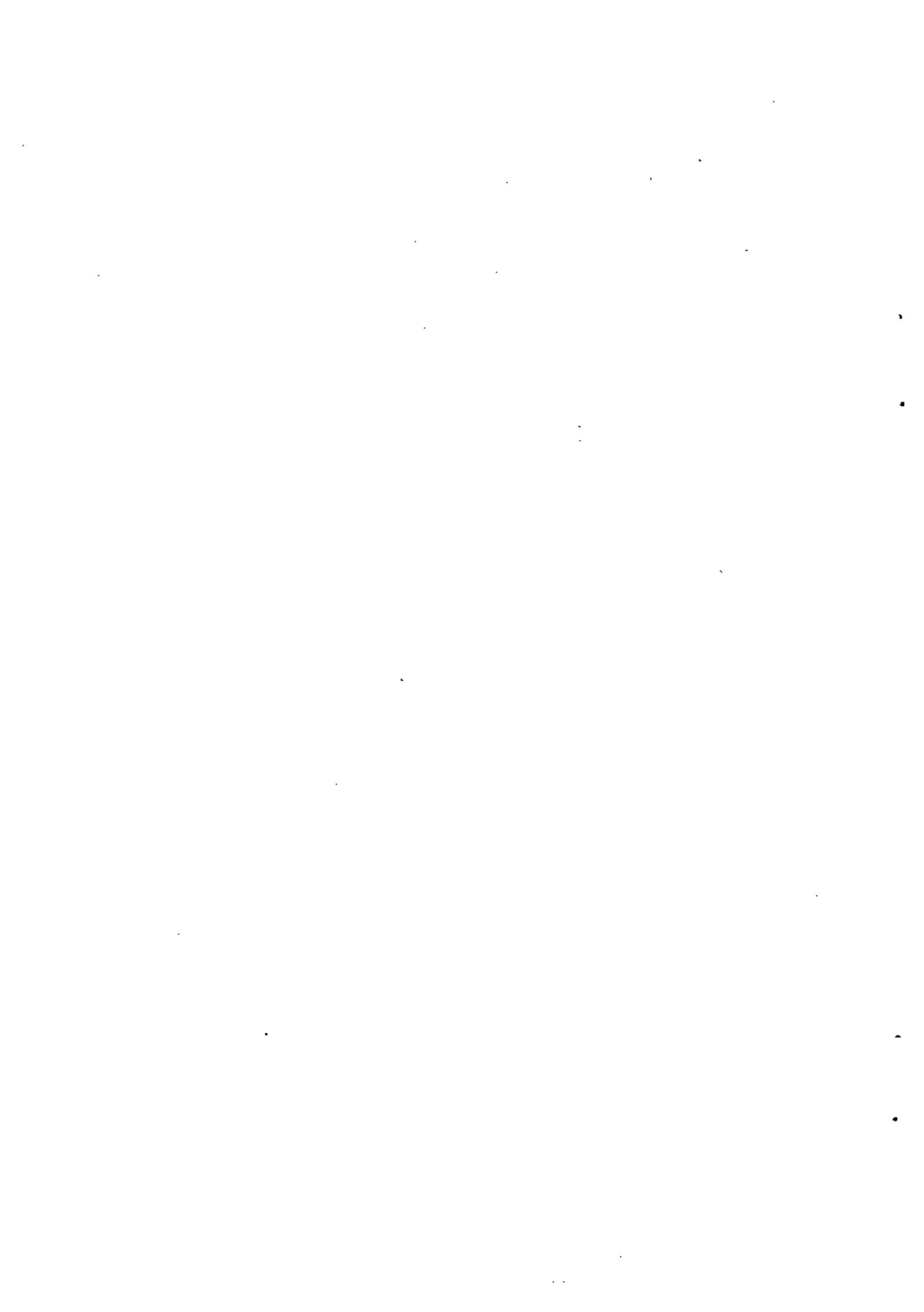
## 第二篇

# 论 事 理<sup>①</sup>

——人类事务上的理性

---

① 本篇系赫伯特·西蒙在斯坦福大学所发表演讲的汇编，西蒙教授亲自整理，并由斯坦福大学出版社于1983年出版。译文除在必要处加了个别译注之外，其他按原文全部译出。原著献辞为：“为纪念J·马斯查克——他对人类理智坚信不移，他的热诚之心无与伦比。”——译注



## 绪　　言

人类理性的本质——其机制、效应及其对人类本身状况而言的意义，是我近五十年来专注研究的中心所在。当我收到斯坦福大学哈里·坎普讲座的演讲邀请书时，我不知道自己在这个课题上究竟还有什么可讲的。如果说存在这方面的专题，不是早已被肯尼斯·阿罗、詹姆斯·马奇和艾莫斯·特沃斯基等斯坦福大学的朋友们彻底考察过了吗？（这里提起的还仅仅是在这一领域某些方面进行研究的许多人中的几个人。）尽管确实如此，我还是撇开了这一顾虑，决定利用这次讲演机会，探索一下迄今一直不在我本人的主要探索道路上、但我感觉有兴趣的和重要的偏远小路。

我特别探讨了三个专题，它们是：直觉理性与情感理性的关系，理性适应与演化之间的类同性，有限理性对社会政治制度的运行有何意义。以下各章就这些专题所做的论述，是在有限理性说的一般观点框架内进行的。

感谢斯坦福大学给了我准备这些文稿的机会。他们的热情好客，使我对访问斯坦福大学永怀良好记忆。我还要感谢唐纳德·坎贝尔、理查德·利文廷和爱德华·威尔逊，他们对第二章初稿提出了宝贵意见；当然，这并不是说他们对该章定稿中的全部内容均无异议。我衷心感谢他们，感谢在进化理论及其他专题上给了我有益指教的许多朋友们。

## 一 几种理性说

有一种乐观主义观点（或者说是想象的乐观主义观点）认为，如果我们充分努力地去思考，如果我们充分恰当地依靠推理，我们就能够解决我们所面临的一切问题。18世纪这个“理性时代”，应当说是充满着这种乐观主义的。至于历史事实是否真的如此，我想留给史学家们去回答；不管怎么说，现在我们对这个世界上的理性所怀的希望，却肯定要比那种乐观主义更现实得多。

本文的目的就是要从现代观点出发，探讨一下理性在人类事物上的运用和理性的限度。为了避免刚刚提到的那种没有根据的乐观主义，我在前两章主要论述理性的限度，而不是理性的运用。第三章再转移论述重点。我相信，随着论题的展开，读者将会明白我为什么要首先讲述理性限度问题。只有懂得了这些局限性，我们才能设计出有效的程序，以利用人类推理能力所给予我们的力量。

在本章里，我首先要集中讨论一下本世纪所产生的有关理性的形式模型。这些模型很有影响，必定属于我们这个时代的辉煌成就之列。由于这些模型广为人知，因此，我仅对它们做一简要描述，主要讨论它们在人类事务中的实际作用为什么不如看上去应起的作用那样大。不过，我并不想在这里对它们进行大量批评。本章的后半部分将对人类有限理性进行更为现实的描绘，并考察一下，有限理性所提供的那种有限的分析能力，究竟能在多大程度上满足人类事务对理性的需要。

在第二章里，我想讲一讲当今常常与生物社会学相联系的一

个论题，那就是：自然选择的更为严格的理性，会弥补理智上的缺陷；这种影响可能是好的，也可能是不好的。在讨论过程中，我将特别注意两个问题：其一，利他主义能否在自然选择力支配下的系统中生存，以及它究竟能在多大程度上在这样的系统中生存；其二，选择过程在多大程度上类同于最优化过程。

根据这两章得出的结论，我想在第三章里讨论一下如何在人类社会事务中有效利用理性的问题。

我们在科学上总应当提出新的真理。科学论文可能遭受到的最要命的评语，就是那种煞有介事的评审者在论文页边上歪七扭八地批道：“此处所谓新的发现并不正确，而正确的东西又不是新的。”不过，本文并不想成为科学发现的报告，因而将不追求新奇。只要我讲的大部分是正确的，即便不是全新的，我也心满意足了。正如我将在论述人类理性时指出的那样，人们的注意力要求人们反复利用重要的既得真理。

与此同时，我也不想简单重复我过去的著作中已详尽论证过的东西，特别是我在《管理行为》和《关于人为事物的科学》<sup>①</sup>两书中周密论述过的内容。那两本书均深刻论及人类理性的概念。在第一本书里，我考察了人类理性限度对组织行为的深意。在后一本书里，我描述了一切适应性系统（即“人为”系统）的共同性质，为我们建立关于这类系统的普遍理论，奠定了基础。在进行讨论的框架上，本文借用了这一现成成果。但在这一框架之内，我集中探讨了那些尚有疑问的或争论不休的课题，以及那些对于理解理性在人类事务中的角色尤为关键的专题。此外，我还扼要指出了其中的某些疑难课题。

## 推 理 限 度

阿基米德的现代子孙们，仍在寻找用杠杆移动整个地球的支

① 后者也有中译本，杨砾译，解放军出版社1987年修订版（1985年第1版）。  
——译注

点。在推理的范畴里，寻找这种支点的困难，就在于“没有无前提的结论”这一老生常谈。推理过程需要输入符号，并经过转换再输出符号。最初输入的是公理，而公理本身不是靠逻辑推出来的，它们纯粹是凭经验观察归纳出来的，甚至干脆是假想的。此外，把输入转换成输出的过程（推理法则）也是规定性的，并非推理的产物。公理和推理法则，共同构成了推理思考这根杠杆的支点。然而，那个支点的特定构造，却无法靠推理方法予以证明。因为，为此所付出的努力，会使我们陷入无穷无尽的逻辑递推之中，而这种递推的每一步又都同其前一步一样，基础总是任意的。

这种无法根除的任意性——这一败坏了推理过程的名誉的“罪魁”及其产物，对我们这里的论题来说有两个重要推论。第一个推论是，推理过程永远无法支配不容置疑的归纳法则，而归纳法则是使我们得以从具体事实出发，确信无疑地得出正确的普遍定律的法则。即使具体事实无穷无尽，也得靠归纳法则。例如，无论我们看到飞过来多少只白天鹅，都不能保证其后是否还会飞来一只黑天鹅。甚至能否就下一只天鹅的颜色问题做出确定的概率描述，也是有争议的。我想，就这一问题而论，否定意见要多于肯定意见。

另外，归纳的基础——事实，无不处在观察、感知和推理等复杂的、有时还是不稳固的依据之上。科学研究尤其如此。在科学上，常常要用仪器来收集事实资料，但仪器本身则含有理论假设。任何一部显微镜都是根据光学理论制造出来的，任何言语学习的口述记录<sup>①</sup>，也无不含有短时记忆理论。因此，推理之所以难免有误，既是因为无法从特定事实当中得到不容置疑的普遍命

① 研究人类记忆的实验手段、实验心理学在这方面做过很多实验，例如让被试者机械记忆一串字符，然后复述出来。显然，在设计实验时必定事先考虑到了人的短时间记忆能力，譬如，若呈示刺激（字符）的时间过短，则根本无法进行这种实验。——译注

题，又是因为事实本身具有暂时性和受理论影响的特点。

对我们这里的论题有着重要意义的另一个推论是，“没有无前提的结论”这一原理永远无法改变规定性叙述（即不可排除“应当”字眼的叙述句）的性质，这种句子总是依赖于同样包含着“应当”的输入。没有任何一种得到公认的推理法则能够单用描述性输入，导出规定性输出。<sup>①</sup> “没有无前提的结论”，意味着“没有单单从是怎样导出的应当怎样”。因此，推理对于寻求实现目标的手段来说，固然是有益的，但它对目标本身却没有什么帮助。

哥德尔首先指出了这里的根本困难所在，那就是：富有逻辑的系统永远不可能是完美无缺的——总是存在着这样的真定理，它无法靠把合法转换应用于输入而导出。这一逻辑不完备问题的重要性，比推理在人类事务中的应用的重要性要小得多，因此我对该问题不做进一步的讨论了。我也不想论及逻辑的标准公理和推理法则本身是否具有某种程度上的任意性。对我们现在的讨论来说，我将把它们视为无可指摘的。

这样一来，只要提供了一组合适的输入或前提，推理就发挥作用了。我们如果要用推理去发现和选择行动方案，那些输入中至少要包括一组“应当如何”和一组“是如何”；或者说，起码要有一组意欲实现的价值，还要有关行动的事实。想用逻辑方法来证明那些“应当如何”和“是如何”的任何努力，除非将归结于同样是假设出来的另外一些“应当如何”和“是如何”。

## 价 值

我们看到，推理完全是作为手段来用的。它不能告诉我们走

<sup>①</sup> 在此我不想申述其理由。好多年前，艾尔（Ayer）曾就此做了很好的论述，参见《语言、真理与逻辑》，修订版（纽约，1946年），第6章。

向何处；顶多只能告诉我们怎么样走法。它是一支可以借用的枪，用来服务于我们的任何目标——好的或者坏的。这就给我们看待人类状况的观点，造成了很大的裂痕：是把我们的困难归咎于罪恶，还是归咎于愚昧和缺乏理性——是归咎于目标本身的劣根性，还是归咎于我们不懂得如何达到目标？

### 狂徒做法

孤立地看，如果我们象是为参加辩论而做准备那样，以分析的态度好好读一下希特勒的《我的奋斗》一书，尽管这件事令人生厌，却是让人加深理解推理只有有限价值的一种有益体验。这种体验很可能是痛苦的，但是它揭示了事实、价值和感情是如何同我们对人类事务的思考相互作用的。我之所以采用这个特殊的例子，是因为在这种情况下，读者的批判精神不大可能被拥护该书观点的意见所磨钝。

我们大多数人都反对希特勒在该书中所列举的许多所谓“事实”，尤其是他对欧洲经济困难起因的分析，以及他关于犹太人和马克思主义者（他还荒谬地认为这两者没有差别）乃是经济困难的根源等大部分论断。不过，假如我们把这方面的疑问暂且搁在一边，姑且承认他的“事实”是真的，那么，纳粹纲领有很大一部分是相当符合德意志民族安全目标的，甚至同德国人的福利目标也十分一致。就此而论，我们之所以不能接受纳粹纲领，并不是因为它的什么罪恶目的——没有人反对关心德国人的福利，也不是因为从那些目标出发进行的推理不完善，而是因为，把目标和纲领联系起来的那些事实前提，是我们不能接受的。这似乎告诉我们，可以用建立在更好的事实前提上的推理，去批判纳粹主义纲领，而使纳粹主义得到补救。

不过，《我的奋斗》这本书给我们造成的罪孽，看来总不是同冷静反应相称的。看来必定还有更多的原因使我们反对其论据。是的，其他理由显然存在。《我的奋斗》所讲的目标，说得轻一

些，是不完整的。人在叙述自己的目标时，通常把目标为之服务的“我们”，同其福利不属于“我们”基本利益范围的“他们”区别开来。希特勒所谓“我们”，指的是德国人——这个“我们”的定义，又基于有关雅利安人与非雅利安人之间遗传差异的某些难以置信的“事实”。撇开所谓北欧日耳曼民族纯洁性这一奇谈怪论不说，我们大多数人恐怕也不会象希特勒那样去定义“我们”。我们的“我们”或许不是指德国人，而是指美国人，要是到了 21 世纪的开明国家，我们的“我们”甚至是指全人类。在这两种情况下，我们都同《我的奋斗》一书发生了实质性的价值冲突。这个冲突是无法靠事实及推理上的改进取得根本解决的。我们关于“我们”的假设——关于我们所关心利益范围的边界的假设，是一个关于是非界线的基本假定。

《我的奋斗》所导致的最大恶意，也许就是由于希特勒在“我们”和“他们”之间划出了一条太鲜明的界线。他不仅给“我们”赋以优越地位，还主张，只要有利于“我们”的目标，对“他们”无论多么残酷都无可非议。就算承认了希特勒的一般目标和“事实”，我们大多数人仍反对他那种损“人”利“己”的作法。在我们的价值体系中，如果不把“他们”视为毫无权利可言的，推理就会向我们揭示出价值冲突——利“己”价值与不损“人”的一般目标之间的冲突。因此，《我的奋斗》必定要受到我们谴责的，并非其推理上的毛病，而是它所声称的事实，和它那卑鄙的价值观。

从《我的奋斗》一书中，我们还可得到另一个教训。我们只有读过好长篇幅之后才可发现，希特勒的推理并不是冷静的，而是热燥的。大家早就知道，当有人夹杂着激情和恶骂宣扬某种观点的时候，特别要仔细检查其前提和推理两个方面。我们虽然知道这一点，但并不总是这样做的。可惜正是在激情、恶骂引起了我们内心感觉上的共鸣时，我们便会失去警觉，变成丧失批判精神的读者或听众了。

希特勒之所以能在德国人中间巧言惑众，恰恰是因为，他的激情、恶骂同德国人心中已有的信念和价值观产生了共鸣。他作说道的热度，使那些读者们丧失了以推理法则和证据去检验其论点的能力。不独德国人同他所宣扬的事实和价值产生了共鸣，许多西方政客那种潜在的反犹太情绪和公开的反共观点，看来也同希特勒的不少论点气味相投。

因此，通过惨痛体验到我们最初直觉判断的失误，我们认识到，不能根据希特勒是一个狂徒就对他不屑一顾，因为其狂妄也是有做法的。他那些狂言妄语符合推理标准；同我们写作时所遵循的习惯做法相比，既不过头，也不欠缺。推理，不是也不可能是我们反对纳粹主义的主要武器。我们的主要武器，是与之对立的求实态度和价值观。

### 争论丛生

一方面，我们承认使用推理方法时的所有这些复杂情况，也承认不能单单从“是怎样”推导出“应当怎样”，但另一方面，我们还必须承认，对人的品行进行推理思考是可能的。因为，我们所宣称的“应当如何”，大多不是衡量品行的终极标准，只是作为走向其他目标的手段而采用的子目标而已。例如，孤立地看“量入为出地过日子”这类目标，好象是毫无问题的，但是，一个学生也可能会明智地举债成学。为未来生产进行投资而欠下的债务，是不同于赌注的。

如果（1）某些价值的实现会在现在或将来影响其他价值，（2）它们是已经获得了的价值，或者（3）它们是实现更远大价值的手段，那么，对这些价值确实可提出质疑。虽然人们对实际运用的推理法则一直有着广泛的共同见解，但长久以来，就应当以什么法则支配那些与相互关联的价值有关的推理这一问题，要达成一致意见，已被证明要困难得多。有些人已经为有关祈使句伦理句的推理法则，提出了好多种模态逻辑，但它们还都说不上

得到了承认，至于哲学界以外的应用，就更说不上了。①

不过，在过去半个世纪里，数理统计学家和经济学家们建立了一整套给人印象极深的形式理论，用不着引进新型逻辑，就可帮助我们对这类问题进行推理。这种理论的基本思想是，把一切价值纳入单一的函数之中，即构成所谓效用函数；靠这个办法，巧妙地解决不同价值如何彼此比较的问题。假如事件的每一种状态都已被指定了效用，那就等于做了这种比较。

这个形式化的理论叫作主观期望效用（SEU）理论。它的建立，是 20 世纪上半叶最令人振奋的知识成就之一。就应用推理方法考虑抉择问题来说，它是一部优美的机器。我们下面的任务就是要对它进行一番考察，并对它的成立性和局限性做出一些判断。

### 主观期望效用

由于文献中已有很多关于主观期望效用理论方面的精深而严格的论述，② 在这里我想仅对其主要部分做一简要的启发性综述。

#### 主观期望效用理论

主观期望效用理论首先假定，决策者有一个定义明确的效用函数，从而假定决策者能指定一个基数，用以衡量和评价其对未

---

① 在我的《发现的模型》（多德雷赫特〔荷〕，1977年）一书第3部分，以及我和西克劳塞合编的《表达与涵义》（英格尔伍德，新泽西，1972年）一书第8章《论有关行动的推理》一文中，论述了有关批驳模态逻辑学的情形。

② 例如可参见萨维奇的《统计学基础》（纽约，1954年）。

来事件任何具体可能状态的偏好程度。第二，这个理论假定，决策者面前摆着一组定义明确的备选方案，抉择就是从它们当中做出的。这些备选方案不一定用于一次性抉择，而是可以有一系列抉择，或者称为策略，其中每一步中间抉择（过渡抉择）都只是在某个特定时刻、利用那时具备的信息而做出的。第三，这个理论假定，决策者能够给未来的全部事件，指定一个具有一致性的联合概率分布。最后，它还假定决策者愿意（或应当）选取那种以其效用函数来看，能导致最大期望价值的备选方案或策略。如此看来，每个策略均与未来可能状态的概率分布联系起来了，用这些概率分布，就可以权衡未来可能状态的效用。

这就是主观期望效用模型的四大部分：一个基数效用函数；一组完备无缺的备选方案；一个与每一策略相联系的有关未来可能状态的联合概率分布；还有一个使期望效用极大化的策略。

### **主观期望效用理论的有关问题**

从观念上讲，主观期望效用模型是理应在柏拉图精神乐园中占有显耀地位的精美作品。但是，人要原原本本地用它去制定实际决策，却面临着多得无法克服的困难，因而是不可能的。关于这些困难，我过去讲过很多了（尤其是在《管理行为》一书中），所以我想只在这儿非常简略地提一下。

按照主观期望效用模型的假定，决策者总是用敏锐的眼光，对面前的一切都深思熟虑。他不仅明白自己当时面临的选择范围，而且对未来的小选择余地也了如指掌。他知道每个可供选择的策略所导致的后果，起码也能给未来可能状态确定一个联合概率分布。他协调了、或者说是权衡了一切互有冲突的局部价值，并把它们综合到单一的效用函数之中，按照对它们的偏好来排列所有未来可能状态的优劣次序。

这个模型十足巧妙地绕开了组成效用函数的价值的由来问

题。那些价值无非就是已经摆在那里了，并且已经组织成了统一的形式，表示对所有可选择的未来可能状态的始终一致的偏好。它还同样十足巧妙地避开了对有关现在及未来可能状态的事实加以确定的过程。这个模型顶多能告诉我们，如何从事实前提和价值前提出发进行推理，至于这些前提的由来，它什么也没说。

把上述假设直说出来，我们就可以清楚地看到，主观期望效用理论从来未曾、也根本不能用于实际——无论用不用最大型的计算机。尽管如此，在数理经济学、统计学和管理科学上，我们还是可以看到很多故意应用的情形。如果我们比较细致地检查一下就会发现，那些应用尽管保留了主观期望效用理论的形式结构，却换掉了该理论所假定的那种毫不可信的决策问题，也就是，要么是通过假想业已给定的效用函数和事件的联合概率分布而简化成几个方程和变量的高度抽象的问题，要么就是从大规模现实问题中割取极其有限的一部分，经过小心翼翼地定义而雕琢出来的微型问题。

### 作为某种近似的主观期望效用理论

既然我本人在管理科学研究中曾用过主观期望效用理论，就让我从自己身上开刀吧。霍尔特、莫地利安尼、默茨和我曾构造过一个关于工厂在不确定情形下决定产量、库存和劳动力的决策程序。<sup>①</sup>这个决策程序符合主观期望效用模型。其中，效用函数是由生产费用、改变产量所引起的费用、订货损失的费用和库存保管费所构成的一个费用函数（取其负值）。我们假定该效用函数是独立变量的二次型，因为这种假设对于设法进行数学推导和计算来说，是绝对必需的。我们还假定，未来每定期的销售

<sup>①</sup> 霍尔特、莫地利安尼、默茨和西蒙：《生产、库存和劳动力的规划》（英格尔伍德，新泽西，1960年）。（此书为控制理论在生产管理上的应用方面的经典著作——译注。）

量期望值都是已知的。（对二次型效用函数来说，这一假定幸而使我们不必知道完整的概率分布。）此外，我们又假定工厂仅生产一种产品，或者是可以合理地表示为单一总量的一组产品。

显然，如果用这一决策程序为某个工厂制定决策的话，那是与运用主观期望效用理论进行实际决策相差很大的。除了一个难点之外，所有难题均被已知二次型指标函数和已知未来销售量期望值等假定，事先解决了。此外，所考虑的生产决策，已经从整整一大批必须制定的管理决策当中分离出来了，并且假定，对于这种决策问题，可以完全不依靠有关其他决策或实际当中任何其他方面的信息，而独立地加以刻画。

我并不急于辩解说，我们这个决策程序是一个有益的管理科学工具。当然，这个工具能够、而且确实已经应用于许多工厂的实际决策任务，看来还应用得不错。但我想强调指出的是，它仅仅是应用于真实情形微小片断的高度简化的表达上。而且，用它所做决策的质量，对其近似假定及其支持数据恰当性的依赖程度，比决策质量对主观期望效用决策法则所规定的最大值计算的依赖程度，要大得多。因此，我们完全可以确信有人能设计出一个不在主观期望效用理论框架之内的、并与之颇为不同的决策程序，用它会制定出比我们那个决策法则所做的决策质量更高的决策。

对于在主观期望效用理论模子里铸成的经济模型，我们可以做出完全同样的评价。对这些模型的成立性和有用性，是不能从形式上符合主观期望效用假定这一点出发，给以评判的。在对它们进行评价的过程中，关键是要知道那些假想的效用和未来事件在多大程度上符合实际。

在任何实际应用上，主观期望效用法则都只不过是为某种抽象提供了一种粗略的近似办法，而抽象的结果对实际问题来说，可能提得出、也可能提不出满意的解答。一旦我们承认了这一

点，我们就可以随意问一问：与主观期望效用不相干的其他什么决策程序是否也能提供满意的解答？尤其是，我们当然可以问一问人们在决策制定过程中实际使用的是什么程序，以及那些实际程序同主观期望效用理论有什么关系。

我希望说服读者相信，在典型的真实情况下，决策者简直就不能应用主观期望效用模型，无论决策者多么想用，都是如此。读者要是对此尚存疑念的话，可以看一看被试人在类似于竞技、但又比实际竞技简单些的情况下做出含有风险和不确定性的决策，考察一下在实验室里进行的实验结果，就能驱散疑云。实验证据毫无疑问地表明，在给定的抉择情境中，人类行为出于某种原因而同主观期望效用理论的规定远远背离。这些证据，很多都收集在艾莫斯·特沃斯基及其同事的多篇文章中。<sup>①</sup>前面我已经指出了造成这种背离的主要原因。那就是：在应用主观期望效用原则所要求的种种条件方面，人既缺乏事实根据，也缺少一致的价值结构，又没有那么强的推理能力。即使是就那些相对简单的情形而言，也是如此。

我们的下一个任务，是要考察一下用什么来替代主观期望效用原则。

## 行 为 说

我想请诸位反思一下，你自己在实际中是怎样做出决策的？我还想讲几个论断，以便诸位能用它们检查一下自己的想法。其一，你的决策并不是在你的广泛生活范围内做出的综合性抉择。一般地说，你的决策只涉及非常具体的特定事物，并且假定这些事务同你生活中也许同等重要的其他方面相对独立，而不管这

<sup>①</sup> 参阅特沃斯基和卡尼曼，〈不确定情形下的决断：启发式搜索与倾向性〉，载《科学》185（1974年）：第1124—1131页，及该文引用的文献。

种假定是否正确。你在花钱买小汽车的时候，并不同时考虑下星期的餐点安排，甚至也不同时考虑你的攒钱计划。

其二，你在做一项具体决定时，可能不会仔仔细细地把未来可能状况搞清楚，并给你想选择的各方案的未来可能状态指定好概率分布。即便你的那项决定是非常重要的，恐怕也是如此。你对自己的生活方式和生活前景有着一般性的了解，也许还经过冥思苦索了解到一两个近期将会发生的大变动，甚至预料到一两件偶然事件。当你盘算买小汽车时，你对自己如何用车，自己的收入情况及其他方面的开支要求，以及是否想去外市谋一份新工作，都有着一般性的想法。你不大可能去料想那些同样可能影响你买哪种车的其他大量可能因素。

其三，你打算买车而不是房子，这也许使你的注意力集中在于你生活的某些方面和某些价值上了，同时相对忽视了生活的其他方面和其他价值。一心想买车，可能会勾起你对外出游玩的眷恋或向往，而暂时不再考虑欣赏立体声音乐或为朋友举行家宴的愉快。因此，你做决策的所有方面，不大可能体现在单一的综合性效用函数当中。相反，决策的具体方面，将要求采用具体的价值考虑。抉择当中存在的严重的不一致性，可能是由注意力的飘乎不定而引起的。大家知道，假如我们想吃饭，就得暂时抑制住自己想干别的事情的愿望，而去找饭吃。让我们的抉择置于那种综合性的、一致的单个效用函数指导下，看来既无必要，又无益处。

其四，你为做出买车决定而花的功夫，大多集中在收集事实资料和思考可能有关的价值上了。你可能会看一看《消费者报》，找朋友商量商量，向汽车商请教请教；以更多地了解可供选购的车型和你自己的偏好。一旦积累了这类事实资料，唤起了自己的偏好，实际抉择可能只须很短时间就做出来了。

### 有限理性

我刚刚描绘过的一般抉择方式，有时被刻画为有限理性的特

例。我们有充分的理由支持这样一种观点：进化过程可以创造出具备有限理性的生物。此外，心理学上的大量研究成果也支持我们对自身进行反思时所获得的那种感受，即：人类决策制定过程——纵然是最认真的决策制定过程——都是以那种方式进行的。让我们把关于人类抉择的这种模型叫作行为模型，以区别于主观期望效用理论那种奥林匹亚山神般的模型。

在有限理性的行为模型里，人的决策不是非最终达到极深刻的地步不可，不是必须囊括全部价值，并把每个问题都同世界上的其他所有问题联系在一起。实际上，我们生存于其中的这个环境，而且一切生命也皆生存于其中的这个环境，乃是一个可以近似分解为各个独立部分的环境。你有时感到饿，有时觉得渴，有时身上冷。幸亏你常常不是三者交迫。就算发生了这种情况，那你也总要先顾其中最紧要的一种需要，而把其他两种需要暂且搁在一边。此外，你还有许多别的需要，但它们也不是全都同时给你造成压力的。

我们生活在一个可说是近乎空旷的世界上，其中虽有亿万种事物在原则上可能是相互影响的，但它们在大部分时间里并非如此。按照万有引力定律，每个物体都受到其他一切物体的吸引，但有些物体要比别的物体引力更强，这或是因为前者质量更大，或是因为它们离得较近。世界上或许真的存在着十分致密的相互关联网络，不过，在我们所面临的大多数境况中，我们所能够察觉到的各种事物，或者优先考虑的东西，只是不多的一部分而已。

这种可分解性即便不是对我们今天生活着的世界的完整描述（但我想对这种说法做些保留），它也肯定描述了含有人类理性在内的另一个世界：原始人及其祖先们生活的那个世界。在那个世界上，除了必须采取定期行动去抵御饥饿、逃灾避难或防寒越冬之外，大部分时间里很少还有别的事情要做。在某一时间，理性只能用来集中处理一个或少许几个问题，同时也抱有这样的预

见，即：倘有其他问题发生了，还会有时间去处理的。①

### 有限理性的机制

生物要使自己有能力察觉到有限的现实，需要具备哪些特征呢？它要有集中注意力的本领，即能在一定时间里避免分神（或至少是避免过多分神），把心思集中在需要关注的事情上。我们可以提出一个很强的论点，那就是：注意力的集中，是我们称之为情感的那些过程的主要功能之一。生理心理学家已经指出了这一点。情感能改变你原来的思考焦点，将你的注意力集中到眼下看来需要注意的其他事情上去。在我们这个社会里，大家固然不必花大部分时间去寻找食物，但是，我们总还需要经常想起食品不可不备。可见，我们具备一种定期唤起饥饿感的机制，将我们的注意力转向对食品的需求。至于其他情感，也可以做出类似的解释。

生物还需要某些连续的活动。人需要呼吸空气——当然，吸取空气可以说是只要短暂时间的、间歇性的活动，但人的血液则必须连续不断地在全身循环运动。人体生理机能自然既涉及这类短暂需要，同时也涉及长期需要。但我们不必为呼吸空气或心房跳动，而留心注意血管里的氧气需要。总的看来，就那些间歇性的、并非片刻缺之不能的需要而言，我们人类很象是分期进行活动，即一时干一事的动物。那样一种需要的规模，大约就是我们的头脑在一时间所能把握的那么大。使我们得以侥幸打破这种限制，并在分期活动性的约束下得以幸存的能力，取决于能够确保紧迫新问题在议事日程上占优先地位的机制，尤其是那些情感机制。

其次，我们需要一种能产生备选方案的机制。在解决问题的过程中，我们所付出的很大一部分努力，都是在寻找较好的备选

① 我在《理性抉择与环境结构》一文中提出了这种理性的形式模型。[参见本书第三章]。

方案，或是在谋求现有方案的改进。认知心理学和人工智能过去25年来的研究成果，使我们懂得了备选方案产生过程的很多机制。我在《关于人为事物的科学》<sup>①</sup>一书第3、4章介绍了其中一些机制。

再次，我们需要一种了解自身所处环境的事实，并从这些事实出发进行推理的能力。当然，这种能力既可帮助我们接近事实的可能后果，又可帮助我们构造备选方案。它使生物能够对那些关系到其经常决策的部分事物，持有某种十分简单的模型，并就该模型进行常识性推理。

关于人类思维和问题求解的这个行为说，这个有限理性说，我们能对之做何评论呢？首先，现在已有大量有力证据表明，这个理论描述了人们实际做决策和解决问题的方式，作为对人类行为的描述，这一理论有着日益巩固的经验基础。第二，它是解释了以下事实的一种学说：无论我们这些生物是凭直感的机敏，还是凭推理和思考，同整个周围世界比起来，我们的计算能力是很有限的，但是，我们毕竟生存着，甚至还日益兴旺发达。这一学说解释了我们这些创造物是如何生存下来的，至少解释了在人类产生后的几百万年间，我们是何以幸存的。在一个近乎空旷的世界上，在一个并非万物致密相关的世界上，在一个可将问题分解成若干小问题的世界上，正是我所描述的那种理性，才使我们得以幸存。

### 有限理性的重大意义

行为模型所描绘的这种理性，当然不是寻求最优的。它甚至不能保证我们的决策是前后一致的。事实上我们很容易证明，具备这些特征的生物所做出的抉择，常常取决于提出备选方案的顺序。如果备选方案甲是先于乙提出来的，我们可能会把甲看成是

<sup>①</sup> 第二版[参见中译本修订版]。

理想的，或至少是满意的；但如果甲是在乙之后提出来的，那么，我们甚至在没有想到甲时，便将乙视为理想方案而予以选用。

行为模型摈弃了奥林匹亚山神式模型的许多形式上的漂亮性质，但作为回报，它所提供的看待理性的观点，向我们解释了这样一个道理，那就是：拥有我们这种思考能力的生物——或者说，拥有我们这种思维能力并辅以硅谷全部计算机的生物——在主观期望效用理论那种奥林匹亚山神般观点看来复杂得无法理解的世界上，是如何侥幸生存下来的。

## 直 觉 理 性

社会科学家们对下述关于人类理性的第三种模型的讨论，要比他们对上面讲过的那两种模型的讨论少得多。不过，这第三种模型在大众心目中似乎更盛行。我把这种模型叫作直觉模型。按照这种模型，人类的思维和人类做出正确决策的本事，大多被归功于人类具有良好的直感和判断能力。由于罗杰·斯珀利①等人的研究工作的激发，直觉和判断的观念当今在舆论界尤其流行。那些研究工作在关于人脑左右半球分工的猜想之下，被人们添枝加叶地大大扩充了一番。

### 人脑的两半球

在某些作者看来，大脑半球分工的观念已成为一种罗曼司。按照他们那种传奇般的解释，脑的左半球沉闷而单调，但分析能力极强。它要么能进行我在前面讲过的那种奥林匹亚山神般的推理，要么能进行上文所述第二类模型那种行为式的思考（如果我们讲的是一个笨人的左半脑）——这两种情形全凭诸位自行决定了。但无论是其中哪种情形，左半球总是一种严谨求实、平淡无

---

① R.W. 斯珀利是美国著名脑科学家，1981年诺贝尔奖获得者。——译注

奇的半球，它也许能进行很深刻的分析，但不能进行奇妙的畅想。于是人脑又有右半球，其中蕴藏着人类的想象力和创造力——蕴藏着解释人类创造本领的一切好东西，人如果依仗这个半球，就能创造性地解决问题。

在着手正面刻画直觉与创造性（注意这两者并不总是同一个东西）之前，我必须先对以上采用漫画手法勾划出来的罗曼蒂克观点评论一番。当我们为这种观点寻找经验证据时，我们会发现自己一无所得。半球分工的证据当然很多，但没有任何证据能真正说明，复杂的人类心理作用在正常情况下是由哪个半球单独干的。总的看来，证据告诉我们：包含吸取信息、处理信息和利用信息等过程在内的任何复杂思维，均以多变的比重和种种方式，运用我们的两个半球。

当然，脑的分工并不是最关键的问题。无论两个半球所发挥的作用是否相同，重要的问题在于，是不是存在着两种截然不同的人类思维方式——分析式思维和直觉思维，另外，我们称之为创造性的那种本领，是不是主要依赖于后者。

### 直觉与再认

直觉究竟是什么东西？它无非是人们有时突然发现了问题的答案。这是一种可以观察到的事实。由于这个原因，人们有着不同程度的直觉感受：“哎，原来如此！”这种现象无疑是真实的。此外，一个人如具备这类经验，那么，当他做直觉判断时，他所获得的问题答案往往也是正确的。

目前我们有好多材料，说明了国际象棋大师们的这种直觉本领。我们从一般棋赛中取一中盘阵势，让一位象棋大师或超级大师看一眼。他只消看上5秒或10秒，一般就能提出一个很好的走法——而且常常是该局势下的实际最佳走法。如果棋逢对手，旗鼓相当，他不会马上就走那一步，他可能会坐在那儿思考三分钟或半小时，以决定他的最初直感是否果真正确。不过，最初直觉到

底还是没错的情形，恐怕十有八九。

关于国际象棋大师的正常直觉，其解释已为心理学家们所熟知，实际上也并不奇怪。①要是你明天在去上课的路上碰到一位朋友，你就会马上认出他来。除非你边走路边沉思，否则你不仅能立即认出他，而且不会认错。对象棋大师直觉本领的解释，同对你的上述本事的解释相比，毫无深奥可言。在通过学习业已获得很多经验的任何方面，我们都结识了众多“朋友”——可以立即再认、识别出来的大量刺激。我们能把这些刺激分类，存放在头脑中起这个作用的某种分类网络当中（其生理结构尚未搞清楚）。而且，一旦碰到一种刺激，就能从全部刺激中，把它辨认出来。我们不仅能按照面貌、外观这么做，而且能用母语词汇这样做。

每个受过高等教育的人，差不多都能识别 5 万到 10 万个不同词汇，并回想起它们的含义。我们成年累月地看很多很多词，年复一年，就结交了 5 万或 10 万个朋友。昆虫学家具有辨认昆虫类别的卓越本领，植物学家具备辨别植物纲目的非凡能耐。在任何专门知识领域里，正是由于有了一张十分精致的辨识网络，人们才能够从千万种不同事物、不同情境当中，判别出其中任一事物或情境，这种网络是专家的基本工具之一，是其直觉的主要源泉。

我们已经数过国际象棋大师所熟悉的“朋友”的数目——他们所熟悉的各不同棋势图案的数目。这个数在数量级上的估计值，大约是 5 万，与使用母语讲话的人所用的词汇量大致相同。直觉，就是认得出一个朋友，并将你过去对他的所有了解，都从记忆中查找出来的能力。当然，如果你对他了解得很深，你就能对他做出很好的判断。借不借钱给他？等你需要时，还要得回来吗？如果你深知其人，你就能凭直觉做出回答。

---

① 有关这方面证据的综述，请参阅我的《思维模型》第 6.2—6.5 章。

## 直觉和判断力的获得

我们为什么认为，再认机制解释了有关创造性的文献中所讲的大多数“哎，原来如此！”之类经历呢？这里有一个重要理由，那就是：真实的这类经历，只有那些具备适当知识的人才能获得。彭加勒曾正确指出：灵感只光顾那些有准备的头脑。今天，我们甚至已经掌握了一些数据，表明为准备和造就一个达到世界水平的具有创造性成就的头脑，需要花多长时间。

为达到世界水平的成就所需要的时间，在不同领域里大致相同。为什么会是这样？乍看起来有点令人迷惘。但无论如何，人的素质高低，是通过人与人之间的比较来评价的。因此，人生的长短是竞争中的一个控制参数。我们可以花掉一生的很大一部分时间，用以增强自己对知识的精熟程度，但时间不能更长。由于这个原因，为达到世界一流水平做准备所需用的时间（对于其天赋允许他们立志达到该水平的人来说），在不同的活动领域里，应当大致相同。

我的同事约翰·哈耶斯所收集的有关象棋大师和作曲家生平经历的实际资料表明，这个神奇的数字是10年。他所收集的有关画家和数学家生平经历的资料，虽然不那么系统，也表明了同一数字。在这些专业领域，差不多没有任何一个人不是首先经过了至少10年光景的紧张学习和勤奋实践，才有可能达到世界水平的成就。

神童的情形如何？莫扎特谱写出世界水平的乐曲，也许是在17岁，而不是更早。（哈耶斯在音乐方面所采用的标准是，在施瓦恩分类目录中出现5条或5条以上的乐曲记载。除了有些莫扎特少年时代的作品之外，莫扎特17岁以前创作的乐曲当中没有任何一部是世界水平的。那些少年时代的作品，也只是因为出自莫扎特之手，才有人愿意听。）当然，莫扎特在4岁上就开始谱曲了。因此，到17岁时，他已经在音乐方面学习了13年。莫扎特的例子，在哈耶斯所考察过的神童传记当中，是很有代表性的。由

此可见，在一定领域做出卓越成就，有一个必要条件，即：要在该领域里勤奋学习10年，或更长时间。

### **概要：直觉模型与行为模型**

思维的直觉模型同行为模型没有任何矛盾。这两者所描绘的思维方式，并不是各在人脑一半球的、争着控制心理活动的不同思考方式。任何严肃的思维，都需要那两种风格，它既是搜索过程，又需要对熟悉的模式做出快速识别。如果没有基于先前经验的识别或再认，在复杂空间里的思索便会象蜗牛爬行那般缓慢。直觉能使我们得到那些通过以往搜索已经掌握了的知识。由此我们可以料到实际发生的情形，如，专家在解决一个需要费劲寻找新东西的问题时，常常能靠直感取得进展。我们也可以料到，在包含新、老因素的问题情境下，求解过程大多是直觉与搜索的协作过程。

### **直觉与情感**

到此为止，我们在讨论直觉过程时，还没有讲到据说是这类过程所具备的另一个重要特征：它们往往和情感联系在一起。问题求解中的搜索过程、辛辛苦苦寻找答案的阶段，看来同强烈情感关系不大，我们可将其描绘为冷静的认知。但是，那种突如其来的发现，“哎，原来如此！”之类的体验，看来是要借助于情感的，这是一种热烈的认知。有时，当我们对某个事物颇感激动的时候，就会产生一些思想，想出某些主意。

### **情感与注意**

因此，要得到完整的人类理性理论，就必须理解情感在理性中所扮演的角色。情感很可能起着好几种不同的作用。首先，有些情感（如快乐）是消费品。它们出现在奥林匹亚山神般理论的

效用函数当中。在行为模型中，它们也必须考虑到我们力求实现的诸目标当中。

不过，就目前的讨论来说，情感之所以具有特殊的重要性，是因为它能从我们的环境中选出一些特殊的事物，当作我们的注意焦点。雷切尔·卡森的《寂默的春天》<sup>①</sup>一书，为什么会有那么影响？她在该书中所描绘的问题，生态学家和其他生物学家们早就知道了。但是，她的描写方式富有感召力，能鼓动起我们的情感，能把我们的注意力吸引到她所提出的问题上去。那种情感一旦被激起，便抓住了我们的心，使我们感到只有为这个问题做些什么事情之后，才会去关心其他问题。至少，情感能使那个问题存留在我们脑海里，成为一种撵不走的心病。

在奥林匹亚山神般的模型中，所有问题永远同时予以考虑（直至解决了为止）。相反，在行为模型中，议事日程上的问题抉择是一件最重要的事情，情感对这种抉择起着相当大的作用。

应当注意，情感并不是总能将我们的注意力引向那些我们认为是正当的目标上去。回顾一下前面讲过的《我的奋斗》的例子，我们可以看到，那本书中的推理不是冷静的，而是热燥的。那种推理所极力追求的，是要鼓动起强烈的情绪，而且往往是仇恨——一种很强烈的人类情感。当然，象《寂默的春天》或毕加索的《格尔尼卡》<sup>②</sup>一样，《我的奋斗》一书的影响，很大程度上是由于它有感召力，能唤起德国读者的注意，并将他们的注意力栓在该书所述的特殊目标上。

行为理论由于把注意的焦点当作了抉择的一个重大决定因素，因此并没有把情感同人类思维割裂开来，也丝毫没有轻视情感对人类问题求解的议事日程的强烈影响。

---

① 此书1962年出版，生动描绘了有毒化学制品对大自然的破坏作用及严重后果。此书有中译本。——译注

② 毕加索名画。格尔尼卡是西班牙北部小镇，内战时曾遭到佛朗哥军队狂轰滥炸。该画即描写当时惨景。——译注

## **情感在教育上的作用**

在此我想稍微偏离一点主题，简要说说情感在教育上的作用。如果说，文学著作或艺术著作有着相当强的感召力（它们确实如此），那么这是否告诉我们，那种感召力在教育过程中能起到什么特殊作用？

大家知道，文学现在有点遭受冷遇。在我们的大学里，有很多学生要求选修法律、企业管理或医学，而文学课程则受到或轻或重的忽视。指责这一倾向的人常常提出这样一个论点，即：学生们通过了解艺术家和文学家看待世界的观点去了解人类状况，要比通过了解科学家的观点更好些，或者说更加有效。当然，我本人的专业身份使我难以与此苟同。但我认为，我们应当更细致地看一看这个问题。人们有效地了解重要问题的最佳条件是什么？是冷静的认知好些，还是热烈的好些？如果其中之一更好些，我们能说它就是同科学学科或文学学科相联系的吗？

这里我想先说一下，我曾听说过物理学家主张在讲授物理课时运用很强的热烈认知。使他们激动，并且激励他们去理解那些深奥事物的，是同基本粒子、天体物理、宇宙构造学说等相联系的宇宙论问题和哲学问题。因此，我也许不该严格地将科学同冷 静认知扯在一起。

不过，还是让我讲一讲能够更清楚、更方便地说明这一点的某些方面吧。诸位当中也许有人读过亚瑟·凯斯特勒的《中午的黑暗》一书<sup>①</sup>。这本小说描写了一个在俄国30年代清洗时期的遭遇。假如你想了解一下两次世界大战之间的西方历史，以及向当今世界过渡过程中的那些事件，你肯定需要了解一下清洗运动。你是愿意读一读《中午的黑暗》去了解它，还是愿意读一读有关清洗运动的历史书，或是到图书馆去查阅有关清洗运动的原始

---

<sup>①</sup> 凯斯特勒是英籍匈牙利小说家（1905—1983年）。《中午的黑暗》是他的一本著名政治小说。——译注

资料抄本？我以为，读凯斯特勒的书，是一个最好的途径。这完全是因为它能唤起大多数读者的强烈情绪。

我还能列出这类图书与有关学科的一张长长的清单，如：《战争与和平》和军事社会学论著，普鲁斯特、契诃夫的小说和有关人情事故的教科书，等等。假如我的专业身份使我必须站出来维护文学在教育中的地位，维护 20 世纪初那种传统的人文、艺术课程，那么，我的论点将基于如下立场，即：如果我们在富有情感的内容中阐明有关问题——给它披上一层感情炽烈的色彩，同完全没有感情色彩的平铺直叙相比，前者能使大多数人更长久地关注有关问题，更认真地思考它们，印象更深刻、更持久。

另一方面，借助热烈认知进行教育，同时还意味着一种责任。小说家要想让我们通过阅读他们的作品，去了解社会科学知识，就必须正确地描写。其科学方面的内容，必须是站得住脚的。正值弗洛伊德心理学受到心理学新知识的根本修正的时候，今天却有大量文学作品依然抱残守缺，大讲弗洛伊德理论。当今，心理学界已经没有几个正统的弗洛伊德主义者了。因此，如果我们在这方面请求文学作品为我们提供学习心理学的情感环境，那就存在着这样一种危险：陈旧的弗洛伊德学说，会对我们的学生造成过强的影响。我们必须重新评价那些有名的古典人文著作，看看它们在多大程度上，已被科学知识的进步所抛弃。

荷马现在仍然活着，因为《伊利亚特》和《奥德赛》所讲的，主要是这样一些事情，在这些事情上，现代社会科学还没有进展到远远超过一般理解的地步。亚利士多德算是勉勉强强活下来了——他的科学著作肯定已经失去了生命力，他的逻辑学也几乎死去了。至于他的认识论或形而上学今天对学生们是否有什么价值，我们可以同哲学家们大大争论一番。琉克理细阿<sup>①</sup>固然

<sup>①</sup> 罗马诗作家（公元前96—55年），——译注

讲过原子，但他已经完完全全死去了。

以上论述的寓意在于，我们如果想把那些能激发情感的著作用之于教育，惟其具备那种能力，它们才对我们有特殊价值——尽管如此，我们不仅必须评价它们激发情感的能力，还必须评价它们在讲到事实时，在科学上是否站得住脚。

因此，人文文学如要把它在人文课程中占据中心地位的主张，建立在它们对人类状况的特殊观点上，它们就必须有能力证明，它们对人类状况所描绘的图景，从生物学、社会学、心理学上讲，是站得住脚的。在这方面，文学作品光能激励学生，是很不够的。它们必须这样激励学生，使学生们能正确认识现实世界中的事实和道理。我不是说，人文文学现在不符合这一标准，对现有各大学文科课程进行仔细评价，肯定会告诉我们，在这个问题上不存在一个简单的肯定回答，或否定回答。我所要指出的是：当我们对不同方面的知识为文科教育所提供材料的恰当作用进行考察时，既要周密注意材料的情感热度，也要认真留心其客观真实性。

## 结语

在这第一章里，我逐次向大家讲了三种理性说：谈论合理性的三种方式。第一种理性说是奥林匹亚山神般的模型，它假定有一种非凡杰出的人，在完整一致的宇宙当中做出全面抉择。这种奥林匹亚山神般的观点也许可以当作老天爷的心理活动模型，但它肯定不能当作人的心理活动模型。就现在的论题而言，我对这种理论采取了十分严厉的批评态度。

第二种理性说是行为模型，它假定人类的理性是非常有限的，受到情境和人类计算能力的很大限制。前面我已指出，目前有大量实验证据表明，这种理论真实地描绘了人类是如何做决策的。这种学说告诉我们，包括人类在内的、具有有限计算能力的

生物，是如何在一个错综复杂的、但大多相当空旷的世界上，做出适应性抉择的，以及他们是如何在这样一个世界上获得生存机会的。

第三种是特别强调直觉作用的直觉模型。我已经指出，直觉模型实际上是行为学说的一部分。它所强调的再认过程，是以一种技能为基础的。人类能通过存储经验事物和在适当情况下再认一定情境，获得那种技能。直觉理论承认人类思维常受情感影响，并且提出了这样一个问题：人的注意力在某一特殊时间被集中在某一特殊问题上的过程里，情感起到了何种作用。

在下一章，我要讲一讲第四种理论：将理性视为渐进适应性的观点。这种进化论模型是一种幕后背后的理性模型，它认为，只有那些寻求适应的、仿佛具有理智的生物，才可能生存下去。这类观点认为，自然选择在人类理性运用上很有效力，并占有中心地位。对此，我将在下一章里予以考察。

## 二 理性与目的论

我在上一章里考察了三种理性概念，但考察的重点不是理性本身，而是实现理性的过程。在那一章里，我们重点讲述了藏在判断与抉择背后的思维过程，以及多年来由关注理性的人们所提出的各种决策过程模型之间的差别。

### 作为理性适应的进化

在各种进化理论当中最为引人注目的，正是这种着眼于进化结果的理性观。进化论用以解释事物如何如何的办法，是指出生物为了生存就必须如此这般。生物究竟是怎样达到适应状态的，当然也是科学兴趣的所在。但从进化观点看，适应或生存的基本事实乃是第二位的。就其注意的是结果而言，这样的理性论，可以和上面所讲的奥林匹亚山神式的过程、行为过程、甚至直觉过程，协调一致。

例如，鸟儿在树上筑巢之所以可被认为是理性的，是因为它们所选择的那个位置有助于保护鸟卵和幼鸟不受地面食肉动物的攻击。这绝不等于说，成鸟是经过了某种决策过程——奥林匹亚山神式的、行为式的或直觉式的决策过程，才做出了选择那个位置的决定。筑巢行为无非是一种本能行为，无非是经过进化过程，淘汰了其他适应性较差的行为方式，从所有行为方式中选择出来的高度适应环境的行为。正是在这个意义上，我们才可以把进化过程的产物，视为理性的一种形式；也正是在这个意义上，我们才确实可将其视为上一章所讲的各种理性形式的一种替代形

式。

### 关于适应性的“仿佛”论

我们之所以关心从进化论立场出发去看待和解释理性的作法，其原因之一在于，有些社会科学家（特别是有些经济学家）认为：了解人们如何做出决策的过程，是不重要的。他们认为，这种了解之所以不重要，是因为我们从人们确已幸存的事实可知，人们实际上等于做出了理性的、适应性的决策。

米尔顿·弗里德曼关于方法论的著名论著，就采用了这一观点。<sup>①</sup>他在那本书里提出了一种关于经济行为的“仿佛”论：企业家和厂商的行为，仿佛是为了一定目的，比如说为了谋取最大利润或最大效用，而做过了正确的、合理的计算。这一论点的依据是：只有在寻求最优的过程中获得成功者，才在企业界站住了脚，其余则从企业舞台上被淘汰掉了。这种观点所关注的，完全是结果——对经济环境的成功适应。它完全没有考虑，究竟是什么样的推理过程，造成了适应——或者就那一点而论，究竟是什么样的随机过程，造成了适应。

我们对世界的运转方式的兴趣所在，如果不限于仅仅关心厂商行为在公共政策方面的意义，那么，上述“仿佛”论的回答，就不能让我们感到十分满意，即使我们知道或相信，只有好象进行了理智计算的那些企业才生存下来了，我们对他们究竟为生存而做了些什么这一问题，仍然是想要追究一番的。在我们这个世界上，奇迹和壮观的巧合也许确实是不断出现的，但是，说不定还有某些内在机制，支配着生存现象。理解这种机制，我们就能更好地判断，怎样就可能使一个演化着的系统，保持在均衡位置附近，以及该系统对均衡位置的偏离，是否可大到足以明显影响政策的地步。

<sup>①</sup> M·弗里德曼：《实证经济学论文》（芝加哥，1953年。）

## 变异与选择

除非我们能确认所达到的均衡是唯一的，否则，“仿佛”论的回答对上述两个问题来说，都不能令人满意。如果说，不同的过程会导致不同的均衡（后面我们将会看到，事实很可能就是如此），那么，在理解现象本身和引出其政策意义这两个方面，过程就再次成为主角了。

在现代达尔文主义的生物进化论中，尽管所强调的是结果（即幸存或生存的事实），但是，这种理论当然并没有把结果设想为奇迹，而是认为存在着一种机制——或者更准确地说，至少是两个机制的组合。这两个机制，就是**变异**和**选择**。前者创造新物种；后者保留其中那些适应环境者。<sup>①</sup> 在这些达尔文主义机制和上一章所描述的行为理性论基本机制之间，有一些很有意思的类同性。

按照行为理论的观点，为了发现适应性的反应策略，理性抉择可能需要经过大量的选择性搜索。最简单的、最基本的搜索过程，要求首先生成和构造出可能的反应，然后检验其适宜性。行为理性论的发生器——检验机制，达尔文进化论的**变异**——是选择机制的一种直接类比。正如生物进化论认为**变异**生成了新物种，关于人类理性的行为理论则认为，存在着生成备选方案的某种作用——也就是，能构造一些简单的想法和主意，并以新的方式将它们联系在一起的某种组合过程。同样，在生物进化论当中，自然选择机制将适应性较差的物种淘汰掉了；与此完全类同的是，对人类思维来说，检验过程把无助于解决问题的想法排除掉了。<sup>②</sup>

① 我之所以采用**变异**的说法，而不用突变一词，是因为，突变仅是创造新物种的达尔文主义机制之一（而且，也许还不是最重要的）。经典遗传学中的另外两个机制，是染色体交杂和倒位。但是，在所有机制当中，最根本的一个机制，就是基本生殖循环，它通过卵的成熟分裂和受精作用，在每一代都产生新的染色体组合。后面我还要更多地讲到这些机制。

② 参见《关于人为事物的科学》第7章“复杂事物的构造”；关于人类思维方面的情形，可参阅该书第3、4章。——译注

在心理学家中间，最有说服力地指出并论证了达尔文进化论与行为理性之间类同性的，是唐纳德·坎贝尔。<sup>①</sup>在经济学界，理查德·纳尔逊和西德尼·温特也以类似的方式，研究了这个课题。他们试图确定一些能解释厂商的进化与适应的机制。<sup>②</sup>纳尔逊和温特的所谓“基因”，指的是企业在处理其事务时所采用的惯例、常规和标准业务程序。随着时间推移，有些新的惯例被想出来了，并且必定要与老惯例一比高低。以这一过程为依托的理性，也是与行为理性极其相似的。因为，这里没有任何东西保证系统能达到或趋于最佳状态。它是适应性的，但未必是寻求最优的。

本章开场白要讲的最后一点是，我们在把进化概念应用于人类社会时，需要稍许留心一下理论的统计假设。在我们那个选择最适者的抽象模型当中，存在着这样一种设想：在接连几代里，有好几个变种被生成和检验，但在选择过程中幸存下来的，只是其中之一。如果我们要把进化概念应用于当代社会，那我们就必须问一问：通过反复尝试进行的统计选择是否可行。例如，要进行多少次核爆炸才能确定，哪些物种的适应能力足以在核爆炸后的地球上继续生存？我们当中有许多人认为，要回答人类在核爆炸过后的生存问题，连一次爆炸都不可做。可见，把大数定律或系列检验概念应用于核事件，没有什么意义。

我们看到，有些实验是难以在统计基础上进行的，因为诸位的生命只有一次。如果你把进化看成是一个试误过程，其中的“误”是由选择力排除掉的，那么，当仅有一次尝试而不可有一次失误时，你的那个模型就不合适了。我们在把进化论模型应用于人种与人类社会的未来发展时，必须记住这一点。

### 达尔文模型

上面我已经讲到了达尔文主义进化模型的几个基本机制。交

① 可参阅坎贝尔：《进化认识论》，载P.A.Schilpp（编）：《卡尔·波普尔的哲学》（拉萨利，伊利诺伊，1974年），第1编，第413—463页。

② 纳尔逊和温特：《经济变革的演化理论》（坎布雷奇，1992年）。

产生新物种，选择过程则对这些物种做出评价，确定其中哪些物种将会续存。有关达尔文主义进化的标准文献，尤其是那些形式化的文献，集中讨论了所谓适应（fitness）的概念。假设有两种生物，都试图在同一生态环境（小生境）里生活，也就是说，都试图利用完全同样的资源，那么，从平均意义上讲，其中一种生物的后代存活率，将比另一种生物的后代存活率要高。后代存活率比较高的生物，便是比较适应的物种。如果这种生物的后代仍能战胜另一种不太适应的生物的子孙，那么，前者的数量很快就会远远超过后者的数量，而且，由于资源总量有限，前者终将把后者逼进灭绝境地。

## 适 应

由此可见，达尔文学说的核心，就是用适应去解释一切的观念。一切问题都归结为，一种生物能否在生殖方面战胜对手。因为，能够最有效地占据某一小生境的物种（这里的“有效”被定义为适应），就是能生存下去的物种。

根据复利计算方面的知识，我们可以轻而易举地计算出，每代仅以 $1.05/1.00$ 的微弱优势胜过其竞争者的某种生物，其子孙数量在以后14代里，就将达到其竞争者后代数量的两倍。如果说，人类在几百万年以前就开始进化了（这看来是真的），那么，人类的适应优势大约已经经历了十几万代。当然，我们人类从狩猎生活方式到耕作生活方式的转变，至今仅有四五百代，但即便是这段时期，也可以发生大量的选择性变化。在四百多代的时期里， $1.05/1.00$ 的适应优势，就可以造成子孙数量上 $250\,000/1$ 的优势。即使是 $1.01/1.00$ 的优势，也可造成 $13/1$ 的子孙数量优势。

另一方面，如果我们认为现代工业社会使人类面临着的条件，与人类在农业社会中的成功和生存条件很不一样，那么，我们还要认识到，这种新条件的生效时间只有十来代（此为最宽的估计）——这样一段时间，还不足以带来实质性的选择性变化。

总之，当我们说到选择力对我们人类的影响时，我们必须要明确，所考虑的是人类历史的哪一阶段，是处在极其原始的条件下的早期发展年代，还是千万年的人类农业社会；抑或是不足两个世纪的现代社会。我们对进化力所带来后果的估计，在很大程度上，将依赖于我们对新条件的认识。也就是，我们认为可能出现适应性特征的新条件同老条件有多大差别。

把最适者生存按上述方式加以概念化，便引出了所谓利己基因的概念。（如果我们可以对基因稍微做些拟人化处理，那么，按照这一概念，我们可以说：）一种基因除非是尽其所能地去适应，否则就一事无成。它如果做适应之外的事，那就会断送其生存机会。因此，那种看起来是以牺牲自己的适应为代价，而为别人寻求福利的基因——利他基因，是很少能在自然界里发现的反常的东西。只有在非常反常的情况下，才能碰到那种利他基因，稍后我将讨论这一点。上述模型中的核心机制，就是争夺小生境。

### 小生境的构筑

对进化过程所做的实验研究，无论是野外考察还是实验室研究，大多是为了理解争夺小生境的现象。但是，除了弱肉强食式的争夺之外，小生境竞争理论还有另外一种观点。这种观点有时是同德克亥姆的名字联系在一起的，<sup>①</sup> 但也可在《物种起源》一书中找到。这种替代观点，是观察到了这样一件事实而产生的，那就是：动物在丛林环境中寻求生存的方式有两种，其一是同试图占据现有小生境的其他种类的动物厮杀格斗，成为竞争中的胜利者；其二则是发现一处完全未被占据的小生境，或者是改造自身，使之具备专门的适应性，使自己能够有效地（适应地）占据一处已被别种动物低效占据了的小生境。

<sup>①</sup> 埃米尔·德克亥姆：《社会中的劳动分工》（格伦科，伊利诺伊，1947年），第2卷，第2章。

我们可以想象出这样一个系统，其中，不同种类的大量生物能够共同生存，因为那里的小生境系统已经增生，使得每一种生物的自身生境都没有多少裂缝。也许还有附加的小生境，它们至少暂时未被占据，或仅仅被那些还不是专门适应该生境的生物低效地占据着。系统增生的一种方式是：两种专门适应的动物，取代一种无事不做的动物（比如说，在原来只有一种中型动物的地方，出现一种大动物和一种小动物）。这种增生现象，在关于海岛上关系紧密的各种动物群体的研究工作中，已经被观察到了。对于获取特定食物来说，两种更加专业化的动物，每一种都比单独一种“事事都做”的动物更有效率。其中比较小的动物，一般以小动物为食；比较大的动物，则一般以大动物为食。那些专业化的动物，既可以独立地进入某一生物群，也可以通过原始种群的变异和自然选择而出现。

在群体遗传学文献中，小生境构筑理论，一直不如多种生物争夺单一小生境的理论发展得那样充分。前者很可能要比后者复杂得多，因为它必须既要解释小生境的增生，又要解释进入小生境的生物的激增。此外，在这样一个系统中，每种生物所处环境的一个重要部分，乃是由其周围的其他生物提供的。这种系统得到改造的途径在于，小生境的真正构成，以及进入小生境的新物种的最终发展，允许构造和发展更多的小生境。

蚤的进化和生存有一个先决条件，即：有狗之类的动物为其提供生存环境。任何动物得以进化的先决条件是要有作为食物的植物，以及那些植物生存的小生境。因此，进化论看来有两个颇为不同的研究方向。其中，范围较窄的理论可集中探讨适应问题——争夺单一小生境或固定小生境系统的问题。经典理论就是如此。但我们还需要一种更为广泛的理论，这种理论认为，小生境系统本身，是与其中的生物同时变化和发展的。这后一种理论的

发展，现在尚处于非常原始的阶段。

因此，要解释今天生活在地球上的物种为何如此繁多（其数量达数百万），起码有两种不同的、但不一定相互排斥的回答。第一种回答是：由于地理环境和气候的多样化，各种生物所能分别适应的不同的微观环境（小生境），可以有亿万种之多。随着新物种不断出现的变异过程，这些小生境可能被逐渐地占据了。有些小生境可能老早就被现在的居住者所占据，但还有大量小生境尚未被占领，或只是被并非专门适应该环境的生物，以比较低的适应水平占据着。

另一个回答所描绘的图景是：地球起初以无机物质占多数，它所提供的不同种类的微观环境是比较有限的，随着新物种的产生，不断出现了新的环境，和环境之间的新差异。这幅图景如果成立，即使是部分地成立，物种激增也可以是无穷无尽的。另一方面，那种小生境数目不变的图景如果成立，我们就会看到，新的、更适应的物种，在进化过程中迟早要把老的、不太适应的物种，逼入灭绝境地。

这方面的证据是自相矛盾的。一方面，据说地球上曾出现过的所有物种，99%以上现已灭绝。另一方面，对当今存在的大量物种，均可在亿万年形成的化石当中找到其同类，或极其相似的同种（这个事实，曾一度被认为是对达尔文主义的严重打击）。人们可以说，有些生物很早就具备适应性了，而且这种适应性还未受到过严重挑战。不过，这毫不妨碍寻找新生境的，即寻找未被占据的生境的大量新物种的出现。这些事实给我们提供了一幅进化史图景，它同那种朴素的“为生存而斗争”的图景，相去甚远。

## 变 异

达尔文所提出的，是进化的选择机制。他没有提出产生变异的具体机制。现在我们对变异是怎样产生的，确实还只有很不完

善的了解。适应性告诉我们，为什么比较好的生物一旦出现就能生存下去。但是在参加竞争过程的高级动物的起源问题上，它没有给我们提供任何线索。然而，没有候选者的来源，怎么会有选择作用？

染色体的发现，以及染色体经过每一代减数分裂而完成的重新组合的发现（至少对有性生殖来说存在这一现象），使科学家们了解到产生新物种的可能机制。然而，20对染色体中（个别生物的染色体数目还要多些），如果每一对都以双等位基因的形式存在，最多仅可产生 $2^{20}$ （约1百万）个不同的种。这个数目还不够大，还不足以解释长时期的重大有机进化。照这个数字，自然界可能出现的物种很快就会用光了。

关于染色体内基因组合和个别基因突变现象的进一步发现，使变异所可能发生的范围，大大拓广了。不过突变比较少见，而且大多数突变并不是适应性的。对于突变本身能否成为一个充分有力的变种机制，生物学家长期以来一直持怀疑态度。

第三个发现更有意义。这个发现就是：染色体不能作为重新组合的单位。相反，单独一个染色体中的基因之间，则可通过交杂、倒位或其他重组过程，发生大量的重新组合，这种重新组合可以进行得非常彻底。例如，可包括整段DNA的复制或消失。这一微观重组的事实，再一次大大增加了变种的可能范围。复杂生物的染色体一般至少载有10 000种基因（有可能多达100 000种）。如果每种基因以等位基因形式存在，那么，就可以产生 $2^{10000} = 10^{3000}$ 种不同的物种。即使是全部地质年代，也不足以完全产生全部这些潜在物种——甚至不能产生其一小部分。此外，重组的发生并不象突变那样少见，有些重组（如交杂）发生得很频繁，它们在每一次减数分裂引起的细胞分裂过程中均有发生。

这样，即使突变不经常发生，正常的细胞分裂和复制过程所引起的DNA的改变，似乎也可提供足够多的变种，其数目之大

足以解释进化过程中出现的许许多多物种。此外，由于在这种情形下，可能产生的生物在遗传学上只有很小一部分在组合过程中实际出现，并受到适应性检验，因此，我们今后应当讲比较适应者生存（较适者生存），而不是最适者生存。看起来，根本没有什么永久性的物种平衡状态，因为新的有效竞争者可以在任何时刻产生于生殖过程当中，而且大多数潜在竞争者还从未产生过（并且永远不会产生出来）。这进一步向我们表明群体遗传学为什么需要注意物种产生的动态过程，而不是仅仅关心现有生物为占据小生境而斗争的结果。

### 表型与遗传型

生物的外形和行为（即表型），是在作为其基础的遗传结构（遗传型）的强烈影响下，才形成的。但是，表型与遗传型之间的对应关系很复杂。自然选择作用于表型，在环境中参加竞争的，也正是表型。遗传变化则作用于遗传型。现在，所有证据都否定了拉马克主义的观点，即：一个表型的经历或变化，可以直接改变其遗传型。事实上表型与遗传型之间的唯一联系，就是自然选择。它通过不同表型的不同生殖速度（生殖率），来改变遗传型的发生频度。学过代数的父母亲，可惜并不能通过他们的经历，去提高其子孙后代的代数天才。

个体表型的特征——如体长，可能受到若干种不同基因的影响，反之，单个基因则可以影响好几种特征发育。此外，一种特征上的具体优势，可通过不同的等位基因组合，在同种的不同成员身上得到体现。对人类群体来说，我们不能认为身高相同的所有人，其控制身高的等位基因组合均相同。在环境不变的条件下，可能有大量基因模式都能造成同一身高的人。

这种遗传多样性，必定也适于解释“智力”之类复杂多样的特征。自然选择可以（而且可能已经）给人类思维行为构造出很多不同的遗传基础。有些遗传类型体现在完成特殊智力任务的不

同本领上，但是同样的能力也可能有着不同的遗传基础。例如，短时记忆的容量和持久性，在大多数认知任务中，都起着重要作用。但能产生效率类似的短时记忆结构的，可能是很不一样的各种基因模式。

对人类来说，在有效功能上具有极大重要性的表型种类，要比其他生物的这种类型为多。它们有非常分散的倾向，而不是趋向于一致。诸如体质、力量、智力、灵巧性、学习能力、脾气等一般的和因人而异的品质，在生物适应性和人类状况两方面，都要比眼睛的颜色、头型或脱发程度等更具体的特征，有意义得多。因此，我们可以料到，在人类历史中，选择力对以上列举的一般品质，有着最为有效的影响。但由于那些品质的不一致性，我们不能指望选择力产生高度的基因齐一性。

### 能源利用

我们还可以把占据小生境的物种，看作是利用能源的物种。它们所利用的能源；说到底，是太阳提供的能量。构造小生境的机会和物种的激增，也许都受到了可用能源总量的严格限制。我们必须考察一下这种可能情形。

生物的一个重要特征是，它们不是简单地烧能源材料（尽管它们也这么做），而是在比较高的温度上，利用能源去发展它们的肌体结构，将能量的一部分转换为有机组织，同时使含能材料逐渐发生热力学耗散。含能材料不是直接经过代谢过程被耗散掉，而是被转化为有机的细胞结构了。就这一点而言，生物在一段时期里所能建立的能量贮备，是没有限制的。随着死亡和腐化过程同新贮备的相抵，最终就达到平衡了。

如果没有生物，地球所收到的太阳能量，实际上只经过一步，就重新辐射回太空了。生物通过在中间层次上贮备能量，延迟了这种耗散过程。能量从植物转移到食草动物，再转移到各种食肉动物，又到再生产者（细菌之类），从而在食物链上得到了

重复利用，有时多达四五次。太阳能所能维持的生命数量，在很大程度上就取决于这些过程的效率：能量的多大一部分没有为新陈代谢所直接消耗，而是被“贮备”起来了；每一次再利用时，有多大比例的能量被降解、耗散掉了。

尽管光合作用可利用直接辐射能的10%或12%，但植物所利用的太阳能，很少超过地球所收到的全部太阳能的1%，其余能量都被大气层吸收或被地面反射掉了。食草动物或食肉动物，可将含于食物中的很大一部分能量，转化为自己的肌体结构，或用于自身代谢。但从代谢过程中的能量损失上看，我们可以算出，可用能量在食物链上每降低一个层次，便耗散大约10%。即使是考虑到植被与动物的能量存贮的可能情形，再循环的时间也不长。从最初的光合作用阶段到最终的有机物降解阶段，所经历的平均循环时间一般不超过一二十年。<sup>①</sup>

这些数字表明，进一步有效利用太阳能的潜在机会，还是很大的，因此不能认为太阳能短缺是影响持续进化的一个严重限制因素。此外，这些数字虽然对地球所能支持的总生物量，给出了一些数量衡量，但它们根本没有表明，可能存在的生境种类有多少，或占据这些小生境的物种有多少。最有可能成立的假设是：物种将持续增加，尝试和发现日益多样化的大量生境所提供的专门适应的机会。

马尔萨斯观察到，无节制的人口增长将是几何增长。达尔文由此受到启发，强调增长的限制和固定稀缺资源的竞争。我们刚刚看到，这种观点并不全面。进化过程可产生新的生物，它们能利用能源，利用从前被浪费的或被低效利用的其他资源。事实上，当动物占据了由植物提供的新生境的时候，当生命在地球上の大本营从大海扩展到陆地的时候，上述情形的的确确发生了。

---

<sup>①</sup> 有关能量流和能量贮备的规模的数据，可参见E.P.Odum：《生态学基础》，第3版，第3章，（费城，1971年）。

我们可以设想，这种扩展今天仍然在较小的尺度上继续着。那种认为我们已接近某种稳定的均衡状态的观点，是毫无道理的。

## 社会进化与文化演变

我们已经看到，人类的生物进化史现已跨越了大约10万代，而自出现农业以来的进化，或许仅有400代。在后一时期里，没有什么证据表明人种有所改变，但是，持续的文化演变则证据极多。有些观察家因此而假定，文化演变已经取代了遗传演变，成为我们人类持续变化的主要过程。然而，这个假设怎么才能与达尔文模型和谐一致，并不完全清楚。在那个模型中，演变的是基因，而不是个人，更不要说是整个社会了。

一个围绕利己基因而组织起来的系统，是否容得下文化演变的概念？卢姆斯登和威尔逊的近期论著，明确地以进化论框架，论述了人类文化问题。<sup>①</sup> 他们把文化“基因”与生物基因并行看待，认为这种文化基因可以在一种文化环境中从一个人传给另一人，可以一代一代传下去。这种传递不是生物学上的遗传，而是社会传播。他们认为，我们可以把一种文化想象成为通过一组生物基因和一组文化基因之间的结合过程而得到发展的。这两部分的结合，意味着它们必须彼此和谐一致。最可能传递的文化基因，是那些最容易感觉到的，以及在社会成员生物构造基础上使用的。例如，语言中有关颜色的词汇，反映了人类眼睛的色觉机制。人的生物特征，对性行为、照料婴儿、认知过程和认知策略等方面，都有很强的影响。<sup>②</sup>

---

① C.J.卢姆斯登和E.O.威尔逊：《基因、心理与文化》（坎布里奇，1981年）。

② 人的生物特征强烈影响人类行为这一点，没有多少争论。争论多得多的当然是：个人生理差异是否能解释个人（或群体）行为差异？如果能够解释，在多大程度上做得到。

卢姆斯登和威尔逊指出：逆命题也是成立的，即在一个社会中，一定的文化特征（用他们更常用的语汇来称呼，就是文化基因）的存在，可以改变同具体生物基因相联系的适应性。这样，一群能传播文化的社会生物的进化过程，要比不能传播文化的生物进化过程复杂得多，因为前者是一个相互作用的过程。一方面，遗传物质决定了哪些文化特征可发展下去，同时，当前存在的文化特征在任何时候都影响着基因变种的适应和生存。关于这种相互作用理论，目前文献不多，卢姆斯登和威尔逊也许是首次写书详述这种理论的人了。

能改变自身文化的物种，是“可程序化的”。正如前面指出的，能最有效地开发可程序化物种的行为灵活性的遗传特征，是那些可以在许多环境下生效的特征，如：有益于强壮、健康、敏捷的那些特征，以及可以增进最重要的思维和学习能力的那些特征。可程序性还有利于社会存在，它在社会环境中（而不是在孤立环境中）能得到最有效的开发利用。我们可以特别考察一下可程序性的一个尤其重要的方面：在社会影响或社会压力下易于接受程序的性质。我们将这种性质称为“顺从性”<sup>①</sup>。在这一点上，一方面是适应社会生活的认知能力和性情，另一方面是作为适应机制的人类社会——我们不必问，它们两者哪个是原因，哪个是结果；哪个是母鸡，哪个是鸡蛋。显然，它们两者彼此强化，相互作用，均有利于适应。

自从我们人类诞生以来，在大多数岁月所一直经历着的条件下，体格强壮和头脑聪慧这两者对人类的成功繁衍或适应，都有着积极贡献。这一点看来肯定是有道理的。人们常常讲到这一点，在此无须详述。不太明显的是：在最近几代人的时期（比如说工业社会的最近百年来），健康聪慧同生育人数之间出现了正

---

① 此词亦可译为“可训练性”，西蒙在《管理行为》中对此词含义作过解释，指“可教化”、“可训练”的性质。——译注

比关系。想以近期文化变革打破这种联系，是在优生学标签下提出种种政策建议的依据。在这里我不想讲这个困难的话题。严肃地思考一个社会为形成其遗传素质而对生育过程进行调节的问题，并不是愚昧无知的主意。不过，在能达到目的的满意的杜会政策得以形成之前，我们还必须对付很多复杂问题，躲过很多陷阱。

如果说灵活性（如通过强壮、灵巧或理智而获得的灵活性）是可程序化社会动物进行适应的主要途径，那么，对这种动物来说，强壮和智慧所可能具备的适应优势，几乎同特殊的文化内容无关。无论面临着什么，强壮聪慧都将是所能去适应的。因此，我们可以认为，文化的演变同支撑着文化的物种之间，有一种半独立关系。文化特征的遗传机制（尤其是个人和社会的各种学习方式），同生物遗传机制很不一样。文化遗传显然是拉马克式的遗传，习得的特征确实可以转移。

上述两种进化机制之间的微弱联系，至少有两个潜在结果。一方面，一些文化特征（如西方工业主义）的成功传播，并不意味着那些特征的开创者拥有优越的遗传适应性，他们可能缓慢地增多，或者根本不增多，并且反而在具备他们最初提出的那种文化的人口当中，成为越来越少的一部分了。另一方面，如果有一群特殊的人具备某种文化，而这种文化使他们在同其他人群的竞争中，占据适应优势，那么，就该种文化并没有给拥有它的那群人造成自我拆台的选择压力而言，它可以给一个真正的“社会进化”提供基础，也就是，给这群人以牺牲他人换取生存提供基础。当然，这只有在各群竞争者之间可以防止广泛的相互借鉴的情况下，才会发生。在近代史上，欧洲人对北美的征服，也许是这一过程的最清楚的例子了。

人类征战史告诉了我们，文化适应性与作为文化载体的基因的适应性之间，关系是何等的复杂。蒙古人在十三四世纪的征战中获得了极大成功，征服了当时世界上可见人迹的很大一部分地

益。但是，这一成功既不意味着文化适应性，也不意味着基因适应性。在文化上，他们大多采用了被征服者的社会模式。在遗传上，蒙古人的人口（如果确有增长的话）是否比被征服民族人口增长更快，根本就不清楚。

## 进化过程中的利他主义

对“利他主义”一词，我们可以用许多不同方式加以解释。其中，最狭义的概念是所谓纯粹利他主义，或称强利他主义，即为了其他生物的利益而无条件地牺牲自己的适应性。与此相对，有一个比较广义的概念，即弱利他主义。我们对这一概念特别感兴趣。我们将看到，这种利他实际上意味着明智的利己。当我们说起弱利他主义时，我们指的是：一个个体虽然牺牲了短期的适应，但却收到了间接的长期报酬，而且，这报酬又大于其直接付出的代价。近年来，群体遗传学家提出了几种能用以解释弱利他主义的演变的机制。<sup>①</sup> 其中每一种机制，都依赖于利他基因获得奖赏的某种特殊的间接渠道。

### 选择利他主义的机制

为了确定弱利他主义存在的起码条件，我们先讲一讲现已得到了考察的两种产生利他主义的机制：亲属关系和结构化同类群。

亲属关系模型向我们提供了一个已经展开大量讨论的弱利他主义机制。个体如果认得其近亲，那么，由于亲属所携带的基因有许多是相同的，所以个体为亲属的生存而牺牲自己的利益，就

<sup>①</sup> 关于群体遗传学的各种利他论，有一个优秀的论述，参见D. S. 威尔逊：《群体与社区的自然选择》（门罗帕克，加利福尼亚，1980年）。

可增强共同基因的适应性。这里所谓“认得”至少出现于两种情形，即无非是居住地点离亲属很近，以及向近邻提供特别利益——这种模型并不要求具备区分亲属和陌生者的特殊本领。

这种亲属关系模型，看来只有在能对很紧密的亲属产生有差别的效应的情况下，才说得过去（也就是，可以解释适应性的增强）。因为，即使是直接堂亲关系，也只有 $1/8$ 的基因是共同的。因此，亲属关系机制看上去最可信的，只是对母亲照料幼子、同胞间的义气和牺牲精神、以及核心家庭成员间其他利他行为的解释。同样道理，以亲属关系为依据的利他主义，在无性生殖情况下，看来要比在有性生殖情况下发生得更频繁。

亲属关系式的利他主义模型，有一个重要的推广模型，即：以结构化同类群为根据的模型。我们假定一个地域被某类群体所占据，并且整个群体（同类群）分成几个小群体（特性群体）。设想该物种的生命循环分为两个阶段。在第一阶段，个体仅与同一特性群体中的其他个体发生相互作用。在第二阶段，整个同类群先均匀地混在一起，直到下一次循环开始时又分成了新的特性群体。为简便起见，我们假定生殖和选择均仅仅发生在第一阶段。

假设突变产生了一个利他基因，使载有该基因的利他个体以某种自我牺牲代价，进行有利于所在特性群体全部成员的活动（但不是对整个同类群中的其他成员有利）。对任何一个特性群体来说，这种情况的结果是，导致了非利他者较之该利他者的更强适应性。但是，利他者对非利他者的数量之比，各特性群体有所不同。利他者数目最多的特性群体，其平均适应性最强（因为，每个利他者的利他行为的代价，从该群体其他大量利他者给他的利益中，得到了补偿，且有所余）。那些有优势的特性群体中的利他者，要比非利他者数量很多的群体中的利他者，适应性更强。我们容易严格证明，如果不同特性群体中的利他者比例之间，方差充分大，那么，利他者的平均适应性将超过非利他者的

平均适应性，其结果是：利他基因将取代群体中的非利他基因。产生这一结果所需方差值要求并不过分。事实上，如果利他者在不同特性群体中的分布的方差，同二项分布的方差一样大，利他基因就会占上风。

在自然界，看来有几种与此相似的系统。为说明起见，我讲一下D.S.威尔逊讨论过的一个系统。<sup>①</sup>有一种昆虫在瓶子草（猪笼草）里生活，它们在那里找到一些很小的水洼，在水中排卵，养育幼虫。在生命循环的第一阶段，这些虫子只同少量同类一起（特性群体）生活在同一片瓶子草叶上，它们与生活在其他叶片上的同类没有相互作用。第一阶段过后，那些昆虫的活动范围就大得多了，它们爬来爬去，在为进行生殖而重新定居到某一叶子上之前，多多少少比较均匀地同所有特性群体中的同类们，混在一起了。

假设某一瓶子草叶上有一只昆虫经历了突变，其活动完全是为了那片叶子上的其他昆虫过得更好——例如，它改变水的酸性，或是在水中添加某种营养物质。由于这种活动所花的能量本可用来生育后代，因此，这个变种的适应性相对来说被减弱了。然而，那片叶子上的非利他者们的适应性，却得到了增强。乍一看，热心肠的变种们似乎是在干最后一件好事了，其实不然。正如我们已经讲到的，如果不同草叶上的变种者比例不同，那就会有更好的事情发生。变种的昆虫在一片叶子上偶然聚首，会使它们形成变种优势。

平均来看，利他者将愿意生活在利他者比例较高的特性群体中，而不是呆在大多被非利他者占据的特性群体里。注意到这一点，便可消除上述结果的明显佯谬。也就是说，平均来看，利他者将处于比非利他者更优越的环境，因而利他者将更适应环

---

① 参阅上注，第21页，第35—36页。

境。①

依结构化同类群模型来看，利他者是有报偿的（从而仅仅是一个弱利他主义者），尽管将他选为受奖对象的唯一机制，只是让他在同样是利他者的地盘上居住。恰恰是这种地域上的集中化，造成了利他者的“识别过程”，以及籍此取得的成功。没有这种有差异的回报，利他者就无法生存下去。

### 利他者的识别

一旦我们引进了用以识别利他者和有区别地给其奖赏的种种机制，我们就可以看到，利他行为（仍指明智利己意义上的行为）得到进化的潜在可能大大地增加了。因此，我们不必假定，一个物种各个成员之间的所有竞争，都是张牙舞爪的争斗。我们还可以解释一下，为什么多少有点儿聪明的人类能在世界上生存。人类，还有其他许多物种的成员，有着很强的识别能力，能够认出以前跟他们打过交道的个体，并能以从前交往所发生的事情为依据，有区别对待它们。

考虑到这一能力，我们不难构造出这样一类模型，用它们就可以预测各种利他行为的共同演化，和利他行为受益者的报答行为。在这种相互交往的社会里，一个生物对其他生物的行为的变化，可导致相互报答的变化，这样也就带来了生境的演变。因此，我们实际观察到的社会高度分工现象，存在一个明显的进化论解释。让我们更周密地看一看这可能是怎样发生的。

### 社会进化中的利他主义

在一个社会环境中，任何一个个体的特殊行为，都可能受到

① 结构化同类群机制，同下一章所讲的俘虏两难策略[又可译囚徒两难策略——译注]是类似的。关于这两个概念之间的联系，请参见艾克赛尔罗德和W.D.哈密尔顿：《协作的进化》，载《科学》第211期：第139—1396页（1981年3月27日）。

其周围邻居们的奖励或惩罚。如果我们把小生境看作是奖励利他行为，从而朝着利他主义的方向改变基因的，那么，我们必须以一种非常特别的方式，去解释“奖赏”。达尔文主义选择过程中的仅有奖赏，就是增强适应性的那种奖赏。人获得大批财富和荣誉，不会有任何遗传上的后果，除非富有和名望能使人生育更多的子孙后代——也就是，除非这种奖赏能增强获得配偶和养育孩子的能力，以及这样的欲求。

事实上，利他行为的遗传基础的发展，看来要求三种特征的共同进化：（1）靠我们的行为表明赞同别人的利他行为（或反对自私行为）的一种倾向；（2）通过负疚感或羞耻心，对所表达的赞同或反对态度做出反应的一种倾向；（3）不仅以赞赏态度，而以更大的生殖机会（或责任）去奖赏利他主义的一种趋向。

所有这三种特征都是不可缺少的。特别是，利他主义如果对适应有利，也只能被认为是在遗传上赢得了兴旺繁荣。如同前面指出的，早期社会中的社会赞许与生殖成功之间的联系，要比当今社会的这种联系更容易被人察觉。不过，直到近代，这种联系恐怕一直很强。如果没有社会支持，不仅会阻碍获得配偶，而且会普遍危及家庭的存在。因此，无论当代社会是否存在这种选择压力的影响，人类对社会压力做出反应的遗传特征，以及人类相互施加这种压力的倾向，都已然经历了很长的发展时间了。

不过，正如我们看到的，社会对个人的奖赏（及惩罚），有许多都同适应本身根本无关。出于进化上的某种原因，只要社会成员认为这些奖赏有价值，这些奖赏便可用以诱发为社会所赞许的行为，包括那些通常被认为是“利他的”行为。例如，设想有这样一个社会，在那里，富人生育的子孙并不比穷人的后代多。富人想得到人们尊重的愿望，会使他们割舍一部分财富用于慈善目的，这通常被认为是一种利他行为。但是，只要这种慈善施舍并不影响遗传上的适应性，从进化观点看，它就不是一种遗传性的利他行为。

由此可见，社会学所定义的利他主义，要比遗传学上的定义广泛得多。在社会学上，社会奖赏可以支持多种社会利他行为，而这些行为从遗传学上看则属于中性的。当然，在认为人类行为存在动机驱使的任何理论中，这种利他主义总可以被解释为一种互利主义，或弱利他主义，而不是强利他主义。富人虽然舍了些财，但换取了喝采。不管怎样，我们没有必要证明他的适应性在平均意义上得到了增强，以使其社会利他行为同其顺从自然选择法则相和谐。这种行为的长久存在，可由如下事实决定：它有利于整个社会的适应性，因而得到了社会的奖赏。

当然，仍从长期上看，象那种想得到世人称颂之类的愿望，只有当它们在平均意义上有利于个人的适应时，才是以自然选择为基础的。但应注意，这种联系并不是直接存在于称颂与利他之间的联系，而是称颂与社会所希望称颂的行为之间的联系。对称颂的反应，是我在前面叫作顺从性的那种普遍特征的一个重要形式。下面我就讲一讲作为利他主义的基础的顺从性。

顺从性可定义为：依照社会赞许的方式去行动，并抑制那些不按社会赞许的方式去行事的一种倾向。跟其他特征一样，顺从性可能也是在自然选择的影响下形成的。也就是，顺从性如有利于个体的适应，顺从性水平便提高，反之则下降。<sup>①</sup>不过请记住，顺从性不是以具体方式去行动的一种倾向，而是以社会认为是合适的方式去行动的一种倾向。因此，在这种机制使个体产生的行为当中，有些可增强其适应性，有些则可减弱其适应。如果顺从性变得选择性太强——只接受能增强其适应性的要求——那它就不再成其为顺从性了，因而将得不到给予顺从性的社会奖赏了。

① 请比较一下E.O.威尔逊《社会生物学》（坎布里奇，马萨诸塞，1975年）一书有关灵活性（易变性）与不可教化性的讨论。另外，请参阅D.T.坎贝尔：《论利他遗传学与人类文化的反享乐因素》，载《社会问题杂志》28（3）第21~37页（1972年）。

通过让孩子们长期享受照顾，依赖成人，并靠学习掌握有效技能，顺从性无疑大大增强了人类的适应性。顺从性当然也可以使孩子和其他年龄的人，产生那些不能增强其适应性的利他行为。但平均来看，比较高的顺从性水平，必定对适应性有积极贡献。

于是，我们不同“利己基因”论发生争吵，就可以给整个社会的演化引进一些机制，它们能把社会准则，加到选择过程之中。这种作法的要求在于，奖赏要同一般化的顺从行为或“遵命”行为相联系，而不是同具体行为相联系。我们引进的机制所需要的全部条件仅仅是，它们的总效果有利于个体的适应。并不要求它们所允许的每一种特定行为均对个体适应性有利。

正如我们从社会昆虫的存在上看到的，社会依赖性的演变，还有着不同于此的其他基础。不过，人类社会行为，同人类思维及学习的能力，特别紧密地交织在一起，它们之间的联系，完全存在于顺从性之类的社会反应机制的演化令中。

### 进化的短见

如果我们能通过下意识的理性计算去筹划一项行动，如果我们足智多谋，我们就可以展望到未来一段时期——至少是比较短期的未来——并设想自己的行动在那时可能产生的后果。原则上，并没有一个固定的时间期限，限制我们对行动后果的计算。

这种“向前看”的能力，同生物进化机制形成了鲜明对照。后者只提供了一种非常短见的合理性。适应性趋向于选择那些能获得直接短期优势的物种。生物从其所处始点开始攀爬的，是局部的适应之峰。<sup>①</sup>

---

① 爬山是隐喻。——译注。

## 局部极大与全局最大

对一个非常简单的世界来说，短期优势可不断积累成为长久优势。诸位如果在这样一个世界上登山，最终将达到世界的最高峰顶。但是，这个结果只有当世界上仅有一座山的情况下，才能得到保证。一个生物所生活的世界如果充满了峰峰谷谷（比如说象加利福尼亚地势那样），那它就会发现，自己处在一个山顶还不行，还必须下山去爬另一座山。因此，把适应性看作是靠进化谋求最优（最大）的任何一种进化论观点，只有对局部极大来说，才可能成立。除非我们认为世界的形状很特别，很简单，否则我们不应设想进化会导致一种可被正确认为是全局最优的状态。

我们还看到，在山峰叠嶂的世界上，一条特殊的爬山路线，可以决定一个系统将费力攀爬哪一座山。哪种突变首先出现，可以决定系统在众多方向上将实际朝何处进化。自然选择理论，对于究竟哪座山将被攀爬这一点，没有做任何预见。由于物种所可能发生的变异如此之多，根本不可能彻底搞清楚地球存在的全部时间里的全部变种，因此，会有许多山永远不被攀爬，其中有些可能还很高。

现已有人提出了加速进化和消除短见效应的种种图式。其中一种是这样的：不是仅仅保留一种最适应的遗传种系，而是保留大量比较适应的种系，让它们平行地进化。<sup>①</sup>由于“次优的”物种不同于最适者，它们在任何指定时间均能攀爬不同的山峰，因此，这种图式能使直接优势避免失落于小山坡上的爬行过程。不过，在将平行图式用于生物进化时，需要考虑某种防护措施，使“次优”物种在同暂时最优者直接竞争时，得到保护。在一定程

<sup>①</sup> 参阅J.H.霍兰德：《自然系统与人为系统的适应》（安阿伯，密歇根，1975年）。

度上，小生境专门适应，就可以提供这种保护。

我们可采用一个稍微不同的方式，得出这个结论。达尔文进化过程，是一个先产生一定的可能变种，然后检验它们，并保留其中较优者之过程。如果要靠这一过程保证达到全局最优，只有实际产生出全部可能变种才行。正如我们已经看到的那样，即使我们只考虑有 10 000 个基因的单个染色体，每个基因都是一对等位基因，那也要有  $10^{3000}$  个可能的变种，这是在世界的整个寿命期里也不能完全产生出来的。

如果爬山努力所依仗着的地形展望不稳定，全局最优的设想就更值得怀疑了。假设我们有一幅描绘演化过程和小生境构筑过程的图景——我们可以说，不断有新的山峰冒出来，因此，进化过程不会导致任何所谓最优状态，甚至也不会达到稳定的均衡状态。在这样一个世界上，进化过程不断开辟新的可能，新的组合。即使是在无机物质的层次上，我们也可发现存在着这样的演化。从一时来看，只有少许几种元素可存在于稳定状态。通过种种组合过程，我们今天在生物世界和社会世界里所看到的巨大复杂性，就逐渐形成了。

关于进化不会导致最优和稳定这一点，有一类重要证据，是达尔文本人提供的，尽管他提出这些证据的本意与此很不同。<sup>①</sup>他列举了许多有名的例子表明：进入新岛屿或新陆地的外来物种，幸亏具有在新环境中生存的优良适应性，才得以到处乱跑，并消灭了原先居住在那里的有些物种。达尔文用这些例子证明自然选择的威力——这的的确确很成功。不过，这些例子也同样证明了，生物进化的世界有着非最优状态和不稳定状态。例如，北美生物群在英格兰麻雀到来之前，如果已经达到了最佳适应状态，那种麻雀在那里就找不到自己的生境了。

由此可知，引进新物种的成功，有力证明了进化生成器的不

---

<sup>①</sup> 《物种起源》，第6版（1872年），第12章。

完备性，从而证明了系统无力达至最优状态。在相对的意义上讲，比较适应者生存。但我们毫无理由设想它们是任何绝对意义上的最适者，甚至也不能定义所谓最适应指什么。

如果我们是生活在白垩纪，我们可能以为恐龙是高度适应的动物，那时它们确实如此。人们普遍认为，恐龙是在无力适应环境快速变化（也许是突变）的时候，失去其适应性的。但是，难道我们能肯定，它们在我们现代世界的某些地方不能再次高度适应吗？也许我们当今之所以见不到它们，只是因为它们还没有被再次创造出来，而不是因为它们不适应当今世界的某类生境。如果这对恐龙来说，以及对曾经被进化过程实际构造了的生物来说，在想象上是正确的，那么，这对于从来没有出现过的众多可能生物中的一部分（大部分或小部分）来说，看来就更不存在疑问了。

### 无终极目标的搜索

由上述可知，进化过程的目的论颇为特别。那里没有目的，只有一个搜索和改进过程。搜索本身就是目的。前面我已经指出，有些人之所以喜欢把进化视为理性的一种解释，仅仅是出于这样一个理由，即：进化不需要详尽细致的解释，重要的是适应，而无论适应是怎么形成的。似乎进化可以让我们不确定手段，便假定目的。现在我们看到，这个问题实际上并非如此。进化确定了手段（变异和选择过程），它们不导致任何可以预见的目的。从没有手段的目的，到没有目的的手段，我们转了整整一圈。

让我们针对人类文化，特别是针对科学技术的发展，考虑一个与上述观念相平行的思想。科学技术不是汽车和发电厂之类的物，而是知识和计算工具，它使我们得以发明汽车和发电厂之类的物，还使我们得以思考是否需要、以及在多大程度上需要制造

和使用那些物。科学技术把我们的时空观延展了；加深了，因为它使我们能够更快捷地构造出备选方案，并且能更好地评价那些方案。

在小生境竞争模型中，尤其是在小生境增生模型中，我们可以认为，知识的进化是总进化过程的一个不可缺少的部分。不过，在这种进化过程中，看得出的唯一目标，就是思想的增生，以丰富世界上的现有知识库。我们可以把知识的增多，说成是整个过程的目的，其自身即为目的。

传统的进化论由于集中考虑对固定环境的适应，故而谈论目的。它们所谓目的，乃是适应这一目的，乃是最佳适应于环境这一目的。另一方面，强调构筑小生境的进化理论，除了也许稍微有些复杂之外，并没有向传统理论看齐，它们描绘了不向任何特定目的演化的系统。

人类，至少我们当中有些人，被一种观念搞得时时不安，这个观念就是：世界可能是一个封闭空间。诸位当中也许有些人因为“新世界”已经被哥伦布发现过了，并且再也没有尚待发现的另一个世界，而感到烦恼。进行太空旅行的一个动机就是：这种活动提供了到达那些尚未被占据的新世界的机会。很多人显然向往一个不是封闭的世界，在这样一个世界上，永远不会有这么一天，使我们到那时可以说：“好啦，要了解的，我们全知道了；要干的，我们都干完了。”一个由演化着的小生境所组成的世界，一个复杂性不断增长的世界，无论它的其余性质如何，都不会使我们心怀这种特殊的牵挂。

## 结语

下面让我把前面的内容串起来讲一讲。关于理性过程的进化观点意味着什么？首先，接受进化论（从而接受进化的理性），并不能把我们禁锢于那种全局最优的观点，那种认为任何东西都

朝某种稳定的最优状态演变的观点。它仅仅使我们相信，存在着许许多多对现有环境的局部适应，同时存在着朝向目标的永恒运动，而目标本身也是不断运动的。

第二，理性的进化模型并没有让我们构造一个特殊的理性过程机制。它仅仅告诉我们该过程可能运动的方向。

第三，达尔文主义的进化论由于假定了变异和选择，因此要求我们非常严格地采用利己基因的概念。特别是在固定小生境模型中，我们很难找到除利己主义之外的其他任何东西的地盘。但是周密考察表明，实际上存在着一些很有效力的机制，就是我在前面描述过的种种反馈机制，它们可以使能够增强适应性的利己主义，变成明智的利己主义。当存在适当反馈时，不明智的利己所碰到的生存问题，同绝对利他主义所面临的问题可能一样困难。因此，在达尔文主义的世界上，我们必定可以看到那些通常被说成是利他主义的大量行为；当然，这些利他行为实际上可通过我们所讲的一条或多条间接途径，得到奖赏。

第四，在真实世界上，最强的和最突出的竞争过程，可能并不是争相占据固定小生境的过程，而是专门适应和构筑小生境的过程。因此，我们不必非得把这个世界看成是到处血淋淋的。

进化论确实给世界的可能面貌的类别，加上了很强（抽空而言）的限制。它确实告诉我们，现有竞争者当中不大适应环境的生物如果大量存在，世界就无法继续存在下去。竞争中的弱者无非是将在竞争中消亡。

因此，进化论的的确确做了预见，至少是告诉了我们，某些物种必然消亡，有些物种根本不可能生存，不值得为它们筹划生计。在这个意义上，进化论是反乌托邦的。无论它有什么限制，它并没有给我们戴上最优化的枷锁。它也没有强求我们把狭隘的利己，当作人类能生存下去的仅有的动机。

最后，如果我们把进化论同上一章所讲的三种人类理性模型做一个对比，我们就会看出，进化论同行为模型最为接近。在这

两种理论中，搜索巨大的可能空间和评价搜索结果，均是适应性变化的核心机制。这两者都是短见的。它们所讲的最优仅仅是局部最优。我们不要将其解释为最优化过程，最好的解释是这样一些机制，它们能够发现“改进”原有状况的新的可能。

在下一章里，我们将考察一下，适应性过程的这些特征对于人类社会事务上的理性应用来说有何意义。我们要问一问，讲一讲，什么样的人类有限理性，能对复杂世界中的抉择和规划做出贡献。

### 三 社会事务上的理性过程

人们或许要问，为什么要讨论社会决策？难道讨论个人决策还不够吗？社会决策为什么是我们所必不可少的？当前，广泛流行着一种自由主义的虚妄观点，它认为，人就是某种莱布尼茨式的单子（某种稳定的小单位），每个人都具有独立的、自身一致的效用函数，人们之间仅仅是通过各自所了解的市场价格而相互影响的。实际上，人并非如此。由于种种原因，我们绝不是单子，原因有很多。我们的价值观，我们的有意识的抉择活动，以及我们对自己的活动所产生后果的了解——所有这些知识和偏好——都产生于我们同社会环境的相互作用之中。我们的价值观和知识，有些是幼年时在家庭生活中学会的，有些是从社会环境中得到的，而且往往是不加批判地接受的。或许，还有一些是在与环境的抗争过程中形成的。不管怎么说，完全孤立地产生知识，恐怕是绝无仅有。

1970年前后，数以百万的美国青年学生以激进分子自居，而到了80年代，绝大部分人则认为他们最好还是走中间道路。如果用一个独立随机变量的模型来处理这些现象，那其中的统计概率意味着什么呢？它的意义就在于，正如此例及无数其他现象所表明的，人们的信仰和价值是可以高速传播和蔓延的。各种信仰，甚至那些我们认为最正确的信仰，其可信性大多也不是获自直接经验和试验，而是由于它们被社会上可信而“合法”的集团所接受，从而才被我们接受的。

在我们的社会中，在大部分其他的现代社会里，人们交换货币和产品的市场发挥着一种异乎寻常的重要作用。然而，市

场并非是在真空中运行的，它们只是更为宽泛的社会制度结构的一部分。而且市场的运行带有大量的外部性，也就是说，人们在市场经济中所进行的活动的许多后果，并不完全体现在市场价格中。这方面的一个典型的例子是，你家烟囱冒出的烟会使你的邻居感到不快；或者邻居的立体声音乐骚扰了你的宁静，他的音乐享受构成了你的噪声污染。在每个社会里，尤其在城市社会里，我们活动的很多方面对他人生活和价值的影响，都是难以用市场价格来调整的。

如同自由放任的市场不能适当地制裁消极的外部事物一样，某些对人们有益的公共产品的生产，也同样得不到适当的报酬。我们不必为自己在社会中享用的许多东西付出代价。例如在匹兹堡，每天早晨我步行去上班时，都可以获得一些公共产品（对我来说非常有价值）。如邻居们修整得平整、碧绿的草坛，以及他们精心栽培的郁郁葱葱的花木，都使我感到极其赏心悦目。一年前，当沿途一些空地的主人们开始建造一些难看的公寓时，我这些免费所得的公共产品就减少了。但是这种减少并不体现在那些公寓的市场价格中，公寓的新主人们不必补偿我的损失，正如我不必为观赏邻居们的花草而付钱一样。因此，如果建房者不受这些间接效果影响的话，那些丑陋的建筑就会造得更多些；而如果考虑到观赏者不付代价的欢悦的话，那些花园也许将不再那么美丽。

外部性，无论是正的还是负的，都被编织在整个社会之网当中。它们是个人所得报酬的重要决定因素。由于外部性的存在，从而损害了自由主义的国家无权干预个人所得的基本宗旨。是什么因素决定着人们的贫富状态呢？对于一个裸体中的婴儿来说，哪些信息能更准确地预测其成年后的生活水平呢？第一，是他所出生的时代；第二，是他所出生的国度；第三，是他出身的家庭状况。用表示因果关系的推理分析，这些因素可以在很大程度上说明，20世纪的大部分美国人或瑞典人为什么富裕，而20世

纪的中国人和印度人为什么贫困。假如我们的运气好，生得是时候，或生于有利的地点或家庭中，那么在人生竞赛的跑道上，我们就可以获得某种好的起跑点；反之，则将不能获得这种机会。

即便我们接受自由主义的论点，认为一切包含个人努力的产品均不受侵犯，这种论点还是把世界收入置于合法的课税或控制之下。不过，如果我们相信国家应对收入再分配加以很大限制，那必定是因为，收入再分配可能削弱人们的生产动机，而决非收入再分配在道义上的“不公正”。

所以，对于那我们称为个人主义的单子主义状态，我们尽其所能，固然可以说它说得完美无缺。

然而，最好的近似实际上也是非常粗糙的。我们所有的行为均发生在一种错综复杂的制度环境中，并对他人造成数不清的影响。市场结构是无法替代社会相互作用的整个网络的，因而也无法证明自由主义政策的正确与否。

舆论界目前对社会制度，尤其是政治制度，评论不佳。特别是，当我们仍用老套话描述政治制度时，我们把他们称为官僚制度，说它们低效无能。不过，当然还存在着其它的观察和评价社会制度的方式。如前两章所述，我们所有的人，在如何完整地筹划我们的行动上，在复杂的世界里所能达到的理性上，都受到极大的限制。而制度则为我们提供了一种稳定的环境，使我们至少可能达到微弱的理性。例如，我们能可靠地预料到，如果朝某个方向走上一程，就可以找到一家食品店，并且它明天还会在那个地方。靠着制度环境的这种稳定性，以及其他许多没有什么疑问的稳定性，我们就可能对自己的行动后果，进行合理而稳定的规划了。

因此，我们的制度环境，就像我们的自然环境一样，在我们周围设置了一套可靠的，可感知的模式。我们没必要非得理解产生这些事件的根本的原因机制，或这些事件本身的全部细节，而只

要了解那些影响我们的生活和需要的少数模式就足够了。我们所生存的环境（社会环境和自然环境）的这种稳定性和可预见性，允许我们在有限的认知能力和计算能力的约束下，去应付我们所面临的问题。

## 制度理性的限度

在这一章中，我想讨论一下社会制度的理性，但我的目的并不在于把制度理性塑造成某种简单化的英雄。恰恰相反，我倒很愿意讲一下限制个人理性的某些方面（我们每个人筹划有效行动的能力）给我们社会制度的设计和运行所造成的问题。我要强调的是，我们据以合理筹划和行动的能力是有限度的，以及这种限度怎样影响到社会制度上，使我们的制度理性也呈现类似的限度。

### 注意力的限度

人类的心理限度给社会行为造成的第一个问题是：我们的政治制度在处理问题时，特别是在处理“重大”问题时，必须采取分阶段的，一时尚一事（或极少的事情）的方式。不幸的是，需要处理的整整一系列公共问题，不可能被同时排上议事日程。其原因在于，当问题事关重大和有争议时，（重大问题往往都有争议），它就必须用民主程序来处理，这种程序需要由立法机构或全体选民的多数通过。结果在某一个时期内，选民或立法机构就必须或多或少地同时注意某一件事。当然，一个立法委员会，是可能同时处理许多问题的；不过，在许多时点上，整个机构必须花费时间在重大问题上取得一致意见。

由于决策在一定时间内难以把注意力集中在更多的问题上，就产生了至少两种现象，这两者，虽然同出一源，但乍一看都是貌似矛盾的。第一种现象是，赶时髦的政治行为。在60年代，环

境问题是时髦的。我用“赶时髦”一词不包含任何贬义，只是意味着大部分可利用的政治注意力都集中在这些事情上。在那个时期，为保护和改善环境质量而设计的名目繁多的新法规，是最容易得到立法机构的批准的。

随后，到1973年，在第一次石油危机的冲击下，我们突然得知，我们不再拥有那些垂手可得的，取之不尽的能源了，或至少必须为购买石油支付极高昂的价格。我们突然又变成了一个只关心能源的社会，为能源短缺，特别是石油的短缺所困扰。在全力以赴对付新危机的努力中，环境保护问题则被严重轻视了（现在仍然如此）。在我们的政治制度中，似乎很难记住一个社会可能同时存在着好几个紧迫的问题。

另一个例子是，六年前，我们非常关心通货膨胀问题。不久，我们几乎所有的经济政策统统转向抑制和减轻通货膨胀的压力。正当我们全力对付通货膨胀时，我们却忘记了经济也意味着生产，意味着必须为人们提供就业机会，以使他们能够养家糊口。在采取强有力措施对付通货膨胀的同时，我们却听任失业率上升到大萧条以来的最高水平，大量生产资源被闲置。在此之后，失业问题便开始与通货膨胀问题争夺公众注意力，不过，此时的一种现实可能性却是，失业问题的解决，可能以放任通货膨胀达到新的更高的势头为代价。看来，我们似乎得熬过一个非常艰难的时期，我们的社会必须把注意力同时集中于这两类相关的问题上，这些问题如此紧密地交织在一起，使我们陷入了那种顾此失彼的尴尬境地。

我们的社会中还有另一类人，他们倒不像我们刚提到过的大多数人那样热衷于赶时髦，他们患的是另一种变态症。这些人的政治兴趣基本上总是集中在某一个问题上，不论是堕胎还是禁止堕胎；管制枪支还是不管制枪支；在学校里作祈祷还是不作祈祷，等等。这种人，无论政治议程发生了什么变化，他们的反应永远是一门心思地在他们偏好的那个问题上转悠。他们对竞选人

的选择，取决于该竞选人对困扰他们所关心的那个问题的立场。

科恩、马奇和奥尔森提出了一个表达此类现象的饶有趣味的模型，他们给它起了一个不太雅观的名称：“组织抉择的垃圾桶模型”<sup>①</sup> 他们的观点是，在某些社会和组织中，存在着某些持续性问题，还有顽固地追随着这些问题的人们。当决定具体事情时，这些持续性问题就会纠缠不休，那些一门心思的家伙就又会把事情扯到他们的问题上去。这样的组织是难以对它想作决定的问题做出决策的。比如说，学校课程设置委员会表面上的问题是，究竟是开设课程甲还是乙对学生更好些；而争论的真正焦点则是，设置课程甲或乙对这个或那个系的教授职位将会产生何种影响。

政治时髦与单边政治，均出自同样的基本原因：人类不具备在同一时间内思考许多事情的能力。结果，政治机构虽被认为是在处理整整一大批社会问题，但有时在平衡其注意力时，却存着极大困难。幸而这种困难被第一章所提到的世界的一个特点——事实上并非所有的事物都彼此密切相关——缓和了一些。为了强调这种困难，我挑选出一些基于人的有限的注意力而产生的困难的事例。例如，比起大部分随机选择的成对问题来，能源和环境之间关系更为密切。解决能源问题的许多办法都有可能产生或加剧环境问题。比如说，如果增加化石燃料的使用量，由于二氧化碳效应，地球的平均温度便可能被提高，而这将是危险的。众所周知，地球有严格的正常温度范围——或至少是人、人工制品为事物和人类制度一般所能适应的温度范围。我们同样也可以用通货膨胀和失业的例子来说明这种困难。总之，对这类成对问题，你只可能在不顾及另一个的情况下，圆满地解决其中的一个。

不过，许多问题之间相互联系的网络并不致密。加之，某些重复的或可以被预见的问题，有可能被平行地处理。也就是

---

① 《管理科学季刊》(1972)17:第1—25页。

说，如果我们能设置有关政策以及与这些政策一致的程序，我们就可以建立某些平行的组织去执行这些程序。例如，仅需要市政委员会给予间断性的注意，各类市政组织便可同时运转。消防部门负责救火，警察部门负责维持治安，公共事业部门负责维修道路。这正如一个人的心脏在正常情况下必然持续而有规律地搏动，而不必给予专门注意（除非心脏出现病变）一样，社会的常规性需要同样可以用平行方式加以处理。但是，对那些新奇的和那些意料之外的适应性变化，则需要予以全神贯注地处置。

甚至连那些彼此无关的问题，当它们共同争夺同一种稀缺资源时，也会变得相互关联起来。军备问题怎么会与社会福利问题扯到一起的呢？那是由于，倘若把钱花在军备上，就不可能再把它用于社会福利了。有鉴于此，政府的预算过程往往成为联结社会的各种需求、意愿及目标的焦点。

### 多重价值

有限的个人理性造成的另一个问题是，在对付相互冲突的目标所代表的多重价值问题方面（像我讨论过的那些例子），我们的社会制度不具备简洁而奇妙的解决办法。在能源和环境问题上，既无公式也无数字可以告诉我们，应该把改善环境强调到什么程度，以及应使我们的能源需要满足到什么程度。同样，对于那些每个人都可能以不同方式权衡其价值的利益冲突问题，我们也缺乏对付它的高明手段。

肯尼思·阿罗的著名的社会福利定理，集中体现了这种困难。该定理表明，社会福利函数必须满足一些似是而非的、非常不一致的假设条件，因此，这样的社会福利函数根本不可能存在。构成阿罗定理基础的合理的前提之一是，假定允许不同的人们用不同方式为其价值加权，其含义为我们并不想强求所有的人都具有相同的价值观。如果我们接受像这样一些假设的话，我们就会发现，我们实在是搞不清应该怎样比较两个人之间不同的价值观，

正如苹果和橙子不可比一样。因而，在某些试图允许人类选择多样化的似是而非的假设下，我们还是无法定义一个有可能解决利益冲突问题的社会福利函数。

### 不确定性

个人的有限认知能力给社会组织造成的第三个问题，是处理不确定性的困难。我们都好战，事实上，在当今时代，战争是特别令人厌恶的，并且人们对战争的厌恶程度超过了人类历史上任何时期。但是，我们对于究竟采取什么行动将可能增加或减少战争爆发的可能性，却缺乏清楚的认识。对苏联采取强硬路线（或温和路线），究竟会增加还是减少战争的可能性？许多人对这个问题都有一些自己的见解，然而几乎没有人能相当肯定自己的见解。对于我们整个社会来说，在这个问题和相似问题上的确定性，是一种自相矛盾的东西，因此，要想达成意见一致的行动方针，是极其困难的。

即使在那些中等程度的不确定情况下，要寻求“最优”的行动方针，几乎也是毫无希望的。在存在价值冲突的情况下，这种冲突差不多总是存在，我们甚至于不知道应该怎样给“最优”下个明确而一致的定义。然而，对不确定性，我们还不至于完全束手无策。如果我们采纳“寻求满意”的观点，如果我们只寻求“足够好”的解，而不是固执得非要最优解不可，那么，调和竞争性的观点，权衡不同价值的工作就变得稍许容易些了。如果我们不是那种强求最优的至善论者，那么，寻求满意很可能是（并且往往可能是），寻找那种能为社会差不多每个人所接受的方案，它们甚至是许多人甚至会喜欢的行动方案。

“囚犯两难问题”的对策模型有助于我们理解不确定性所产生的许多问题。假设有两名共同作案的罪犯被警察逮捕，警方指控他们犯有严重罪行。然而，如果他俩都拒不坦白，警方只能因证据不够，按轻罪发落他们。于是警察便单独劝说每位囚犯：如

果他肯和警方合作，供出他们的罪行，他就将受到宽大处理，对另一个顽抗的同伙将严加惩处。不过，如果他俩都坦白交代，警方将根据其实际罪行，给予他们公正的惩罚，这种惩罚当然将比其共同抗拒时为重，而比其中一人单独抗拒时为轻。那么，对于这两名囚犯来说，合理的行动策略应该是什么呢？

在不能串供的情况下，这种选择是很困难的。从囚犯甲的角度看，如果同伙乙招供的话，他最好也招供，免得落入重罚境地；而如果乙抗拒的话，甲的坦白将为自己带来较大的好处（当然以牺牲乙为代价）。于是，在这两种情况下，对甲来说，坦白交代都是合理的选择。同样道理，从囚犯乙的角度看，坦白也是合理的。但是，假如他俩都坦白的话，他们的结局将比共同抗拒时更糟糕。看来他们都陷入了两难选择的困境。

“囚犯两难问题”与核对抗问题具有某种惊人的相似之处。我们不清楚，怎样做才能使核对抗中的双方保持明智，使每方都采取克制行动，而不贸然发动第一次核攻击。不过，两难问题并不一定以这种极端形式出现，它也表现在许多存在利益冲突的双边对抗情况中——比如，在劳资谈判中，达成妥协避免罢工，对劳资双方来说，差不多都比陷入罢工更好些。不过，禁止罢工可能并不利于现行社会体制的稳定。

既使假定博奕过程不是不可逆的，而是可以重复进行的，也很难有助于问题的改善。在你的对手攻击你之前，采取先发制人的行动——至少以理性的许多定义来衡量——自然不失为先手之策。然而，对由被试者扮演的“囚犯两难对策”反复进行的经验研究，以及对采取不同策略的两名棋手之间对弈的计算机模拟结果，却显示出一幅不那么冷酷无情的图景。<sup>①</sup>竞赛者们经常采取较为仁慈的策略。他们这样做往往也可以得到对手适当的回报。

① ——参阅拉珀波特和钱玛合著的《囚犯的两难问题》（安阿保尔，密歇根，1965年）。阿克赛尔罗德所著《囚犯两难问题中的有效选择》，刊载于《消除冲突季刊》1980年24期：第13—25，379—403页。

在两种计算机策略的对抗赛中，针锋相对的策略特别有效。这种策略要求以其人之道还治其人之身：只要对手不先攻击你，你便对他采取仁慈行动；如果对手采取敌对行动，你也转而攻击他；其后，只要他放弃敌对行动，你也转而采取仁慈行动。

罗伊·拉德纳<sup>①</sup>已形式化地指出，如果博奕者的目标不是追求最优的报偿，而仅仅是满意水平的报偿，那么采取针锋相对的策略就是合理的。他的研究成果，为我们提供了一种关于人类经常倾向于选择这种策略的可能的说明。尽管如此，“囚犯两难问题”的基本范例仍向我们显示出：在存在着不确定性的局面中，理性机制是何等之脆弱；特别是，在存在着党派利益冲突的局面中，涉及其他党派活动的不确定性时，就更是如此。

### 制度理性的加强

以上所述的社会制度的理性限度，因起源于个人理性的限度，所以是相当基本的。然而，有些制度上的安排，能有助于我们更好地处理社会选择问题。组织，可用于处理有关决策之间的内部关系。市场结构，可以减少经济活动者对综合信息的需求。对抗式程序，可以保护被有意或无意忽视的有关事实和价值。社会制度中存在着一些众所周知的方式，这些机制和另一些机制可以用来加强社会选择的合理作用。下面，我想简要评述一下其中主要的几种。

### 组织和市场

首先，对于那些常规的、重复性的社会需要，可以通过建立专业化的团体和组织，使之平行地处理各种问题。每个组织处理一种问题，别的组织则同时处理其余的问题。如果这种现象不是

① ——私人通信。

如此明显的话，我们甚至可以称它为“组织理论的基本定理”。

其次，在一种广泛的范围内，我们可以用市场和价格，来限制每个人决策时所需要掌握的信息数量。我去本地超级市场时，无须过问面包和麦片的生产过程，也不必打听制造商有什么为难的问题，便可以决定需要买什么和想吃点什么。在这种个人决策过程中，我只需要了解那些可以买到的商品的价格。由于这个原因，市场和价格被证明是现代社会的极其有力的机制，因为它能帮助我们每个人独立地作出决策，而无须了解与此有关的所有的人决策的全部繁琐细节。人们为了进行市场交易，只需注意有关的市场价格，因为所有的相关信息都被概括在我们必须支付的价格中了。

这里我所讲的是另一种市场观，它与人们在某些经济学教科书中见到的那种最优化的市场观非常不一样。后者认为：在某些极严格的假定下（包括完全竞争和全面理性假定），市场可以导致帕累托最优——也就是说，社会达到不能再使任何人的福利同时增加的均衡状态，因为如果再增加一些人的福利，另一些人的福利就必须减少。帕累托最优并不是唯一的，还可能存在许多个最优解，不同的解取决于各种市场的参与者的不同偏好。不过，帕累托最优并不是我这里所关心的问题，我所提出的，是很多年前由哈耶克所阐明的那种更基本和一般的市场观。哈耶克认为，无需完全竞争和全面理性的假设，市场就能提供一种限制信息量的方式，使我们只按自己的需要做出决策，不必了解别人行动的细节。因此，即使最优可望而不可及，现实的市场机制仍能为我们提供一种过得去的社会安排的途径。

所以，我们可以把市场视为这样一种机制，它能使人类凭借有限的信息和计算能力，对社会和经济制度进行大体明智的操作。今天，我们看到了一个饶有趣味的现象：社会主义国家正在用广泛引入市场和价格机制的改革，来解决他们国内的计划和管理问题。为使价格成为资源配置的主要工具，他们正在试图把市

场问题与公有制和私有制的问题相分离。当然了，在讨论价格机制的这种用途时，我们必须牢牢记住前述外部事物是无时不有的。市场只能与其他社会控制和决策方式一同使用，它并不能提供一种独立的社会选择机制。

在存在外部事物的领域，我们将可以比我们今天在价格机制的有效利用方面有更多的创新发明。在这方面，经济学家们已提出了很多建议——例如，收取和烟尘污染程度在边际上相称的排烟费。不过，即使我们把这类外部处理程序扩大到其最大可行限度，也仍将保留许多外部事物（无论是消极的外部事物还是公共产品）。其中某些外部性，例如，当今时代的，那些事关重大的公共卫生或公共安全问题，总是被置于政府的直接管理或调节之下。

### **对抗式程序**

对抗式程序是加强制度理性的另一种方式。我想每个人都能说得许多法律程序的名字，尤其是法庭申诉和法庭辩论，以及对抗式程序。然而，在司法制度中，我们用得最广泛的还是对抗式程序，那里的理性准则是相当有意思的。司法的基本准则是遵循特定的程序，这种标准肯定是寻求满意而不是寻求最优。它的基本假定是，如果这些程序被遵循的话，那么，在某种长期意义上，就能实现过得去的，甚至称心如意的决策。因此，在法律制度中，我们倾向于对结果的评价委婉些，不像在程序公正性上那么直截了当。

对抗式程序的作用很像市场，它能减少每个成员采取合理行动所必需的信息。因而，它为司法制度提供了一种极其有用机制，因为在司法制度中信息的分布是相当广泛的，而且每个成员都有其不同的目标。对抗式程序中的每个成员，应该充分了解自己的利益及有关的重要事实，而不必知道其他成员的利益和处境。每个成员都陈述自己的理由，在此同时，也就使大家对有关

事实的认识和理解变得更全面和深刻了。

为了使对抗式程序充分发挥作用，必须足够宽泛地规定，在诉讼中什么人有权参加审理过程，以便使所有的受到重大影响的社会成员，都有机会陈述其证据和表达其偏好。我们的社会已日益承认相互依存的意见，法庭已逐步放宽确定哪些人能够参加哪些案件的审理的规定，甚至包括那些开始在两个特定成员间有争议的人们。在这种情况下，法庭考虑到那些类似在市场体系中出现的外部性问题。

### 决策的技术工具

近三十多年来，可用于多变量和多关联变量的决策工具已经取得了很大发展。这些新的分析工具通常属于运筹学和管理科学，以及当代的人工智能技术。这些分析工具最突出的特征是，它使我们能够建立公式和模型，来处理拥有成千上万的变量和约束的大型复杂问题，并且可以考虑所有变量和约束在解题过程中的相互作用。

运筹学和管理科学技术在应用方面的一个重大限制是，它们要求问题必须能数学化，即已知的数学工具可适用于所欲求解的问题。例如，如欲用线性规划来求解问题，就必须首先把问题翻译成（或叠加成，或压缩成）以线性方程式、线性约束式和线性目标函数表达的形式。如果现实世界不具备这些性质，或者不能用这种方式适当地近似，就不能运用线性规划来求解问题。相反，人工智能技术一般却没有给问题强加数学化条件，因而可以对付定性因素，处理完全定性的问题。这就使得现代计算机在提高人类分析能力和拓宽分析问题的范围方面，有了实质性的扩展。

尽管存在种种限制，这些新的分析工具，还是它使我们具备了考察这个世界的某些复杂问题的付作用及其相互影响的能力，而这在战后初期引入这些分析工具之前，简直是无法想象的事。

情。假定我们找到了环境保护和能源等难题的解决办法（我认为我们很可能找到）——也就是，同时处理两组相关问题的解——那很可能是由于我们能建立一种模型，它能反映上述问题的众多侧面之间的主要相互关系，从而有助于我们清醒地权衡利弊之故。

新的分析工具，至少标志着一种有限的进步，同时，它也至少为我们对于世界上日益复杂的问题进行处理的能力，提供了一个乐观主义的依据。

## 公共信息库

在重大政策问题上，我们的制度进行合理决策所涉及的另一个重要问题是，我们的知识和信息是否正确和适用。我在第一章中论证了，推理作为决策工具的有效性，关键取决于作为其输入的事实——数据、知识和作为其假设的理论。除非这些输入是有意义的，否则处理它们将一无所获。如果你把错误的数据和知识灌输到人的头脑中，那么你所得到的只能是一些与事实相反的错误结论。

### 新闻传播媒介

作为公共政策决策之用的知识库怎样才算是适用的？我想，我们大家恐怕将不约而同地指出新闻报道的反复无常，并说得出自己所讨厌的某个最可恶的新闻媒介的名字。看来异口同声的意见是，如果将新闻媒介用作公共政策制定过程所需事实和知识的主要来源，是有着严重缺陷的。

或许，最基本的缺陷在于，新闻媒介的视野很难超越当前的新闻和时尚。它们强调新闻价值，追逐耸人听闻的、猎奇性的报道。在这方面，电视所起的作用甚至比老式的传播媒介更为糟糕。因为它不仅可以左右地区性的舆论倾向，而且还可以影响全国或全世

界的注意力。然而，就连那些老式的传播媒介也仅侧重于报道新闻消息，而不是增进对事实的理解。比如说，要是有人打算探讨美国对华政策的有事实依据的观点，我建议他最好能去读一两本有关的好书，这要比去阅读明年全年的《纽约时报》上所有有关中国的报道效果好得多。因为报纸只能给他一些短促性事实的杂录，书籍则能为他提供一种稳定的，仅发生着缓慢变化的知识框架。通过这种背景知识，当前事态的发展就变得有条理和易于理解了。

在这个例子中，为了研究对华政策，真正需要知道的是中国社会制度和历史——这种正确的信息是难于从新闻媒介中获得的。新闻媒介正忙于告诉人们今天或本周所发生的活动。然而，今天中国所发生的事件，恰恰是中国社会的基本特征和倾向的产物，那些没有读过有关内容的适当书籍的人们，对今天中国发生的事件是很难有正确的理解的。

从新闻媒介可以得到的信息，大多只具有短暂的价值，假如人的注意力不是一种极其稀缺的资源，那么这种短暂性就算不上什么大问题。然而，人们的时间如被花在读报或者看电视上，就不可能再用以获取概念结构和背景知识——这正是那种能使短暂事件变得更易于理解的信息。我想，假如一个社会对注意力的稀缺性非常敏感的话，它就应该改变其社会成员的阅读习惯，以便更有效地分配注意力。尽管在我们自己的社会里，人们对信息泛滥的抱怨已相当普遍，但还没有证据表明，他们正在认真设计对策，以减少那些大量出现的昙花一现的事件的侵扰。不“因为报上登了”就轻信，不简单地靠新闻媒介来获取新信息的想法，对大多数人来讲恐怕还是个新思想。

## 专家

然而，即便我们从所需知识清单上排除了那些仅有短暂价值的信息，我们还有一个怎样选择阅读书目的问题。一位认真负责

的读者，在读完了有关中国的书后，还必须读关于阿富汗的书，以及没完没了的其他书籍。我们都感觉到获得适用知识所面临的无法克服的困难。但是，我们还是有处理这种困难的办法。在公共政策事务中，就像我们找医生看病或者请管子工修管道一样，我们也转而求助于专家。当我们难以确定有关事实时，我们就去找了解这些事实的专家，听取他的意见。有时候，就干脆采纳他的建议，而不再刨根问底。

我们怎样寻找真正的专家呢？怎样鉴别那些合格的专家并给予其合法地位呢？在这些方面，虽然我们并不总是干得很出色，但是，我们的社会和其他发达的社会一直在学习如何干得更好些。例如，美国国会愈来愈多地向美国科学院及其附属组织——美国工程学会，美国医学协会，以及美国研究委员会——进行咨询和索取资料。这些研究组织在审议同当前公共政策有关的课题中，能够找到并利用我国的大部分科学、医学和工程技术方面的专门知识。

然而，我们怎样确信这些专家（或另一些专家）是适当的人选呢？怎样确保这些专家实事求是，即确保他们所提的建议不受其自身利益的影响呢？在某种程度上，带有利益色彩的建议问题是易于解决的；从另一种角度讲，则是很难处理的。在向专家们索取信息和咨询时，要求他们公开那些可能产生利益冲突的财务关系或专业职务是容易做到的。政府在通过上述组织寻求咨询时要求专家们公开这些关系，在我们今天的社会还只是一种例行公事。

但是，利益冲突的更微妙的问题则直接产生于人类的有限理性。事实是，假如我们被卷入某项具体活动，并为从事该项活动花费了相当精力的话；那么，就可以肯定，这项活动将比我们不曾与它打交道时，具有更大的价值和重要性。如果某人的职业是设计核电站，那么我敢打赌，他不大可能在反对将核电站建在你们镇上的请愿书上签字。而你或许连问都懒得问他，就直接

到别处去征求反核运动的支持者了。

我们往往极难从这类事实中得出结论，说人类是一种非常虚伪的动物。人们往往是干什么就吆喝什么。这种由职业和先入之见所造成的许多偏见，正确地说不应被描述为虚伪的根源——这或许使得这种不虚伪的偏见比真正的虚伪更隐蔽和危险。

人类看不到整个世界，我们所见到的不过是其所处的一小部分世界。人们能够虚构一个局部世界的形形色色的合理性，并极大地夸大了它们的重要性。让我们继续讨论核电站的例子，因为它向我们提供了此类现象的许多非常好的说明。十几年前，当利弗摩尔①实验室的两台呼呼作响的通风机所产生的一些统计数字显示出，核电站附近的辐射性对人身健康的危害，实际上比以往所认为的还要大时，人们的最初反应是赞成关闭核电站。几乎无一例外，没有人说：“让我们更仔细地调查一下，让我们指定一些德高望重的人组成委员会去调查事实真相”。相反，公众的普遍反应几乎都是：“那些不负责任的家伙们为什么还在夸夸其谈呢？”

当时，我作为总统科学顾问委员会的成员，与这些事件关系较为密切。回想起来，处理事件的许多“圈内人”对公众深切的忧虑竟然麻木不仁得惊人。许多处理这些事件的“圈内人”是我的朋友或熟人，他们都是一些非常正直的人，其品质不容怀疑，绝对不可能被收买。蒙蔽他们的，使他们对公平地看待事实的要求视而不见的，是他们已有的“知识”，即那些伴随着核能能源的多年发展所获得的核技术知识。他们深信核技术会造福人类，能开发新的生产力，能减轻我们对日益枯竭的化石燃料的依赖性。并且，发展核能对人类健康绝对不会产生不可预见的、无法解决的那种异乎寻常的危险。正是这种对核技术的深刻

---

① 位于加利福尼亚州奥克兰附近，美国有名的核反应堆建在那里。——译注

了解，妨碍了他们客观地考虑证据是否有利于他们。

在问题变得争议极大的情况下——当问题被不确定性和冲突的价值所包围时——就极难得到中肯的专家意见了，而且也很难使专家的地位合法化。在这些情况中，我们发现既有持肯定意见的专家，又有持否定意见的专家。我们不能把这类问题委托给某些特定的专家组去处理。我们最好能把争议转换为一种对抗式程序，在这种程序中，我们这些外行们听取专家们的 意见，但必须在不同的意见观点之间进行独立判断。

### 关于政治制度的知识

妨碍我们有效地参与政治过程的知识缺陷，是我们对政治制度本身知之甚少。对那些为有效和负责地参与政治过程而真正需要了解的有关人类的大量情况，我们却仅仅是一知半解。

例如，政治制度的正确设计，往往取决于对人性的可改善程度的合理估价。某种新的政治制度和经济制度安排(社会变革)，如果要行得通的话，只能假设处于其中的所有的人或大部分人愿意采取利他行为，或至少顺从社会需要。有人认为：新的社会必然产生崭新的人。在我们的一生中，至少见过两场重要的社会革命——俄国革命和中国革命——是建立在假设革命能通过变革制度改变人的行为这一推断上的。而我们大部分人都相信，这两场革命并不曾如所预期地改变了人的行为。问题仍然是：在现实中，存在不存在某种社会制度的变动，能引起人类行为基本方式的改变——就是说，使他们变得更利他些，或更自觉地服从法律。这个问题恰恰也是改造罪犯的讨论中争论不休的问题。

第二章的进化论的观点提出，人的行为确实可以被制度所改造，至少在增强或减少那种微弱的利他主义方面。而那种产生持久而深刻变动的可能性——就是说使人性变得更驯服——是相当成问题的。迄今为止，几乎没有什么经验基础可以对这些问题予以明确答复，除非把那些从革命中得到的反面教训作为确

证。

不过，相对来说的重要问题，倒是还存在着一些科学知识，那些被政治科学家们积累和检验多时的知识，可以有助于我们设计和选择更有效率的政治制度和政治程序。在我们的社会里，我们有一种不幸的嗜好，即习惯于为我们的政治制度贴上两种不同的标签。当我们喜欢它们时，便称其为“民主”；当不喜欢它们时，便称其为“政治”。在贬义的用法中，我们不愿选择的那种公认的“政治”，其实只是标志着我们恰好不抱幻想的民主政治制度的某些特征。实际上，无论“政治”或“民主”均不能完整地描述政治制度的全部特征，用这种随意为自己想要的或不想要的东西贴标签的方法，是解决不了任何问题的。

几年前，我受聘担任一个委员会的主席。该委员会负责审查宾夕法尼亚州的有争议的控制牛奶价格法案。委员会的成员包括奶牛场的农场主、牛奶商、送奶车司机工会的官员，还有两位据说是代表消费者的委员，和两位与乳业无直接利益关系的“公众代表”。在委员会的会议上，几乎总是出现几位委员拍着桌子慷慨激昂地大骂“政客们”的场面。实际上，有些时候，差不多所有的委员们都加入了这种对“政治家”的攻击和漫骂。他们这样做完全出于不自觉。因为他们忘记了，作为该委员会成员，他们自己也是“政治家”（更何况他们中许多作为院外活动家也是政客）。对他们来说，“政治家”或“政客”简直就是一种诅咒，一种根本没有想象可能加剧自己头上的污辱性的词。

实际上，这种对政治和政治家们的幼稚想法遍及我们整个社会。它对我们的政治制度是非常有害的。如果我们能更成熟老练的话，我们就能够更好地观察政治和社会制度，也就能更好地识别其肌体上的痼疾。虽然我们可以尽力消除那些附赘悬疣，然而，我们必须认识到，某种政治现象——那些对立法或执法机关施加影响的企图，那些对特殊利益的辩护——对于一个社会的政治制度的运行乃是不可缺少的。因为实际上，社会存在着多种多样的利

益，并且可以预料，社会中的大部分人，更为关心的是其自身利益。我们所谓的“政治”活动，只不过是人们为使社会认同其个人目标，并企图以合法方式实现这些目标的另一种表现形式。

我们所抱有的一些关于选举的基础观念，与上述我们对政治和民主的谬见有关的是，我们的社会普遍相信（或者至少人们的行为符合这样的信念），选民在电视上见到和听到竞选人的言行后，就可以决定在选举中，他将投谁的票。与此相反，大量的社会心理学的试验证明，人们观察他人时（尤其是在观察那些谈吐中流露出企图对他人施加影响的人们时），是非常难以判断其言语的含义，以及那些隐含在行动中的意义的。比如，有很好的证据表明，如果观察者已倾向对支持某一位竞选人，他将把该人的竞选宣言解释得符合自己的立场，而不管其本意如何。反之，如观察者不打算支持某竞选人，那么他将把同样的竞选宣言解释成不符合他们的立场。

所以，我们的社会相信的选举过程是，大家都盯着自己的电视机，听着竞选演说，出于某种莫名其妙的原因，便可得到了自己应当投谁的票的信息。我们甚至还有一个关于投票行为的口号：“投某人的票！不投某党的票！”如果说，我们对某位成功的竞选人在当选后将会作什么决策有兴趣的话，那便是，我们特别希望这些决策尽可能地与我们自己的价值判断相一致。我们从电视或别的新闻媒介中了解竞选人的个人品质，是否能比了解其党派的背景更好地判断其当选后的所作所为呢？从我所知道的所有证据，以及其他大量的证据来看，竞选人的党派背景才是更为可靠的判断依据。

美国公众的那种不断增长的作为无党派选民的自豪感，已经大大削弱了美国的政党。这不仅扩大了美国政治制度易受蛊惑的弱点，而且也增加了制定和执行公共政策的困难，尤其增加了形成以团结选民偏好为基础的多数派的困难。在近两百年中，政党一直做为合理有效的机制发挥作用，在妥协与谈判的复杂过程中

中，它们以形成多数派和制定公共政策的方式来处理社会难题。今后，它们可能不复如此了。那种基于单子主义幻觉的独立选民观念，正在摧毁着政党组织所维系的竞选行为的可预见性，降低了公众的理性水平。

我倒是很希望能有一付医治民主社会的那种不负责任的公民行为的万灵药方。在政治舞台上，那些受过教育的大学生们虽然上过标准的政治学和经济学的入门课程，其表现却并不比那些没受过这类教育的选民们更成熟老练些。这话既可以理解为对这类课程的评论，也可以理解为对学生们难以造就的议论。不管怎么解释，显而易见的是，我们还没有找到进行公众教育的有效手段。我们还不曾发现一种方式，能用人们愿意花在公众教育上的有限的时间和注意力，使人们了解我们的政治制度如何运行，从而产生出一种成熟的公众理性水平。目前我们还不了解，选民是怎样有效地吸取竞选信息——那些关于竞选问题和竞选人的有用的信息；或者，他们应该怎样去选择自己信得过的专家。

如果我们忽略政治过程中产生的犬儒哲学，那么“政治”的贬义用法恰恰是政治制度的一种征象，而倍受尊崇的“民主”只不过是处于理想和现实的尖锐冲突中的一种观念。或许，对那种犬儒哲学最好的解毒药是，向人们揭示民主政治制度的现实图景，并教给人们实现这种制度可行目标的正常的讨论是什么。但是，我在新闻媒介和教育机构中都几乎看不到有这种改进的迹象。

### 知识是万能的吗？

在公共政策的重大问题上，我们是否具备可以进行合理而明智决策的足够知识呢（如果我们将发现它们，就能利用）？答案形形色色，因情况不同而异。这里我想举三个例子，它们多多少少可以组成一个连续的故事。

第一个例子是当今世界上最引人注目的、生死攸关的战争

与和平问题。对此，我们实在只能抱悲观主义态度。因为没人能搞清楚，为了比现在更容易看清和识破维持和平政策的大量复杂性，我们应当搜集哪些信息和知识，或者应当进行哪类科学的研究。如果我们同时还有好几个目标的话，就像我们大部分人那样，那么维持和平就会变得格外困难。我们既想维护和平，又想维护我们的制度和自由的基本特征。我实在想像不出应该对我们的实际知识应进行何种改进，才能减轻这些问题的复杂程度。

困难的主要根源在于，战争与和平问题不仅涉及到我们自己在变化莫测的环境中的行为的不确定性，而且还涉及到其他民族的行为的不确定性，以及我们所进行的那种斗智斗勇的竞赛的不确定性。我不知道凭我们现有的科学知识，怎样才能科学地处理这些问题。

不过，谈到我的第二个例子，即有关能源和环境问题时，我们就可以大为乐观了。我们目前已发现了整整一大批研究和发展程序，这不仅能有助于我们更好地理解已知的技术方案及其后果，而且也拓宽了我们的抉择范围。例如，比起50年前，我们对二氧化碳的增加对大气和气候的影响，或酸雨对植物生长和湖泊生物种群的影响，已经了解得更多了。而且，我们对如何解决这些问题，也了解得更多了。

我的第三个例子是经济政策问题，可以认为它处于乐观主义与悲观主义之间的某种位置上。原因在于，经济体系运行的关键，取决于人们对未来的预期，以及人们对这些预期的反应方式，这是一个相当困难的研究领域。

目前流传着一种时髦说法即如果有五位不同的经济学家凑到一起的话，那么，在关于对经济怎样运行以及应该怎样改善其运行的问题上，就会有五种不同的观点。在某种程度上，这是对的。通过适当地选择专家（至于要求不要求所有的专家们都拥有经济学博士学位，随你的便），就能得到你所想要的有关国民经济政策的任何一种意见。经济学家们之间的意见分歧主要限于一小部分

关键性的问题上，并主要集中在人们如何形成对未来的预期上。供应学派经济学家会对你说，如果你使得投资有利可图，不管是用降低利率，还是用减税或别的什么办法，投资都将有实质性的增长。合理预期学派经济学家也会对你说：在有关未来的问题上，人们是不可能被愚弄的，人们的预期代表着他们对经济体系均衡位置的运动趋势的现实估价。凯恩斯主义者对预期则有其他不同的假说。

那么究竟谁的意见是对的呢？不幸，我们不得而知。我们完全不具备关于人们怎样形成其预期，以及预期怎样影响人们的事实依据。而这正是我们为检验供应学派、合理预期学派、预算平衡论者、货币主义者、或凯恩斯主义者的假说所需要的经验依据。预期问题是当今各种不同的经济思想之间意见分歧的主要领域。它并不是一个很大的领域，然而，它却在经济理论结构中，在经济理论对公共政策的应用中，占据了一种令人不快的战略位置。

正如这三个例子所表明，生机勃勃的自然科学和社会科学的研究和开发工作，可以在那些以知识为主要限制因素的决策领域，给我们以重要帮助。然而，科学知识并不是那种能够解决一切问题的“点金石”。

## 结语

我一直主张：人类理性，较之它作为探索特定的局部需要和问题的工具而言，远不足以成为构造和预测全世界系统的一般均衡模型，或者创造一种包罗一切时代的所有变量的宏大总模型的工具。从隐含在主观预期效用理论（SEU）的理性模型的那种奥林匹亚山神般的图景中，我几乎见不到多少有益之处。我在第二章提出的反对纯利他主义的进化论中指出，在形成公共政策和我们自己的个人决策中，或许有理由假定，作为一种基本的近似，人

们的行动是以个人利益为出发点的。这样，任何社会的主要任务就是创造一种使个人利益有理由变得更明智的社会环境。倘若我们希望由“看不见的手”把社会带到某种和谐状态，我们就应该相信，首先，我们的社会制度应更有利于我们的自我实现；其次，这个制度并不要求很多人长久地牺牲较重要的个人利益。

理性，就其本身而言，是工具性的。它既不能代我们选择终极目标，也不能调解我们在追求终极目标过程中的价值冲突——我们必须用其他方式处理这些问题。全部理性所能做的，仅仅是帮助我们更有效地达到既定的目标。不过，至少在这方面，我们可以学得更好些。在一些有限的范围内，人类理性力量本身也取得了一定程度的发展，特别是我们同时处理几种关系的能力。在我们的理性工具中，这些新的发展可以说是代表了人类思想的一种质的变化，正如书写文字的发明，使我们能把自己的思想形诸笔端，从而使我们具备了应付新的复杂问题的能力一样。上述新的发展，也使我们预测自己行动后果，以及设计新的方案的能力，取得了并不断取得着进步。当然这些进步尚远不足以应付世界的全部复杂性。但幸运的是，这个世界，甚至现代世界，很大程度上是空旷的。世界上的大多数事物与其他的事物之间，仅存在微弱的联系。人类的理性所需要对付的仅仅是这样的一个世界。

我们的社会，或其他什么社会，都可能达到某种没有任何危险的稳定状态，在这种状态中一切问题都已获解决。无论怎么看，这种状态都是相当单调乏味的。因此，我们非常有必要为子孙后代保留足够宽广的选择空间，就像我们的祖先们曾传给我们的那样，使我们的子孙们有机会在解决足够的发展问题中发现乐趣，从而不至于抱怨他们的天地比我们的更狭窄。<sup>1</sup>在我看来，这恐怕是比现世的乌托邦（甚至来世的乌托邦）更现实可行的社会政策的目标。这比起下述假设也更为合理些：即认为我们称为“人类问题”的东西是与所谓“解（答案）的东西”相对应的，并且一旦

我们发现了答案，问题就将不复存在了。

在社会实现其有限目标的过程中，提倡明智的利己主义有足够的理由吗？这取决于我们给明智所加的约束。成功取决于我们拓宽人类视野，从而使人们在决定自己利益的过程中，考虑更为宽泛的后果的能力。取决于我们每个人能否认识到，我们自身的命运与整个世界的命运休戚相关。也取决于我们是否认识到，明智的或过得去的利己主义，不可不与我们所生活的整个环境协调一致。