

## 第 1 章 什么是制度？ 我们如何理解制度？

在人类社会的大棋盘上，每个个体都有其自身的行动规律，和立法者试图施加的规则不是一回事。如果它们能够相互一致，按同一方向作用，人类社会的博弈就会如行云流水，结局圆满。但如果两者相互抵牾，那博弈的结果将苦不堪言，社会在任何时候都会陷入高度的混乱之中。

——亚当·斯密《国富论》，1759 年

传统上，经济学家一直致力于分析市场机制的运行及其影响。毋庸置疑，市场可以被认为是人类所创造的最引人注目的制度之一。然而，在近来大约十年的时间里，为了理解不同国家迥然不同的经济绩效，认识到“制度是重要的”这一点已变得日益重要(North 1990；世界银行 2001)。这里所引的制度一词，不单单是指市场。的确，在 20 世纪最后十年所发生的一系列与制度相关的事件和现象已经并将继续对相关国家的经济绩效产生深远的影响。举例来说，苏东社会主义国家的崩溃以及随后的经济转轨，硅谷现象及电子商务的出现，欧元统一和市场一体化，日本和东南亚金融危机，非洲持续的种族隔离和长期的经济停滞，金融市场的全球一体化与不断出现的货币危机，对由成员国组成的国际组织作用的重新审视，以及全球非政府组织影响的加深，等等。这些例子当中的一些就其表面而言可看作是纯粹的市场现象。但如果我们试图更深入地理解这些事件和现象的深层原因及后果，我们不得不将它们的制度方面纳入考察的范围。

什么是制度？我们能不能把制度等同于法律条文、非正式规范、组织、合同、人们的意识或所有这些因素的部分或全部的组合？给诸如“制度”之类的任何概念下一个合适的定义将取决于分析的目的。作为一个具体例子，我们不妨考虑下面的问题：既然制度对经济绩效如此重要，那为什么其他国家不能学习和采用经济绩效较好的国家当中最佳的制度呢？这是诺斯在一本开创性的论制度的书中提出的主要问题(North1990)。为了分析这一问题，诺斯把制度定义为“博弈规则”。他把博弈规则分为两类：正式规则(宪法、产权制度和合同)和非正式规则(规范和习俗)。即使能从国外借鉴良好的正式规则，如果本土的(indigenous)非正式规则因为惰性而一时难以变化，新借鉴来的正式规则和旧有的非正式规则势必产生冲突。其结果，借鉴来的制度可能既无法实施又难以奏效。

因为这个原因，经济学家对可实施性问题产生了浓厚的兴趣。博弈规则何时是可实施的呢？是不是当实施者(enforcer)降临之时？可是，实施者又怎样才能被激励去实施他理应实施的博弈规则呢？简而言之，实施者怎样才能被驱使去恪尽职守？为了避免这种无穷尽的循环推理，一种办法是力图说明，博弈规则是内在产生的，它们通过包括实施者在内的博弈参与人之间的策略互动最后成为自我实施的(self-enforcing)。从这种观点出发思考制度的最合理的思路是将制度概括为一种博弈均衡。近年来我们陆续看到一些以博弈均衡制度观为基础的重要文献问世，虽然其中大部分文献是从历史案例的研究中汲取灵感(我们在下一节会讨论其中一些代表性工作)。我们能不能将同样的思路运用到作为多样制度复合体的当代经济呢？这种复合体仅仅是各种相对自主的制度的混合呢，还是内在一致的整体，如某种均衡结果那样？

显然，把制度及其复合体视为均衡现象并不意味着制度是一成不变的；它们会发生变化。中欧和东欧国家社会主义经济的崩溃和随后计划经济的转型就是一个显而易见的例子。问题是，我们如何在理论上解释制度的起源或变迁。在理论上，博弈模型可能存在多重解(均衡)，或者说模型解高度依赖于对模型本身的设定。给定博弈的结构，可不可以将制度的突现(emergence)或变迁解释为：从许多等可能性的均衡中选择其中一种的过程，或者从一种均衡到另一种均衡的转型呢？如果是这样的话，均衡选择或转型过程是否由技术或市场诱导并因为技术性规模经济而最终被锁定(lock-in)？还是说，制度是由“文化基因”编程的结果？制度能否被政治企业家设计或被创新性经济企业家引发产生？不可预料的政治事件会不会对制度选择发生随机性影响？尤其是，新制度诞生通常具备的新奇性究竟从何而来？

因此，在本节一开始提出的基础研究设想可以更明确地概括为下述两个方面：首先，将当代不同国家整体性制度安排的复杂性和多样性理解为某种多重均衡现象(共时性问题)；其次，在与均衡制度观相一致的框架下理解制度变迁的机制，同时又允许新奇性出现的可能性(历时性问题)。请让我再进一步阐释其中的具体含义。

理解制度的多样性和当代经济的复杂性需要研究在经济、政治、组织和社会诸域制度之间的相互依存性以及联结这些域的制度的性质。在这项研究中，不仅要在正统经济学的框架下思考问题，而且还要借鉴邻近学科如社会学，政治学和认知科学对制度问题的重要贡献，这是非常必要的。但不同于传统制度经济学的地方在于，我们试图在一个统一的博弈论框架下分析制度多样性的源泉和影响，而不只是简单停留在积累丰富的制度类型、然后对它们进行任意和随机的分类。发展统一的理论分析框架，同时结合其他学科的重要贡

献，这对深刻理解经济制度的运行大有裨益。

与此同时，我们也必须意识到，博弈论分析作为系统研究制度的理论工具本身尚不完备。从该框架出发考察制度的相互依存性可能会得出制度安排的多重性、次优性和帕累托不可比(Pareto-unrankable)。也就是说，即使面对相同的技术知识和被相同的市场所联结，制度安排也会因国家而异。因此，为了理解特定的制度安排在某特定国家演化生成的原因，单单囿于博弈论框架本身是不够的，我们必须依赖比较和历史的知识(Greif 1999)。也就是说，制度分析在本质上是比较性的，因而被称为**比较制度分析 (comparative institution analysis 简称 CIA)**<sup>1</sup>。

在本书第 9 章考察制度演化的历时性过程时，我们将偏离传统的博弈论(包括古典博弈论和进化博弈论)，对其作出重大的修正。我们将放弃如下假定，即认为博弈参与者(及其分析者)对博弈过程的客观结构具有完备的知识。相反，我们假定，博弈参与者对于博弈结构只拥有个人的不完备观点——我们称之为主观博弈模型(subjective game models)。当参与者基于主观博弈模型选择的行动决策在各个时期相互一致(即均衡化)时，那么，他们的主观博弈模型将可以被他们行动共同决定的可观察的事实证实，并作为未来行动决策的指南而不断再生产出来。所以，我们把制度定义为参与人主观博弈模型中显明和共同的因素——即关于博弈实际进行方式的共有信念(shared beliefs)<sup>2</sup>。当这些主观博弈模型所导致的行动决策未能产生预期的结果，一种普遍的认知危机便会随之出现，并引发人们寻找新的主观模型，直到新均衡实现为止。理解制度变迁过程就等价于理解参与者协同修正其信念的方式。从这种观点出发，我们能够分析技术和环境变化、政治因素、法律条文、创新试验和文化遗产等等方面在制度变迁过程中的作用，当然这一切必须在我们分析了共时性问题之后才真正有可能进行。

本章剩下的部分安排如下：第 1.1 节是对经济学家制度观的一个概览。第 1.2 节将在均衡制度观的基础上提出我们自己的制度概念，其更细致地讨论将在后面(第 7 章)进行。第 1.3 节介绍制度分析的博弈论框架将涉及的一些基本概念，如“博弈形式”和“域”的类型等，它们在本书中是至关重要的。

### 1.1 博弈论视野下的三种制度观

如上所述，经济学家目前似乎都同意“制度是重要的”。但除非我们对于制度是何物以及它们如何形成有一种共识，否则，认识到制度重要并不能说明什么。撇开旧制度学派不

论<sup>2</sup>，直到最近，经济学家基本上满足于对制度只下一个模糊定义。现代社会学的先驱杜克海姆曾经定义社会学为“关于制度的科学”，定义经济学为“关于市场的科学”<sup>3</sup>。经济学家一直忙于有关市场的分析研究，杜克海姆的学科划分似乎从未让经济学家感到不安。然而，经济学家对理解制度的性质、起源和影响可以做出自己独特的贡献。事实上，越来越多的经济学家近年来开始从事制度研究的任务。我们下面将看到，经济学家赋予“制度”一词至少三种不同但相互联系的含义。我们这里所关心的当然不是对制度一词进行语义上的澄清，而是寻求一种有助于更好地理解经济制度运行的理论概括。

为了区分经济学家提出的三种制度含义或定义，将经济过程比喻成博弈可能是合适的。我已经指出，博弈论是比较制度分析不可分割的一部分。本书从进化博弈论和重复博弈论借鉴来的研究共时性问题的博弈分析工具是相对较近才发展起来的。但是，从博弈角度理解经济过程的思想至少可以追溯到亚当·斯密，本章开始的引语就是明证。在亚当·斯密看来，博弈是个体参与人从各自的动机出发相互作用的一种状态，这正对应着现代博弈论所研究的情形。

通过将经济过程类比于博弈过程，不同的经济学家分别将制度看作是博弈的参与人、博弈规则和博弈过程中参与人的均衡策略。人们在日常交谈中所涉及的制度，通常是指重要的组织机构。一些经济学家沿袭这种习惯，将制度明确等同于博弈的特定参与人，诸如“行业协会、技术协会、大学、法庭、政府机构、司法等等”（Nelson 1994：第57页）。诺斯支持第二种观点：即制度应该被视为博弈规则，以区别于它的参与人<sup>4</sup>。他在论制度及制度变迁的开创性著作中开宗明义：

“制度是社会的博弈规则，或更严格地说，是人类设计的制约人们相互行为的约束条件……用经济学的术语说，制度定义和限制了个人的决策集合”  
(North 1990：第3至第4页)。

这些约束条件可以是非正式的(如社会规范、惯例、道德律)，也可以是有意识设计或规定的正式约束。正式规则包括政治规则(宪法、政府管制)、经济规则和合同。经济规则用来界定产权，即使用和处置经济资源并从中获取效用或收益的权利束(the bundle of rights)。合同是一种关于物品使用和交易的(可执行的)协议，它受产权规则的制约。经济博弈的正式规则不可能被正在博弈之中的参与人自己制定(变更)，它们的确立必须先于博弈过程。由于我们关注制度的起源，马上就面临一个问题：谁来制定经济规则？正是在

这里，诺斯对博弈规则和博弈参与者(组织及其政治企业家)作了明确的区分，后者是推动制度变迁的主体，即规则制定者。根据诺斯的观点，现存的博弈规则决定了参与者如何交易及创新的激励，因而在根本上导致了伴随相对价格变动而产生的对新规则的有效需求。这些新规则将在“政治市场”上经各方协商而确定，政治市场则由政治规则决定。诺斯声称“正是政治过程本身界定和实施产权。”(North1995，第23页)<sup>5</sup>。

关于博弈规则论更为技术性的定义是由赫尔维茨(Hurwicz 1993, 1996)给出的，他的定义更侧重于博弈规则的实施问题。根据他的观点，博弈规则可以由参与者能够选择的行动(“决策集”)以及参与者决策的每个行动组合(profile)所对应的物质结果(“后果函数”)来刻画。他将这一对设定称为“机制”(mechanism)或“博弈形式”(game form)<sup>6</sup>。作为一个例子，不妨考虑一种价格限制机制，即由政府规定销售商所能索价的最高限制。在这种情况下，销售商所面临的决策集的限制可以表示为特定的参数值，即限制价格<sup>7</sup>。然而，根据赫尔维茨的观点，关于制度更为合适的定义应该对它施加进一步的限定。他认为规则必须是可实施的，或者用他的术语说，是“可执行的”，惟有对人类行动的一组人为的和可实施的限定才构成一项制度。他运用纳什均衡概念使可实施性这个概念形式化。如果在别人将遵从所设定的策略的前提下，没有任何一个参与者有偏离其选择策略的动机，此时参与者的策略组合便被称为是纳什均衡。为了使博弈形式中一组人为设计的限定成为可实施的，而参与者又允许从所有技术上可行的决策集中自由选择其行动，该博弈必须存在一个纳什均衡。

赫尔维茨主要关注的问题是探究“设计”一项实现既定社会目标的制度的可能性，其中制度在一系列环境(技术、偏好和资源禀赋)下与参与者的激励是兼容的。社会目标(如效率、公平、清洁的空气和水源)可以表示为在每一种经济环境下预期实现的一组结果。假定由立法者负责设计实现既定社会目标的机制。我们无法事先保证机制是不是可实施。例如，立法者也许希望价格限制有助于实现价格稳定和分配公平的社会目标，但是，总是有销售商会发现将产品以高于限定价卖给黑市更有利可图。价格限制无法自我实施，因而不是可执行的。

如果一种机制为了达到某种社会目标被设计出来却无法自我实施，那就需要附加一种额外的实施机制。添加具备特定行动集合(如将犯人投入监狱等)的实施者(如法庭、警察、调查员等)必然改变博弈形式，后果函数也将相应地加以变更。然而，这种情况恰恰给机制设计者制造了一种困境。为了使实施机制行之有效，一方面，实施者必须被给予适当的激励，使其忠于职守。另一方面，实施机制的运行消耗社会资源，从而相应减少直接为社

会目标作贡献所需的资源。其结果，最初的社会目标的实现程度将不得不大打折扣。

当赫尔维茨在考虑实施者的激励问题时，他对制度的认识实际上已经接近第三种观点，即关于制度的博弈均衡观。第三种观点最早的倡导者之一是肖特(Schotter, 1981)<sup>8</sup>。近年来在博弈均衡制度观方面有两项主要进展，其中每一项进展都基于不同的博弈均衡概念——进化博弈论(evolutionary game approach)和重复博弈论(repeated game approach)。前种的代表性的工作是萨格登(Sugden, 1986, 1989)、扬(Young, 1998)和鲍尔斯(Bowles, 2000)<sup>9</sup>。进化博弈论认为，参与人的行为习惯可以自我形成，不需要第三方实施或人为设计。当惯例(convention)演化时，参与人在进化选择的压力下，倾向于发展某些适应性更强的特征(traits)(如环境认知、偏好、技能等)。这样，惯例和参与人的特征共同演化。惯例也许最终会以法律条文的形式固定下来，从而节约了因变异和错误带来的失衡成本。另外，用文字清晰表述人们行为中已经习惯化的准则也有助于给出具体环境下具体的行动指南。但是，萨格登继承休谟的传统，认为将法律理解为源于政府限制公民的行动是误入歧途的。相反，“法律反映了大多数人自愿施加的行为准则”(Sugden, 1986: 第5页)。

第二种博弈均衡制度观得力于格雷夫(1989, 1994)、米尔格罗姆、诺斯和温加斯特(Milgrom, North 和 Weingast, 1990)、格雷夫、米尔格罗姆和温加斯特(1994)和卡沃特(Calvert, 1995)的工作，他们运用了一些较为复杂的均衡概念，如重复性囚犯困境博弈下的子博弈精炼均衡(subgame perfect equilibrium)。有关子博弈精炼均衡的精确定义将在本书的稍后部分(第7章)给出。不过，现在值得强调的一点是，子博弈精炼均衡和其他相关均衡概念有助于澄清预期或参与人共有信念在博弈中的作用。子博弈精炼均衡为每个博弈参与人界定了一种行为策略，该策略是在所有可能的博弈状态下行动决策的完备计划，简单地说就是一部“脚本”<sup>10</sup>。完备行动计划的任何一部分，即每种特定可能性下所规定的行动决策，必须在该可能性化为现实之时也是纳什均衡，因而能够自我实施。运用子博弈精炼均衡策略的一个后果是，某些状态在博弈实际进行的过程中永远不会被观察到。出现这种情况不是因为导致这种状态的博弈路径被某些外部约束条件排除在外，而是因为一旦均衡的“脚本”被采用，参与人的策略性计算使得他们相互避免选择这条博弈路径。由于均衡策略中有一部分规定了在非实际路径上应该采取的策略在实际中不能被观察到，这部分策略可以解释为代表其他参与人所持的理性预期或信念，它们是关于相关参与人在非实际路径上将采取的行动的预期或信念。

这一点可以从格雷夫、米尔格罗姆和温加斯特(1994)提出的商人行会的模型中得到清

楚的说明。该博弈是在中世纪背景下关于交易——商人和城市或交易中心的统治者之间的重复博弈。交易中心为了扩展商业机会，必须确保来访商人的人身和财产的安全。城市统治者也许会发誓为来访商人提供各种安全保证，但是在交易关系建立之后，统治者可能又会说话不算数。现在假定商人们自己组成行会，并采取如下策略：他们在既定时期内在该城市做交易，当且仅当他们当中没有任何一个商人以前被统治者欺骗过；否则，他们将组织罢市(让我们假定实施集体罢市没有问题)。除非行会宣布罢市，否则统治者不会欺骗任何商人；但一旦宣布罢市，统治者将欺骗任何一个来访的商人。这三位作者证明了这样的策略组合构成一个(精炼)均衡。在实际的博弈路径上，欺骗和罢市均不会在正常情况下被观察到。但这不是因为它们事先就被博弈规则排除在外，而是因为统治者意识到，一旦他不信守诺言，行会一定会举行罢市，因此欺骗对他不利。这样一来，行会的形成就成为一种迫使统治者信守诺言从而鼓励交易扩张的机制。在这个例子中，行会(商人的组织)和它在欺骗发生时(均衡策略中非实际路径的部分)组织罢市的预期作用被认为提供了一种可信的执行合同的机制。

基于上述模型和其他理论结果，格雷夫从博弈均衡角度给制度下了一个简明定义：

“在博弈论框架中，两个相互联系的制度要素是(关于别人行为的)预期和组织……组织是非技术因素决定的约束，它们通过引入新的参与人(即该组织本身)，改变参与人所得的信息，或者改变某些行动的报酬来影响行为。”

(Greif 1996)

这里的“组织”指的是如商人行会之类的社会单元，它们构成博弈参与人集合的一部分，受到博弈均衡所衍生的约束制约。在这种特定的意义上，格雷夫的定义结合了第一种观点，即博弈参与人的制度观。

关于制度的起源，我们已经看到，持博弈规则论的经济学家倾向于设计的观点，即规则制定是立法者、政治企业家或从事机制设计的经济学家明确设计的结果。在持博弈均衡论的经济学家当中，关于制度起源问题，在一开始似乎并未达成一致的意见。那些相信进化博弈论的经济学家明显赞成制度是“自发的秩序”(Menger 1883; Hayek 1973)或自组织系统。与此相对照，子博弈精炼均衡概念假定，个体参与人在认识自己决策和别人决策之间的反馈机制方面具备完备的演绎推理能力。然而，个体参与人如何能够联合选择相互一致的策略、促成制度的建立，尤其是在有多重均衡的情况下？子博弈精炼均衡理论根本

无法解释某种制度出现在此地而非别的地方的原因。例如，在上面讨论的中世纪交易的例子中，统治者每期选择欺骗和商人罢市（因此所能观察到的是无交易发生）也构成一个子博弈精炼均衡。这就不难理解，那些运用超理性均衡观点（如子博弈精炼均衡）的经济学家无非想说明，特定的策略组合（即实际的博弈路径和参与人预期）一旦确立，就将是自我实施和可维持的。

在我们支持博弈均衡制度观之前，还有一个理论悖论需要解决。如果制度的作用在于以一定方式制约参与人的决策，那么，这种约束是如何被参与人意识到并被认为是相关的呢？是不是在均衡出现之后？如果是的话，那么，个体参与人在知道均衡之前因而尚不受它的制约的时候，又是如何发现和选择适当的均衡策略的呢？换言之，参与人关于未来情形的信念与参与人基于这些信念做出选择所导致的实际后果之间的一致性究竟是怎样建立的呢？这个问题乍看起来象是一个常规的均衡稳定性问题。但是，本书将说明，其实这个问题比均衡稳定性问题更具根本性，不能简单对待。这是为什么我们要对传统博弈论做出重大修正、提出一种新的制度观的原因所在。我们下一节就来讨论它。

## 1.2 制度的诸方面：共有信念、均衡的概要表征和博弈的内生规则

### 1.2.1 我们的制度观

如上所述，关于制度的定义不涉及谁对谁错的问题，它取决于分析的目的。因为本书的目的在于理解制度安排的多样性以及制度变迁的过程，我们现在就要引入一种适合于分析这些问题的制度定义。在一开始试图用简短一句话说清楚我们的定义及其全部含义是相当困难的，但我们先不妨将制度概括为关于博弈重复进行的主要方式的**共有信念的自我维系系统**。我们可以把“博弈重复进行的方式”等同为博弈规则。不过，不同于博弈规则论的观点，我们不认为规则是外生给定的，或者由政治、文化和元博弈（meta game）决定。我们认为博弈规则是由参与人的策略互动内生的，存在于参与人的意识中，并且是可自我实施的，这就如同均衡博弈论者所认为的那样。制度作为共有信念的自我维系系统，其实质是对博弈均衡的**概要表征（信息浓缩）** [summary representation (compressed information)]，它作为许多可能的表征形式之一起着协调参与人信念的作用。制度也许存在于人们的意会理解之中，也许存在于人们头脑之外的某种符号表征之中。但在任何情况下，某些信念被参与人共同分享和维系，由于具备足够的均衡基础而逐渐演化为制度。

通过强调均衡“信念”，我们的制度观与上面引述的格雷夫的制度观非常接近<sup>11</sup>。不过在目前阶段，我们的定义除了要求均衡是纳什均衡之外，未对均衡概念的内涵做出任何进一步的说明。随后我们会提出一种关于制度的更精确和更广泛的定义，它不仅将包容古典和进化博弈论两种观点<sup>12</sup>，而且还具备两者一些更自由的变种形式。另外需要注意的是，我们强调制度的“概要”和“信息浓缩”的性质。在后面将看到，这一界定对于理解制度演化过程非常有用。总而言之，贯穿全书，我们关注哪种制度在什么条件下才能变得是可行的，以及制度关联的方式。虽然现在就对这种新制度观做一个准确的界定还为时尚早，我们还是要在这一描述一下这项研究背后的基本想法。

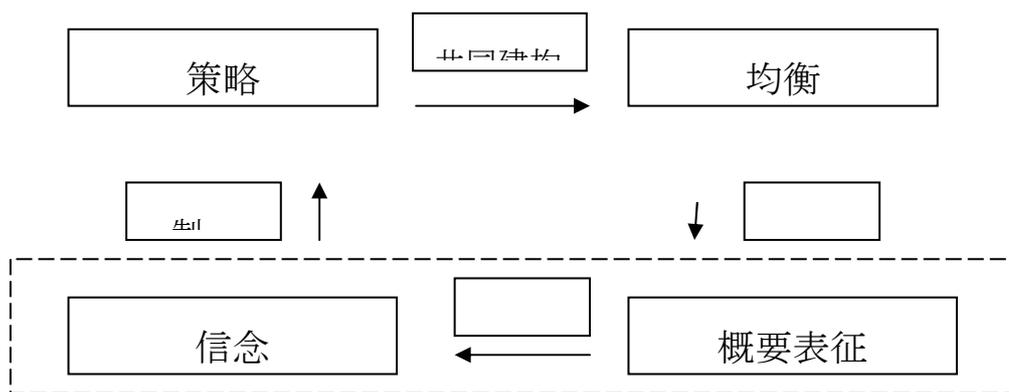
每个参与人设想博弈在一固定集合的参与人当中进行，每个参与人面临一个技术上（和认知上）可行的行动集合。在一定时期，所有参与人选择的行动加在一起——我们称之为行动组合(action profile)——决定了每个参与人的报酬分配。让我们暂时把参与人的行动集合的总和称为博弈的域(domain)，把规定每个行动组合产生报酬分配结果的规则（功能）称为博弈的外生性规则。给定博弈的这些特征，每个参与人都试图使其报酬最大化。不过，最终结果不能由单个人的行动惟一决定。每个参与人最优的行动决策取决于别人的决策。可是，参与人如何能够推想别人在博弈所有可能状态下的决策情况呢？

暂且假定博弈是重复进行的，由此演化出一个稳定结果(行动组合)，每个参与人基于个人经验对博弈进行的方式形成了大致的认识。换句话说，参与人无法、甚至也不需要推断出别人行动决策规则的全部细节，但知道有关参与人在行动决策时可能采用的规则的一些显著特征。参与人依靠这些浓缩信息得出自己在域的各状态下的行动规则——即策略。所有的参与人根据他们对别人行动规则的主观认知(信念)形成自己的行动决策规则，虽然是以一种不完备和浓缩的形式形成的。惟有在关于他人行动规则的浓缩认知稳定下来并不断再生的时候，他们自己的行动规则才能趋于稳定，成为参与博弈的有用指南，反之亦然。当参与人的信念与其行动规则形成一致时，我们称这种状态为纳什均衡。只要参与人对别人行动规则的信念是可维持的，偏离自己形成的行动规则是不合算的。于是，制度对应着（几乎所有）参与人共享的那部分均衡信念，其中信念是关于博弈将实际进行的方式（或者说域中参与人使用行动规则的方式）的预期。虽然制度是一种均衡现象，但不应把它们看作是在一次博弈下不完备演绎推理的结果，也不应视为一种根本不需要归纳推理的完全的静态平衡。它们代表了实际上重复参与博弈的当事人自我维系的基本预期。制度是“由有限理性和具有反思能力的个体构成的社会的长期经验的产物”（Kreps 1990：第183页）。

均衡制度是一种社会建构的现实，因而它必然内生于域之中。它通过其概要表征——

意会 (tracey needham) 的或符号的——协调着参与人的信念。制度作为一种均衡现象，只要别人不忽略它，任何人都不敢不正视它的存在，从而对人们的策略选择构成影响。参与人基于共有信念而做出的策略决策共同决定了均衡的再生，均衡的再生反过来又强化了关于它的概要表征。这样，制度成为自我维系的，浓缩于其中的信念也被参与人视为当然，除非发生了动摇共有信念的事情。(参见图 1.1，虚线方格代表制度，请暂时忽略点线围成的方格)。通过这种方式，制度虽然是内生的，但同时又客观化了<sup>13</sup>。基于均衡分析，我们能够理解制度的双重性质——内生性和客观性，正是这双重性质导致了人们对制度观做出内生规则论和外生规则论这种多少有些让人疑惑的区分<sup>14</sup>。

图 1.1. 作为共有信念和均衡概要表征的制度



一种均衡状态的显著特征可能客观化，凝结为一种制度，因为这一点，制度可能表现为明确的、条文化的以及（或者）符号的形式，如成文法、协议或言示系统界定社会不同角色的社会结构和组织等等。不过，关键点在于，一种“竞争性符号系”只有当人们相信它时才能成为制度<sup>15</sup>。在此意义上，成文法和政府规制如未及有人把它们当作行事就不构成制度。举例来说，政府颁布某项法令禁止进口某些物品，但如果人们相信，贿赂海关官员可以绕开此项法令，那么，与其把这项法令视为制度更为合适。另一方面，一些没有正式化的实践只要参与人认为它们是域内在状态的相关表征，就可以看作是制度；而当参与人对它们的信念动摇了，它们就不再作为制度存在了。

从均衡的概要表征来概括制度有助于澄清制度既制约又协助的双重性质 (dualistic constraining/enabling nature)。制度的作用通常被外生博弈规则论理解为是制约参与人行为的因素（例如前面讨论的诺斯和格雷夫的定义）。的确，制度通过协调人们的信念

控制着参与人的行动决策规则。在多种可能性(即多重均衡)中,共有信念引导着参与人朝某一特定的方向行动。在此意义上,控制或说制约自然是制度化的内生特征。然而,制度只是通过概要和共享的方式来协调参与人的信念。在一个信息不完备和非对称的世界里,作为博弈将如何进行共享和概要的制度的制度,又“协助”理性有限的参与人节约决策所需的信息加工成本(见图 1.1)。<sup>16</sup>

这里借用经济学家熟悉的价格机制作个类比可能是合适的。在市场机制下,个人无需知道决策的内部和外部环境的每一个细节,只需知道相对价格即可(Hayek1945)。撇开产权和合同的执行成本不谈,如果存在一个市场的完备集合,相对价格可以严格地称为“充分统计量”(sufficient statistics),它们概括了社会最有效实现社会最优结果所需的全部信息(消费和转换的边际替代率)。相对价格的维数不超过交易产品的种类减去一,因为总有某一特殊产品用作记账单位(Koopmans 1957; Hurwicz1960, 1973)。当然在实际生活中,市场远不完善。个人还需要其他手段去获得决策所需的信息。在市场之外,各种制度应运而生,以补充市场之不足。简言之,个体参与人不仅受制于制度,同时也被制度所协助。正象市场以相对价格的简略形式传递有关经济环境(技术、偏好和资源禀赋)的信息一样,制度以其简略形式传递着相关信息(参见第 6.2 节)<sup>17</sup>。除此之外,由于信息浓缩的特点,均衡的概要表征在域的环境发生微小变化以及导致的相应的均衡移动下具有刚性<sup>18</sup>。体现在制度当中的浓缩信息使得理性有限的参与人能够有效地收集和利用信息,这些信息是使其行为与变动的内外环境保持一致所必须的。

### 1.2.2 我们的新制度观适合于分析目的五个理由

对制度的概括体现了理论家的偏好,而与正确与否无关。但是,我认为,至少存在以下五个理由让我们相信,从共有信念均衡的概要表征角度看制度适合于比较制度分析。前三个理由适用于一般的博弈均衡制度观,最后两个理由在一定程度上专门适用于均衡的概要表征制度观。前两点理由已经由格雷夫(1994, 1997, 1999)做了富有说服力的表述,这里重述一遍只是为了论述之完整。米尔格罗姆和罗伯茨 (Milgrom 和 Roberts 1990a, 1992)对第三点理由的贡献也是需要在这里指出的<sup>19</sup>。

(i) **对制度起源和实施进行内生性分析**: 博弈均衡制度观可以从内生的角度分析制度的起源和实施问题。如上所述,如果人们接受传统的外生博弈规则论,一个直接的问题

是，制度究竟从何而来，如何产生和如何实施。人们不得不在规则发挥作用的域之外寻找制度的起源，如在经济域之外的政治域，或者作为一种理论可能性，在理性的参与人从诸种可能的规则中集体选择一种规则的元博弈的域中寻求制度起源<sup>20</sup>。可是，政治域的博弈规则又是如何决定的呢？元博弈的参与人又是如何知道所有这些可能的规则呢？他们如何博弈？元博弈的规则又是怎样确立的？这样，无穷尽的循环推理问题又会自然浮现出来。部分解决该问题的办法是，把制度看作是博弈过程的内生稳定的结果，其中的域既可以是经济和社会域，也可以是政治域，并尽量在一开始不界定博弈的非技术性规则。

但有一点警告，必须说明清楚。虽然我们致力于把制度理解为域中“内生的博弈规则”，我们并不希望建立一种模型，使每一项制度都同时是内生的。换言之，构成“外生博弈规则”的因素，即参与人集合，行动决策集合和参与人决策组合与后果的对应规则不可能完全由技术、资源禀赋和参与人的偏好来描述。A. 费尔德(Field 1979, 1981)首先指出了这一点。为了稍微换一个不同的角度来理解其中的含义，让我们假想某种博弈的外生规则完全可以由技术来描述。可即便这这样，也还有重复博弈中多重均衡解的可能性，如前所述，究竟哪一个均衡被选中无法由模型内生决定。我们有必要获知历史事件和过去流行的规则，以及邻近域所流行的规则<sup>21</sup>。这是因为参与人认为可行的特定行动也许受到历史先例的制约，而参与人在某一域行动决策的后果有可能被该域的制度环境(即邻近域流行的内生性规则)决定。从来就不可能存在一种免于制度影响的世界，使得分析能以此为出发点，完全不依赖于任何外生给定和人为设计的规则系统。因此，没有人可以避开无穷循环推理的问题。但我们可以将无穷循环推理引向从过去和历史继承下来的社会结构，而不是引向元博弈的逻辑结构。

外生博弈规则论关于实施问题同样面临无穷循环推理的困境。撇开自我实施的规范和惯例(诺斯意义上的“非正式”规则)，如果博弈规则(即“正式”规则)必须由附加的第三方(实施者)实施，那么，实施者的激励问题又不得不考虑。谁来监督实施者呢？或者说，是不是需要另一个实施者来监督初始的实施者呢？正如赫尔维茨的分析所表明的那样，对该问题的解决又必须诉诸于分析一个包括实施者在内的博弈，将他视为一个博弈参与人，然后在给定其他参与人的策略选择所构成的一个均衡的条件下，给实施者规定的行动规则是否成为他的均衡策略，从而可自我实施，反之亦然。

(ii) “历史是重要的”：在特定博弈模型存在多重均衡解的情况下，博弈均衡制度观可以揭示制度中“人为设计”(North 1990)的一面，而不只是生态、技术或文化所决定的

那些方面。如果对应于博弈结构的技术设定，博弈只存在一个均衡解，那么该均衡就仅仅是技术条件的一种表征，而不是制度<sup>22</sup>。例如，东亚社区规范的演变通常被归结为东亚独特气候和生态条件的产物，它们使小农家庭种植和集体使用灌溉工程更具生产效率。事实上，朝鲜和日本享有相似的生态条件，但在乡村社会结构和社区规范方面发展出各不相同的制度演化路径，对两国随后制度变迁轨迹产生了深远的影响(参见第 2.1 节和第 9 章)。

均衡多重性通常被博弈理论家看作是一个困扰性的结果，他们付出了大量的研究努力，试图“改进”均衡概念，使单一纳什均衡解成为可能，目前尚未获得重大突破。我们认为，博弈均衡的多重性在比较制分析中不应看作是一个理论负担。通过将制度置于均衡分析，我们可以清楚地表明，制度是人力所为，虽然绝非任意设计或随意执行的产物。更进一步说，制度分叉一旦发生，即使两国随后面临同一技术和市场环境，他们的整体性制度安排仍可能会相差甚远，其结果取决于各自制度发展的历史轨迹——这就是著名的路径依赖现象(David 1985)。因此，均衡和历史分析在比较制度分析中是互补的和不可分割的。

考虑到不可能在同一时间将所有的制度现象都视为内生的结果，格雷夫(1998)建议在运用均衡观点分析制度时，采用如下分析程序处理历史信息：(1)以历史和比较知识为基础，首先区分出哪些技术和制度因素可以看作是“外生的”，哪些制度因素需要当作“内生变量”即需要解释的变量；(2)建立对应特定背景的博弈论模型，其中哪些被界定为外生变量的因素定义了博弈的外生规则，然后解出所有的均衡解；(3)考察是否存在一些均衡解有助于解释内生性制度的性质；(4)研究哪些“历史”因素促成了对特定均衡的选择，由此确定历史对制度形成的作用。

(iii) **制度间的相互关联和相互依赖**：均衡制度观提供了一个分析经济中各制度相互依存的可驾驭的理论框架。当政府为了引进一项新“制度”而颁布法令，法令的实施在一定的经济、政治和社会背景下经常产生意想不到的后果。比如在后共产主义国家，政府为模仿发达国家的公司控制权市场，颁布私有化法令，其中的一个意外后果是公司控制权普遍被内部人掠取，如前工业官僚和原国有企业的经理，他们在经济“转型”前已经掌握了实际的控制权<sup>23</sup>。这种情况有些类似于将实验室研制的药品用在人体身上，因人体的复杂性而出现副作用。意外情况出现的一个主要原因是在理想计划和现存制度环境之间缺乏必要的“耦合”，后者反映了制度发展独特的历史轨迹。这说明，只有相互一致和相互支持的制度安排才是富有生命力和可维系的。否则，精心设计的制度很可能高度不稳定。公司治理域的共同决定模式(co-determination)和政治域的社会民主社团主义(social

democratic corporatism)这两种相互支持的制度安排同时出现在德国，而主银行制、终身雇佣制以及产业协会和相关行政官僚部门的密切联盟在日本共同演化，相辅相成，两者均区别于所谓的盎格鲁 - 撒克逊模式，这都不是偶然的。

我们在第 2 编将从制度关联和制度互补的角度对制度的相互依存性进行理论概括。采用均衡制度观使得我们能够将这些直观上非常意味的概念置于严格的理论分析。特别地，我们将考察经济中不同的域的博弈，如组织协调、商品交易、人力和金融资产的劳务交易、政治交换和社会交换等等。然后运用托普吉斯(Topkis 1978)、米尔格罗姆和罗伯茨(1990)发展的分析工具，分析参与人在一个域的均衡策略如何与同一域或不同域其他参与人的均衡决策成为互补品，或者以后者为条件。通过这种方式，我们既可以理解一个国家整体制度安排之所以具有耐久性的条件，同时又可以理解这种制度安排的多重性<sup>24</sup>。

与上述系统观点相比，代理理论(agency theory)作为比较制度分析的工具，其优越性和局限性都同时显现出来了。代理理论将经济某一域内参与人的经济互动视为一种委托 - 代理关系。它试图研究在委托人和代理人信息不对称的条件下，哪一种类型的自我实施(激励兼容)的合同安排能够成为对环境和激励约束的次优反应(a second-best response)。该理论的模型解不仅对技术环境的设定非常敏感，而且对隐藏在委托人和代理人目标函数内的参数设定以及刻划代理人外部机会的参与约束中的“制度环境”非常敏感。在解释和运用委托 - 代理模型的结果时必须注意到这一点。这些结果只是适用于模型隐含假定的特定制度环境，并非在任何域都可以成为由技术惟一决定的次优解。代理理论和我们的博弈论框架的区别有些类似于给定外生的价格参数分析个人决策的“局部”均衡观点，和在瓦尔拉斯经济中分析市场价格决定的“一般”均衡观点之间的区别。在委托人和代理人互动的特定域，当其他域的制度安排被假定为外生环境的时候，代理理论提供了一种强有力的分析制度的工具<sup>25</sup>。但是，为了真正理解某种制度在一国经济的域而非在别国类似的域出现的原因，我们还必须明确探讨每种经济中不同制度跨域的相互依存的机制。

(iv) **不同符号系统竞争而诱致的制度变迁**：均衡的概要表征制度观提出了考察制度变迁的一个新方法。如上所述，制度传递信息的能力不是完备的。但对于理性有限的参与人来说浓缩的信息也许就足够了。在浓缩信息的指引下，参与人开发出与博弈内生规则相一致的技能和性情。可是，当环境发生重大变化时，参与人业已形成的技能和决策习惯可能同时出现问题，不再适应了，从而引发“制度危机”。于是，制度被“视为当然”的部

分，即关于博弈实际进行的方式的共有信念开始受到怀疑，参与者被迫参照现存制度之外的新信息，重新审视他们自己的决策规则。

新制度的出现只有当参与人的决策规则在新的条件下相互一致，其概要表征导致的各人的信念系统相互趋同时才能实现。这种转型并不仅仅是在给定的博弈结构下从一种均衡向另一种均衡的过渡。它可能还包含一种新奇的变化，即从参与者一定的决策规则集合下的均衡向其他(如扩展的)决策规则集合下的均衡过渡(参见第9章)。在过渡过程中，包含新行动选择的各种决策规则被参与者不断尝试并相互竞争。在这种情况下，信念的收敛过程和新决策的协调过程这两者如何同时实现呢？在第7章我们将看到，经济学目前尚无法证明，实际决策和信念的双重收敛通过一种相互作用的合理机制(即实际决策由信念引致，而被观察到的实际决策又导致信念的形成)是可能的，尤其在行动决策包含新奇因素的情况下，证明双重收敛就更为困难。

但是，在过渡过程中，如果某种特定的符号系统从众多竞争性系统中脱颖而出，被公认为是重要和显著的，参与人的新决策策略不得不与之相协调一致，那么，借助该符号系统的指引，双重收敛还是有可能实现的。当参与人的决策趋于均衡时，一种指导性的符号系统将与人们的经验相互一致，并不断被后者核实和确认。于是，这种指导性符号系统成了镶嵌在参与者稳定信念中的均衡的概要表征，即制度(参见图1.1中带点线的方格)。关键点在于，某种符号系统存在于新均衡的演化之前，被相关域所有参与者的经验所接受。它可以是“尚处于形成中的、(有可能)确立行动新策略或新风格的文化或意识形态 - 即明确表述和高度条理化的符号系统”(Swidler 1986: 第278页)，也可以是“激发最终冲破组织能力和环境约束极限的行动的企业家眼光”(藤本 Fujimoto 1999)，甚至可能是革命党的政治纲领(如“一切工厂归工人所有，一切土地归贫农所有!”)。第9章我们将说明理性有限的个体参与者形成自己的“主观博弈模型”的方式，并把制度变迁的机制作为这些主观模型不断修改、完善和最终收敛于共同的表征系统的过程来讨论。

(v) **成文法和公共政策的作用**：最后，究竟把构成制度的博弈规则看作是内生于相关域还是在政治域外生设定，这对解释公共政策的作用具有重要的影响。如果人们接受制度是政体决定的规则并对经济绩效影响重大这一观点，那其隐含的推论将是，绩效不佳的经济应该通过设计和推行更好的规则来改革自身，这些更好的规则可能从别国借鉴而来；但如果这一目标未能实现，则应拿政府是问。这种观点存在两个问题。首先，政府自身是一个组织，它由具有自己动机和抱负的个人构成。在政治域的博弈中，政府是一个内生的参

与人，任何政策制定的结果都应理解为政府、政治家、私人等参与人策略互动决定的<sup>26</sup>。谴责政治家道德和能力低下，这本身无助于问题的真正解决。

其次，如上所述，一项政策若与其他域的现存制度不相耦合，或人才积累不足，均可能产生政治家意料之外的结果。本书将把成文法和政府规制视为定义博弈形式的外生参数（即博弈的外生规则），然后考察在此条件下参与人策略互动的结果<sup>27</sup>。成文法和政府管制也许会引致制度的变迁，但它们自身不是制度。一项制度结果也可能与立法者或政府的初衷不尽相同。以下一些问题特别值得进行仔细和系统的研究：诸如初始的“制度”条件——比如旧制度的遗产、现行的非正式规则（规范、社会伦理等等）、人力资产存量的水平和种类等等因素如何影响公共政策对制度变迁的作用，政治域的规则设定又如何与其他域内生的博弈规则的演变相互作用。<sup>28</sup>

目前讨论的微妙之处在于，内生地看待政府并不必然意味着博弈结果将完全被决定了，不存在任何政策建议的空间——这被称为“决定论悖论”（Bhagwati, Brecher 和 Srinivasan 1984）。前面曾经暗示过，当某域陷入“制度危机”的时候，参与人关于博弈状态缺乏清晰的预期，即便他们自认为有清晰的预期，其预期也可能不相一致。这时将存在一定的余地让竞相争当“聚焦点”（focal point）的外生性符号系统发挥影响力（Schelling 1960）。这些符号系统可以是竞争性政党的施政纲领或职业政治家的建议。因此，在政治域内外的政治言论对制度演化将产生一定影响（Weingast 1997）<sup>29</sup>。至于影响大小，其中一个关键因素在于它们和政治域之外流行的实践“耦合”的程度。

### 1.3 全书的安排

#### 1.3.1 域、博弈形式和制度

我们已经介绍了本书将发展的关于制度分析的基本概念架构。本书的根本目的是将这个概念结构运用于当代制度问题的一些基本方面，并考察其相应的含义。为此我们需要系统地分析制度在不同域之间的相互关联。与此同时，如果我们能够结合经济学、社会学、政治学和认知科学关于制度的重要理论贡献，发展一种内在一致的博弈论分析框架，那将是非常有意义的。在本节，我们想进一步为本书提出的雄心勃勃的目标做一些概念性的准备，然后说明本书的整个安排。

首先让我们对迄今为止所介绍过的基本概念重新做一番严格而准确的表述。如前所

述，我们分析的基本单元是博弈的域。它由参与人集合和每个参与人在随后各个时期所面临的技术上可行的行动集组成，参与人可以是自然人，也可以是组织<sup>30</sup>。记住在域内所有参与人在一个时期所选择的行动被定义为一个行动组合。给定每一期开始外在的环境和历史决定的域状态，该期的一个行动组合将产生在状态空间的一个结果，状态空间定义了在该域与参与人有关的所有可能的物质状态。该期的结果决定了随后一个时期的初始状态。根据一个函数关系(或规则)，对于每一个行动组合和历史给定的初始状态，都有一个状态空间下的结果相对应，该函数(规则)被称为**后果函数**。各种环境因素，如在其他域流行的技术和“制度”以及由政体决定的成文法和公共政策，定义了后果函数的具体参数值。一定域和相关的后果函数则界定了一定的**博弈形式**，它代表了博弈(或机制)的外生性规则。

每个参与人对状态空间下每期所有可能的结果都有一个偏好序。后果函数和参与人的偏好函数结合在一起定义了博弈论通常意义上的参与人的**报酬函数**。我们之所以区分客观性的后果函数和主观性的偏好函数，原因在于界定博弈的外生性规则这一概念，以区别于作为内生性博弈规则的制度。如果域中所有的参与人，给定其自身关于别人策略选择的预期，均选择一定的行动计划，使其当期报酬或当期和未来报酬现值之和最大，那么，这种情形就称为是一种博弈。参与人选择的行动计划就称为他或她的策略。策略可以只是一个特定的行动，也可以是依域状态(或历史状态)而定的全面的行动计划(脚本)。

我们将涉及六种基本的域类型：共用资源、交易(经济交换)、组织、社会交换、政体和一般性组织场。我们将考察何种类型的制度或规范在相关域或跨域是有生命力的。第1.2节曾提到，诺斯把组织排除在制度类型之外。他认为，组织是博弈的参与人而非博弈规则。他这样做的原因是强调组织是政治域制度变迁的主体。相反，本书相对来说不那么强调政治决定的正式规则以及从政治到经济过程的因果性，虽然政治域的博弈结果和经济交易域的博弈结果之间的反馈机制是非常重要的，我们将在后面讨论它们。另一方面，本书强调企业和其他私人机构(如金融中介)的组织惯例作为经济整体制度安排的一个不可分割和内生的一部分。我们主要根据参与人集合和不同参与人决策规则的不同特征区分域的类型。我们尽量鉴于它们的技术特征作出区分，以便于界定在域内和不同域之间内生演化的制度的差异性。但是，正如已经指出的那样，在理论上，试图从一种与制度无关、纯粹属于技术的地方开始我们对于制度的探讨是不可能的。因此，在下述域的分类中，一些一般性制度如所有权、角色预期和权力分配将不可避免地出现或隐含地包含在其中，虽然是以非常原始的方式。

**共用资源域**(Commons Domain)：在该域内的参与者集合是由那些使用共用资源的个人构成的，共用资源是任何参与人都可接近的(或由他们联合生产提供的)资源。这里假定，要排除一些潜在的获益者从共用资源得益在技术上是非常困难的，因而参与者集合假定是固定的。他们对共用资源的使用，可以是同时进行，也可以是按顺序但不必需是联合进行的<sup>3126</sup>。关于共用资源的使用，我们不假定参与者具有一个公共目标或使公共价值内在化，而是假定他们是策略性参与者，每人只感兴趣于如何在预期别人行动的条件下是自己的报酬最大化。由于共用资源的存在是该域的一个主要特征，所有参与人的行动集合基本上是对称的，即每个人的行动决策都包括生产，维护和使用共用资源(如增加或减少共用资源的积累，或以不同的密集程度使用它)。因为参与者集合的固定性，个人的行动决策将通过域导致外在经济或外在不经济(例如，使用共用资源的拥挤现象，维护共用资源时的搭便车行为，或者惠及他人的提供公共物品的个人努力)。一种内生于该域的制度可能是(习俗性)产权规则(见第2章)。

**交易(经济交换)域**：该类型的域是由那些私人拥有可以自由交换或处置的物品的个人构成。虽然他们在初始拥有的物品是千差万别的，但其决策集在性质上是对称的，即包含所有技术上可能的用特定数量的物品或货币交换其他物品的价格，接受或拒绝别人的出价，遵守或违背商定好的交易条件。交易博弈的一个重要特征是所有参与者都有不交易的选择权。该域还可以进一步区分为金融交易域，劳动交易适用域，供应域和产品市场域等等(见第3章)。

**组织域**：在本域中，参与者可以通过联合行动生产产品(收益)并在他们之间进行分配。虽然这些联合行动可能涉及某些共用资源的使用(如品牌、公共信息、组织的基础设施)，但该域还是在两方面不同于共用资源域。首先，参与者可以选择是否参与该类型的博弈。换言之，如果本类型的博弈是重复进行的，参与者有选择退出的权利，或者一个参与人在任何一期结束时有权让其他参与者退出本域，因而参与者集合不是固定的。其次，参与人的行动集合彼此可能存在巨大差异，具体视操作性和管理性劳动分工的情况(例如劳动分为管理、工程和操作性任务)，但与此同时，有一个聚焦性(集中化的)参与者——经理，他的决策对该域所有其他人的决策都有直接或间接的影响<sup>32</sup>。交易域和组织域有一个共同的特征，那就是参与者拥有参与博弈的自由。但是，交易博弈的协调任务可以通过两个交易者之间自愿协议来达到，而在组织博弈中，所有参与者都必须赋予必要的激励，

以吸引他们加入博弈同时适当地协调其行动决策（见第4章）<sup>33</sup>。

**组织场**(Organizational Field)：我们还考察一种称为组织场的一般性的域，它介于交易和组织域之间，但又区别于两者。这是一种相对来说较为原始的域，其中组织的建立是通过域内人们的配对(matching)实现的。组织结构类型高度依赖于参与人在配对之前投资于信息加工的人力资产的类型。参与人可以自由退出该域，但他们的行动决策集合假定是对称的，即选择人力资产的类型，决定是否接受某种配对结果，在配对之前没有对层级安排(hierarchical assignment)的具体规定。我们可以把本域想象为一种理论建构，它对于理解组织规范和人力资产共同演化的内在逻辑意义重大（见第5章）。

**政治域**：本域包含一个独特的中心参与人——政府，它拥有和私人参与者不对称的决策集合。私人参与者可以是公民、利益集团、产业协会、工会、阶级等等，具体视场合而定。政府的行动集合包括单方面将私人参与者的财产转移给自己或其他参与人(税收、补贴和罚金等)，强制动员私人参与者的服务(如服兵役和参加陪审团)，对私人参与者施加的有组织的暴力(如死刑、逮捕)和垄断供应某些公共服务，如司法。私人参与者无法通过选择逃避政府行动的影响，后者具有排他性的管制权力。但私人参与者可以选择是支持还是抵制政府。如果政府行动引起私人参与者的强烈抵制和反抗，其结果对政府将是代价昂贵的(如失去权力)，虽然反抗也许对私人参与者也是有代价的。当政治域出现某中心参与人被认可为中央政府这一稳定结果时，我们将这一显著特征称为**国家**。换言之，我们区分作为博弈参与人的政府和作为博弈稳定结果的国家（见第6章）<sup>34</sup>。

**社会交换域**：这是社会学研究的领域，本书不准备专门探讨它，但它对于理解某些制度如社区规范和同质团队或组织的地位分层现象扮演了一个重要的互补性角色<sup>35</sup>。在该域，那些直接影响接受方参与人报酬(如尊严、赞同、反对、同情、谴责、善意的疏忽等)的非经济物品(如社会符号、语言)被单方面提供，或与“非明确的回报义务”相交换(Blau 1964/1998)，有时甚至伴随着礼物馈赠行为(gift-giving)<sup>36</sup>。当交换以多边的形式发生在相互认识的一群固定的参与人中间时，我们就称之为一个社区(a community)。它可以是乡村社区、商人社区、职业社区等等。这些域连同其他类型的域一起导致了各种各样的社会规范的产生（见第2章）。

下面的矩阵从两个维度表述了博弈的六种域的区别。纵向维度对应着域参与人行动集

合的差异，而横向维度描述了参与人是否有选择退出或被强制驱逐于博弈之外的可能性。六种域的位置根据上面的说明在矩阵图应该是一目了然的。

图 1.2. 博弈域的六种类型

参与人集合 决策集合	可变的 (退出 - 驱逐的可能性)	固定的
	经济交换	共用资源
对称的 (水平的)	组织场	(社区)
	社会交换	
非对称的 (层级的)	组织	政治

如上所述，任何关于域类型的划分不可能是在纯粹技术意义上进行。因此，上述域的分界不免多少有些任意。如果我们从策略互动的参与人中选取某个特定集合，他们因具备多个域的特征，有时不容易恰好归入某个域类型。例如公司企业，毋庸置疑，其最显著的特征组织性的，是涉及操作性和管理性劳动分工的组织。但合同理论经济学家强调企业作为“合同的联结”这一方面(Jensen 和 Meckling 1976)：即理解为作为交易域一种制度安排的方面。在一般意义上，企业的本质也可以概括为组织场与人配对的结果。公司企业同时还具有社区的特征，其中公司成员的社会交换形成一定的公司文化(如 IBM 人)和各种亚组织规范(如车间工作规范)。公司企业也分享共用资源域的一些特征(如使用无形的信息资产)。最后，公司企业被认为镶嵌在类似于政治域内的制度的治理结构之中。在一开始，我们不可能发展一种企业理论，将上述所有的特征均考虑在内。企业不同的各个特征将在本书的不同章节中涉及。我们最根本的目的是在历史的和环境的条件下，试图理解企业的各个方面以不同方式相互适应的逻辑结构。

最后，让我们基于共有信念和均衡的概要表征的观点对制度这一直观的概念重新下一

个定义，关于该定义的严格和重要的发展将留给后面的章节。假定参与人在一个域或跨域策略性地选择行动决策的计划，并且，在该域或跨域内已演化出可持续的稳定结果。那么，给定其他均衡(或更一般地说，另一个均衡序列)存在的情况下，我们定义制度如下：

**制度是关于博弈如何进行的共有信念的一个自我维系系统。制度的本质是对均衡博弈路径显著和固定特征的一种浓缩性表征，该表征被相关域几乎所有参与人所感知，认为是与他们策略决策相关的。这样，制度就以一种自我实施的方式制约着参与人的策略互动，并反过来又被他们在连续变化的环境下的实际决策不断再生产出来。**

上述定义涉及制度的五个特征：**内生性**(隐含在“自我维持”、“自我实施”和“不断再生产”三个词组中)，**信息浓缩**(隐含在“浓缩性表征”一词中)，对于环境连续性变化和微小动荡的**刚性**(“均衡路径显著和固定特征”、“被所有参与者所感知”和“在连续变化的环境下……不断再生产出来”)，与相关域几乎所有参与人相关的**普遍性**(“共享的”“制约着参与人策略互动的方式”“被所有的参与人所感知”)和**多重性**。一项制度可以简要地代表域内不同类型的行动决策规则或所有参与人共同的决策规则，这具体取决于域内参与人的行动集合是否对称。举例来说，在政治域，参与人的行动集合不是对称的，概要表征就由关于政府和公民或利益集团依状态而定的不同行动决策的预期构成(回忆一下商人行会的例子)。类似地，在组织域，概要表征则由组织中处于不同位置的参与人的预期角色(如经理、工头和工人)构成。另一方面，在共用资源和社会交换域，行动决策集合是对称性的，概要表征采取的是规范和自我实施的合同的形式，它们受到有关博弈重复进行的方式的共有信念的支持。

经济可认为是由不同域——共用资源、经济和社会交换、组织和政治——的混合体组成的，其中有些相互重叠和交叉，有些则可能寓于其他域之中，凡此种种，不一而足。对于同一外生性博弈规则(如技术)，多重制度可能共存于每一个域之中。不仅如此，制度因为参与人策略决策的关联性也可能跨域而演化。我们把经济中跨域的共时性制度集合称为**整体性制度安排**(overall institutional arrangement)。其结构毫无疑问是非常复杂和多变的，但尽管如此，它们并非随机组合在一起。本书的主要目的之一就是揭示各种各样的整体性制度安排背后的规律性。

### 1.3.2 本书的安排

本书由三大部分组成。第1编将依次讨论六种域最原初的形式，界定每种域里作为稳定多重均衡现象的制度原型(prototype)。这将为第2编建立一个一般性博弈论框架奠定

基础，该框架将用来定义制度，分析制度跨域和随时间的相互依存性，从而理解制度变迁的机制。然后，我们需要考核该理论框架是否真正能够帮助理解当代经济制度安排的复杂性和多样性以及制度变迁。这是第3编的任务。

第1编从第2章开始，讨论共用资源域以及嵌入其中的社会交换域。从中我们将推导出习俗性产权规则和社区规范，它们作为这些域参与人策略互动的内生结果而出现。第3章是关于交易(交换)域，从中可引出在没有政府干预下治理和促进交易和市场的各种各样的自主性制度类型。第4章集中分析组织域，界定各种富有现实意义的组织和准组织结构，然后研究其相对的信息效率以及治理问题。第5章分析组织场，其中不同的组织结构连同不同的人力资产类型(信息加工能力)共同演化成一定的组织规范，然后讨论如何利用来自组织多样性的收益。第1编的结束——第6章转向政治域，探讨在政府作为博弈参与人的情况下，各种国家形态作为政治交换博弈的稳定结果出现的条件。在第1编所界定和讨论的制度被称为**元制度**(proto-institutions)，原因在于它们均是在单一域得出的，没有借助域之间的相互依存性，故不可避免地表现为原始的形态。这一部分主要依赖于许多学者——包括我自己在内——以前的研究结果，为当前关于制度的分析观点提供一个概览(当然这决不是一个全面而广泛的文献概览)。

建立在第1编准备性的分类研究基础之上，第2编着重于建构关于制度的一般性分析框架，其中伴有一些阐释性的例子和案例。主要目的在于提供一个分析经济整体制度安排的系统特征及其变迁的统一架构。第7章将给出关于制度的一个精确的博弈论定义，我们把制度视为共有信念的自我维持系统和均衡的概要表征，然后是关于制度作用的讨论。第8章系统地研究跨域的博弈关联如何导致新制度形式产生以及(非帕累托有效的)制度安排的多重性，以揭示其中的深层逻辑。在第1编使用的重复博弈或进化博弈模型有助于对单个制度进行严格的分析。但正因为其内在的逻辑力量，这些模型在把握制度变迁的基本特征——新奇性和创新性时表现出一定的局限性。第9章将放弃参与人行动集合客观固定的假设，引入参与人用于主观地认识博弈结构的个体主观博弈模型这个概念。通过研究参与人在面对外部冲击或内在危机时如何相应地在认知上修改其主观博弈模型，我们试图描述制度变迁的一种可能的机制。第10章转向制度变迁的客观机制，并讨论导致路径依赖现象的制度间的历时性关系。

在建立了一个分析制度间历时性和共时性关联的理论框架之后，第3编开始对当代经济中现实且更复杂的制度安排进行比较制度分析。第11章根据组织结构的不同类型，将公司治理机制划分为几种相应的形态，然后讨论它们与其他域的制度可能存在的互补性。

第 12 章将给出关于关系融资的新定义，论证在日益加深的全球金融市场一体化过程中，一些基于意会信息的关系融资类型仍将保持经济活力。第 13 章结合关系融资的经典例子 - 日本主银行制，考察一种制度起源、兴盛和陷入危机的过程。第 14 章将试图理解硅谷模式，讨论在什么条件下和在何种意义上它构成了产品系统创新和治理方面的制度创新。第 15 章作为结束语，结合前面的分析结果，首先归纳整体性制度安排的几种重要模型，然后对全球范围内整体性制度安排在全球一体化加深和通信信息技术大发展的情况下仍然呈现多样性的原因提出一些猜想。

---

1 Greif(1994, 1997b, 1998b), 青木昌彦(1995/2000, 1996), 青木和奥野-藤原(1996), 冈崎(Okazaki)和奥野-藤原(Okuno Fujiwara 1995/1999)。格雷夫把这(CIA)称为历史比较制度分析(HCIA)。

\* 这一词组经作者、译者、编者和一些学者反复推敲，也未能像英文那样简洁达意，作者认为，beliefs既有信念、信仰的意思，又有习以为常的思维定势或观念，还有人们的预期。有学者建议创造一个词“共信”，但考虑到作者关于“beliefs”有一系列深入讨论，最终放弃了这一想法。在日文本中，直接译为“共有预计”；法文本中为“connaissances partagees”(知识共享)。显然都不令人满意。——编者注

2-3 维布伦(Veblen)对制度的定义是“对于所有人而言普遍和稳定的思维习惯”(1909[1961]:第239页)。关于承袭该传统的当代表述，参见霍奇森(Hodgson 1988, 1998)。

4 奥斯特罗姆(Ostrom)也是这种观点的著名倡导者(1990:第51页)。

5 虽然正式规则的实施对博弈过程是外生的，由第三方执行，但博弈的绩效取决于监督和惩罚的有效性。诺斯问道：“当共同的规则集合施加于两个不同的社会，其结果会怎样呢？结果会不一样，虽然规则相同，但是，实施机制、实施方式、行为规范以及行为者的主观模式都不一样。于是，真正的激励结构和被认知的政策后果也就不同。因此，同一的相对价格的基本变动或者施加同一的规则将对制度安排各异的社会产生迥然不同的影响。”(North 1990:第101页)。诺斯为此提出了一个更为广泛的“制度框架”的概念，它包括“法律规则、组织形态、实施机制和行为规范”(North 1991:第33页，也请参见第58, 101页)。在其它场合，他有时用“制度架构”(institutional scaffolding)表达同一意思(North 1998)。

6 后果函数的值域是物质结果空间而非效用报酬。赫尔维茨认为，参与人的效用函数不是博弈规则的一部分。因此，他使用“博弈形式”一词而不是“博弈”来指代参与人集合、决策空间及后果函数。

7 严格地说，这种参数表示对于界定限价制度太有限和狭隘了一些，因为政府拥有改变限制价格实际水平的自由。为此，赫尔维茨允许参数值在一定范围内变动，把制度定义为一组机制(如一组限制价格)而非单个机制(如某一特定限价)。

8 肖特把社会制度定义为“一种被社会所有成员认同的社会行为的规律性，它规定了在一些特定和经常出现的情况下的行为，它要么是自我监督，要么由某种外部权威监督。”(1981:第11页)。根据这种定义，如第二种观点那样，制度被认为是行为规则。但这些规则被假定是内生于经济过程，作为博弈的结果而产生的，而不是由政治或立法过程外生制定的。

9 霍奇森给出了一个关于制度经济学进化论的一般性概览(1993)，但他未明确使用博弈论框架。

---

10 “脚本”的说法也被社会学家尚克(Schank)和埃伯森(Abersson)(1977)独立提出,他们用它们来描述被特定角色或情景所召唤的行为模式和程序,自他们以后,该词在论制度的社会学文献中被广泛使用。

11 关于格雷夫最近的定义和阐述,参见格雷夫(1998b)。他的分析与我们的分析有一个细微的差别,就是本书明确地分析了制度化过程中的认知机制及其对制度演进的意义(见第 5.2、7.2 和第 9.2 节)。

12 在第 7 章我们将说明,博弈均衡的各种概念,尤其是古典的和进化的,其实并不相互排斥,而是互补的。另外,均衡概念应依域特征的设定水平而变化。

13 这种观点和伯格(Berger)和卢克曼(Luckman)(1966)所发展的知识社会学的关于制度的现象论解释是一致的。根据他们的观点,“制度化发生在各类行为者习惯化行动相互典型化的场合。”(1966:第 54 页)。但是,“这种制度世界又被作为客观现实来经历。”(同上:第 60 页)。

14 斯科特对社会科学关于制度问题的研究方法作了一个比较研究(1995),他认为,虽然不同的社会科学领域和流派分别研究了制度的规制性方面、规范性方面以及认知性方面,但是,这些方面是总和地体现在一切制度上的。我们的计划也是把这三个方面合并起来思考:一个制度是规制性的,因为它限制了个体参与人的选择;它是规范性的,因为通过共有信念限定了所有参与人行为选择;它是认知性的,因为它是共有信念的建构。

15 因此,《社会现实的建构》(The Construction of Social Reality)一书的作者希尔勒(Searle)指出:“所有的制度事实在本体论上都是主观的,虽然一般而言它们在认识论意义上是客观的。”(1995:第 63 页)。但是,值得注意的是,制度的现象论观点的创始人伯格和卢克曼在《社会现实的建构》一书提到“[制度世界]并未从产生它的人类活动中获得一个本体论的地位”(1966:第 60 页)。我们均衡导向的制度观承认人类活动的结果和信念都作为制度而相互建构(参见图 1.1)。

16 这可以看作是经济学家考察制度的认知性方面的方法,注 14 中斯科特(1995)提到了相关问题。

17 阿罗指出:“预期本身可以看作是个人心理的一部分,但在实际的社会制度中发挥着导引和形成各种预期的主要作用。……存在别人不会利用每一个可能获取短期利润的机会的共识,复杂的金融服务网络将提供预测并为人们解决暂时的资金困难。”(Arrow 1997:第 6 页)。

18 随后我会把耐久性作为制度的一个要求提出来。在这个意义上,价格向量作为偏好和技术的概要表征并不算做制度,因为它们随这些参数变动而变动。因此,类比到此结束。但是关于人们可以凭协定的价格成交并按协定交货的信念可算是制度。

19 接下去的论点是从我自己的著作中发展而来的(见青木 1994a, 1995/2000, 1996)。格雷夫也在自己独创研究的基础上发展了同样的观点(Grief 1989, 1994, 1998b),并且以简洁的形式将其表述出来(Grief 1997, 1998a)。下一句论断清楚显示了他对我的思想的影响。同时,米尔格罗姆和罗伯特对第三点作了贡献,我对它们非常感谢,这是必须注明的。

20 关于后一种观点,参见雷特和休(Reiter 和 Hugh, 1981)和赫尔维茨(1996)。扬曾从均衡博弈论的角度采取了这种观点(1998:第 9 章)。

21 巴苏(1998)为说明这一点提供了一个有趣的比喻。让我们比较两个“全球”博弈:其中一个——我们称之为“法制博弈”——有法律和政府官员,而另一个称为“无政府博弈”——两者皆无,因此是无制度的。但假定在参与人的策略集合和报酬及其它方面,两种博弈不存在什么差别,在第一种博弈中被称为“政府官员”的参与人可行的任何策略在无政府博弈中同一种参与人(无具体称呼)也同样可以选择。又假定我们把前者的均衡称为“法制均衡”,后者称为“无政府均衡”。这样看来法律规则就是均衡结果。然而,仍然有一个什么因素决定前者选择了“法制均衡”,而不是“无政府均衡”的问题。一定存在某些超技术的方面影响着均

---

衡的选择。

22 萨格登强调均衡多重性应作为惯例的一个必要条件。他论证说，“一种自我实施的规则可以看作是一种惯例，当且仅当我们能够想象还有其它不同的规则存在，它们一旦确立也能自我实施” (Sugden 1986 : 第 32 页)。

23 这一点将在第 10.3.2 节的实例 B1 中有详细的讨论。

24 关于米尔格罗姆和罗伯茨对于比较制度分析的贡献，参见米尔格罗姆和罗伯茨(1990b, 1992, 1994) 以及霍姆斯特朗(Holmstrom)和米尔格罗姆(1994)。也请参见帕加诺 (1993)、帕加诺和罗索恩关于早期制度比较分析的两个文献 (Pagano 1993, Pagano 和 Rowthorn 1994)。

25 交易经济学把交易视为分析的基本单位，以区别于代理理论把个体参与人视为基本单位 (参见 Williamson 1996, 第 7 章关于这两种观点的比较)。在这个意义上，交易经济学和我们把博弈域视为分析单位是相似的。交易经济学的权威威廉姆森明确指出了他所研究的分析单位的环境的制度性质——“这些制度的定义，如诺斯、肖特及其他人主要停留在制度环境的层次上，即所谓的博弈规则。制度经济学研究的第二个即更微观的层次是治理结构制度(市场、组织混合体、层级和政府管理部门)”(Williamson 第 4 至 5 页)。还请参见威廉姆森(2000)。

26 巴格瓦蒂等(Bhagwati 1984), 巴苏(1997), 卡沃特(Calvert 1995)和温加斯特(1995, 1997)。

27 格雷夫 (1998a, b); 巴苏 (1997b, 1998), 和科恩豪瑟 (Korhauser, 1991a, b)。

28 格雷夫 (1998a, b)。

29 诺斯 (1990, 1995) 和巴苏 (1997a)。

30 在第 9 章，我们将讨论参与人的主观认知定义其行动决策的启用集合(activated set of action choices) 的情形。

31 这里使用的共用资源概念接近于奥斯特罗姆使用的“公用资源”(common-pool resources)。但她把后者和公共产品(由政府机构提供)，即必须“联合使用”的服务区分开来(Ostrom 1990 : 第 30-33 页)。只要相关域的参与人的行动集合是对称的，我们的概念并不必然把集体物品和集体次品 (collective goods/bads) 如环境污染从(负面的)共用资源中排除出去。如果共用资源由政府机构单方面提供，我们就称之为公共产品，将它放在下面涉及的政治域内讨论，象奥斯特罗姆所做的那样。

32 这里我们主要关心组织的协调方面，当然必须指出，组织还有其他的方面。从这个角度看，即使是平等主义的或集体主义的组织就治理结构而言也是有聚焦点(管理层)的。

33 这被巴纳德(Barnard 1938)论组织的经典著作视为组织的决定性特征。

34 通过将政府视为博弈中自主的参与人，我们的观点和一些政治学家如斯科波 (Skocpol) (1979, 1985)、克拉斯内(Krasner 1978, 1984)和温加斯特(1997)研究的思路是一致的。但这些学者均把政府(主权政府)当作“国家”而不是政治的均衡状态。关于政治学各种相关概念和观点，参见 Krasner (1984)。

35 这一类域一直是各种学派的社会学家制度研究的对象，如社会交换学派如布劳 (Blau 1964/1998), 赫曼斯(Homans 1961); 现象学派如伯格和卢克曼(Berger, Luckmann 1966); 新制度主义学派如梅耶和罗文(Meyer 和 Rowan 1977), 扎克(Zucker 1977), 梅耶和斯科特 (1983), 蒂马基奥和鲍威尔(Di Maggio 和 Powell 1983); 社会嵌入学派格拉诺威特 (Granovetter 1985) 和理性选择学派如科尔曼(Coleman 1990)。

36 在符号和语言交流影响参与人报酬这个意义上说，他们与博弈论中所谓的“廉价交谈”(cheap talk)是相区别的。关于廉价交谈理论的通俗性概览，请参见法勒和拉宾(Farrell 和 Rabin 1996)。