



上海市学术著作出版基金

陈常荣 著

互惠的美德

博弈、演化与实践理性



上海人民出版社



上海市学术著作出版基金

陈常荣 著

互惠的美德

博弈、演化与实践理性

 上海人民出版

图书在版编目(CIP)数据

互惠的美德:博弈、演化与实践理性/陈常榮著.

—上海:上海人民出版社,2017

ISBN 978-7-208-14682-2

I. ①互… II. ①陈… III. ①哲学—研究 IV. ①B

中国版本图书馆 CIP 数据核字(2017)第 168979 号

责任编辑 毛衍沁

互惠的美德

——博弈、演化与实践理性

陈常榮 著

世纪出版集团

上海人民出版社出版

(200001 上海福建中路 193 号 www.ewen.co)

世纪出版集团发行中心发行 上海商务联西印刷有限公司印刷

开本 635×965 1/16 印张 22 插页 5 字数 315,000

2017 年 8 月第 1 版 2017 年 8 月第 1 次印刷

ISBN 978-7-208-14682-2/B·1287

定价 68.00 元

互惠的美德
博弈、演化与实践理性



力不若牛，走不若马，而牛马为用，何也？曰：人能群，彼不能群也。

——《荀子·王制》

那些了解狝狝的人，可以对形而上学作出比约翰·洛克还大的贡献。

——查尔斯·达尔文

前 言

作为一部哲学著作，本书借助博弈和演化的视角探讨“行动理由”这个实践理性的核心问题，从交互行动亦即“互动”(interaction)切入对行动的考察，以纳什均衡、合作解以及演化稳定策略(ESS)切入对理性的考察，并在演化与建构两相对照的框架下，反思社会秩序和个人美德的哲学基础。从博弈和演化的角度看，人类交互行动的内部理由是社会习俗规范制约下的个体动机，而社会习俗规范又是自然—文化共生演化(nature-culture co-evolution)的自发式渐进过程，为此我们将“互惠性”(reciprocity，亦译为相互性)概念当作博弈、演化与实践理性三者之间的交汇点，作为人类社会契约和道德规范的一个核心概念，而“理性的建构”对于人类行为和社会规范的范导和制约作用，都应该被置于这个背景下来理解。

人类诸多个体(以及由此派生出的诸多利益共同体)之间纵然纷争不已，但不影响我们本质上仍然是共在(co-exist)、共生(symbiosis)的存在者。这种共在共生关系的实际发生场域，就是所谓的人类社会。然而从博弈和演化的视角看，对人类社会的分门别类的科学研究(如经济学、社会学、人类学)稍嫌支离破碎，而哲学研究——典型的是柏拉图的“哲学王”——则过多遮蔽于个人主义的流行话语或国家至上的意识形态，为此我们建议，首先从博弈和演化视角重新理解社会，始终不离开人类之为社会动物这一背景来考虑行为的实践理性议题，并且在这个框架之内为我们进一步思考伦理和政治问题另辟蹊径。以族群或社群为基本单位的社会共同体，应当成为不管是对人类行为的实证描述还是规范研究的一个

现实立足点,为此我们不再满足于对实践理性的玄奥思辨,转而强调概念考察与经验反省之间的相互补充和反思均衡(reflective equilibrium);一方面,实践理性的哲学研究很大程度上乃是概念分析的工作,实证和实验并非哲学之所长;另一方面,实践理性不应成为不食人间烟火的纯粹思辨体系,关于社会实际运行的“健全的常识感”(罗素语)不应成为理论上追求大体大思精的牺牲品。在不同层面上,包括大卫·休谟(David Hume)和亚当·斯密(Adam Smith),都是我们的精神先驱。

本书聚焦于“互惠性”概念,探讨了自利论的弱互惠版本、利他主义的强互惠版本、道德情感的一致性版本以及实践理性的相互性版本,分析作为人类美德之源的相互性如何坐落于博弈、演化和理性之中并且成为沟通三者的重要枢纽,同时兼顾其科学(博弈论、道德心理学、演化人类学等)与哲学视域。非合作博弈论中的合理性(rationality)、相关性(correlation)或互动性(interaction)、各种均衡解(譬如纳什均衡解),合作博弈中不同层次、不同类型的“合作”解,演化博弈论对经典博弈论中的“理性人假设”的克服,以及演化博弈中的“演化稳定策略”,还有不同层次、不同类型的演化视角,所有这些概念和方法,都可以用来反思和修正传统的实践理性概念。譬如,只要将博弈和演化中常见的个体理性(individual reason)、交互理性(mutual rationality)、模块理性(modular rationality)、合作理性(co-operation reason)等种种理性形态考虑进来,那么诸如实践推理(practical reasoning)、普遍理性(universal reason)和公共理性(public reason)这些关于实践理性的常规语汇,都不得不面临重新审视。此外,通过引入合作博弈论中不同层次、不同类型的“合作”,以及当代人类学等社会科学中不同版本的利他主义理论,进一步可以帮助我们丰富和反思以合作、利他为旨归的道德哲学研究。从实践哲学的思想资源上看,我们在休谟的情感主义伦理学和“反契约论”政治哲学、尼采的道德谱系学,尤其是在经过重新阐释的亚当·斯密的道德情操论中,找到了博弈、演化与实践理性这三者之间的契合点。

我们的实践理性研究仍然在某种意义上延续了康德(Immanuel Kant)的“批判”或“划界”精神——尽管我们的结论不是康德主义的,因

为基于博弈和演化视角的经验研究,对于哲学上关于实践理性的规范性研究而言,不仅是一种可能的补充,更是一种必要的限制。譬如,在个人主义的流行话语和国家至上的意识形态之间,必须为人类社会的独特话语留有一席之地,这意味着我们不仅能像大多数哲学家和伦理学家那样,从个人或者国家出发进行理性建构,更能够从社会出发,重新理解人类的互动,以及一个人如何自处、人与人如何相处以及全人类如何共处,对契约论和建构主义传统的政治哲学提供某些反思维度。当理性的建构被视为一项文化事业时,实际上它也参与到自然—文化的共生演化之中,对人类社会产生了不可逆转的根本影响。所以人类理性能力对于社会发展的作用既有消极一面,也有积极一面,关键在于如何利用它,限制它。借用赫伯特·金迪斯(Herbert Gintis)的话说,理性的边界不是非理性的,而是社会的。

最后,对实践理性的考察对于博弈论研究并非毫无助益。譬如,从实践理性角度看,博弈论回答三个根本问题:(1)理性博弈(互动决策)如何可能?(2)合作(博弈)如何可能?(3)公正(博弈)如何可能?这三个问题之间是一个递进的关系。不过仍有必要澄清,这是一本哲学著作,不是一本博弈论专著,它从博弈、演化及其他学科领域借来的更多的是方法、观点和旨趣,而对于那些翔实有力的调查数据或缜密周全的数学模型,在不得不涉及它们时都做了“极简化”处理,本书只为广义经验主义背景下的实践理性议题探寻某些思路,并不奢望在非哲学领域有所创见。

目 录

前言	1
导论	1
一、“个人—国家”夹缝中的社会话语	4
二、囚徒困境与社会困境：社会秩序的博弈演化及理性建构	10
三、实践理性与自然主义	14
四、互惠性的四种版本	22
五、要素分析与溯源分析	24
第一章 博弈论的理性基础	31
第一节 博弈人：理性人、知性人和目的人	31
第二节 博弈论与实践理性	43
第三节 从模块理性到交互理性	60
第二章 博弈论与行动解释	80
第一节 行动解释与实践理性	81
第二节 博弈与因果性	93
第三节 理性与自利	105
第四节 理性与自然	118
第三章 博弈的合作变换与对称变换	131
第一节 合作博弈与反合作博弈	131

第二节 合作博弈中的理性与演化	145
第三节 非对称博弈与对称变换	153
第四章 演化、规范性与社会秩序	167
第一节 社会演化的休谟主义纲领	167
第二节 从语言演化到社会契约	182
第三节 演化均衡与社会契约	192
第四节 演化的结果	207
第五节 尼采与道德的演化	216
第五章 博弈演化的独特话语	232
第一节 自然—文化共生演化及其理性因素	232
第二节 均衡、规范与实践理性	240
第三节 从“看不见的手”到“看不见的握手”	248
第四节 从“常胜之道”到“常不败之道”	274
第六章 演化与建构	285
第一节 理性的建构	286
第二节 演化与建构相互作用	293
第三节 公平、正义与道德	302
第四节 现代性与理性的建构	319
参考文献	332
后记	340

导 论

如果说,20 世纪是一个学科专业化的时代,那么 21 世纪也许将成为一个跨学科综合的时代。当不同的学科聚焦于一个共同的议题时,各自的研究结论——不管它们是从人类学家的田野调查,还是从经济学家的数学建模,抑或从哲学家的思想实验得来的——将会达成一致、互相支持。本书尝试从合作/非合作博弈(cooperation/non-cooperation game)以及自然—文化共生演化的跨学科视角,重新审视西方哲学中的实践理性议题,以期在对人类——人类学意义上的“智人”(homo sapiens)——行为的哲学分析与行为科学分析之间寻求某种融贯性解释。诚如赫伯特·金迪斯《理性的边界》一书结论所言,博弈论(主要涉及非合作博弈、合作博弈和演化博弈)是人类行为建模中不可或缺的工具,抛弃或排斥博弈论的行为科学在理论上是残缺的;反过来,博弈论也需要广泛的社会理论来获得解释力。^[1]进一步,我们认为,就实践理性事关对人类行为的理由诉诸和理性建构而言,这个结论同样适用于博弈论与哲学的关系。

作为对决策论(decision making theory)的一种拓展,博弈论(game theory)是对不少于两个理性且知性的行为决策者之间的相互冲突或/和彼此合作的数学模型研究。^[2]作为运用数学的一个分支,它被用来解释有理性、有知性、有目的的人类行为,因此很自然地被视为哲学(尤其是伦理学、政治哲学这样的“实践哲学”分支)和社会科学的一种方法论,在经济学、管理学、政治学和生物学等广泛领域具有重要的运用价值。这里的“game”,也就是“游戏”,泛指一切人类活动(尤指便于数学建模的经

济、政治等社会活动),约略相当于德语中的“*Spiel*”(活动)。

博弈论首先在经济学中得到广泛运用后,生物学家从中得到启示,尝试运用博弈论中的策略互动思想,建构各种生物竞争演化模型,包括动物竞争、性别分配以及植物成长发展等。接着,生物学家根据生物演化的自身规律,对经典博弈论进行改造,比如将经典博弈论中的收益函数(payoff function)转换为生物适应度函数(fitness function),引入突变机制将纳什均衡拓展为演化稳定策略(evolutionary stability strategy),以及引入选择机制的复制者动态(replicator dynamics)模型。^[3]

经典博弈论的一个预设类似于经典牛顿线性力学,认为时间是可逆的,即无论在什么时候,只要给定的条件相同就会得出相同结果。理性博弈人具有无限的计算能力,在给定信息下,能够迅速计算出最优反应动态,这种最优化结果只与外部条件相关,而与时间无关,最多在局中人的策略互动过程中加入折现因子(discount factor)考量。^[4]演化博弈论基于生物—文化演化理论,降低了博弈人的理性要求,取而代之地加入了时间和经验的考量。从时间上看,博弈不再被视为一步到位的“静态”过程,而是一个渐进—生成的过程,演化稳定的实现需要更长的时间,然而一旦达到就具有更强的内生稳定性,除非遭到外生不稳定要素的强大冲击。从经验上看,博弈人通过试错、模仿和学习,将自己的决策过程与其他的博弈人以及整个博弈的结构、演化紧密地联系起来,而不再将个人决策看作给定信息下的个人选择,从而将博弈的主体从“个体”(individuals)的维度推进到“种群”(populations)的维度。在《经济变迁的演化论》一书中,里查德·纳尔逊(Richard Nelson)和悉尼·温特(Sidney Winter)以经济现象为例,指出人类行为演化的惯例性特征,而不单纯是“理性选择”或“市场自然选择”的结果。这种惯例性行为甚至被认为是“基因在经济学中的对应物”,因为演化是沿着惯例而非理性的方向前进。^[5]

如果经典的非合作博弈论及合作博弈论通过用“多线互动”取代传统的“单线突进”思维,对行为科学中的理由解释和预测领域实现了一次重要的方法论革命,那么,基于生物—社会—文化演化模型的演化博弈论,则通过用“生成演化”的因果解释模型取代了理性人假设下的“理性互动”

解释模型,完成了行为科学中的“二次革命”。生物—社会—文化演化包括生物演化、社会演化和文化演化三个密切相联的层面。生物演化的主要理论贡献是查尔斯·达尔文提供的,法国生物学家拉马克等人作了重要的推进。^[6]前者强调自然选择,变异在某种程度上是被动的,而后者通过提出“获得性遗传理论”强调变异的主动性,使得演化论解释更加适用于拥有主观能动性的人类。

奥地利学派的代表人物哈耶克提出了社会秩序自发演化的原理。哈耶克将社会秩序看作是社会成员相互作用的一种状态,依赖某种规则得以形成和延续。规则本身是自发演化——哈耶克称之为文化演化——的选择机制的产物,这种选择机制全然不同于个体理性的选择,而是一种试错、模仿和学习的过程,这些规则作为一种“共同知识”(common knowledge)弥补了理性的有限性。在文化演化中,规则既是选择的,也是被选择的。成功的规则将取代不成功的规则。^[7]

当代以道金斯为代表的新达尔文主义者通过将基因视为“复制者”(replica),把生物体比为“运载工具”(vehicle),抹杀了哈耶克对自发秩序(spontaneous order)与人为秩序(artificial order)的重要区分,因为从达尔文主义观点严格来看,即便是遵循人为秩序的人类行为也是自发秩序的组成部分。然而,对单纯的生物组织过程与复杂的人类社会组织过程进行区分仍然是有必要的。首先,人为秩序和理性建构是一种现实的存在,难以将它们完全纳入狭义的生物演化或自发秩序的解释框架中。其次,与哈耶克有所不同的是,我们主张人为秩序和理性建构不仅是可能的(现实的),还是必要的。譬如,理性建构并不总是多余的。人类是拥有能动性、价值观和规范性的动物,其特有的实践理性能力势必对其行为进行规范、调节。人类的行为不再是单纯的自发行为,就实践理性的谋划而言,它还是自觉行为。

博弈和演化的思维在哲学中的运用,为当代实践哲学注入了鲜活的血液。许多博弈论专家、经济学家、人类学家乃至生物学家所做的工作,实际上有助于当代哲学的多元拓展和纵深推进;他们试图探究以利他主义和普遍合作为前提的人类美德起源于何处,或者社会契约又是如何可

能的。这些人通常自称休谟主义者,对实践理性秉持一种得到精致改进的休谟主义进路:对康德式“原则先行”的实践理性版本总体上持怀疑主义态度,在继承休谟式实践理性版本中的自然主义基本精神的同时,剔除了其中的心理主义要素。^[8]

一、“个人—国家”夹缝中的社会话语

一种稍带激进的观点认为,个人观念在很大程度上乃是现代个人主义意识形态的产物。^[9]根据奥克肖特(Michael Oakeshott)的分析,自笛卡尔以来,在过去的几百年里,个人主义被抬高为一种伦理学、认识论,甚至是形而上学理论,据此还出现一种对政府职能的相应理解,改变了政治模式及种种制度,它在艺术、宗教、工业、贸易和一切人类关系上确立了自身。此外,几乎所有近代关于道德行为的著作,都以假设选择并追求其自身的活动方向的个人开头。^[10]个人被赋予独立自足的人格和尊严,有充分的实践理性能力,也享有自由、独立和平等的政治权利,这些构成了个人主义政治建构的一幅耳熟能详的理论图景。将这种个人主义发展到极端,国家和政府乃是为保障个人权利服务的,其中最重要的除了自由和平等,还有私人财产神圣不可侵犯的权利。在契约论者看来,个人为了保障其权利而订约,建立政治联盟——国家;离开了个人的权利诉求,国家和政府的存在便丧失了合法性。当现代自由主义者们探讨一个正义的国家如何可能时,实际上他们关心的是国家如何才能满足个人的权利诉求,比如通过树立一个至高无上的政府,消除家庭、教会、社群乃至整个传统对个人的公共压力。

在这幅图景下,社会也是由个人得到定义的。考古学界通常认为,距今约 4.5 万年以前,文化意义上的现代人类(culturally modern human)已经出现在非洲大陆了,这些人拥有与我们一样的大脑,一样灵活而先进的文化,由他们所构成的狩猎—采集者群体可以被视为人类社会的第一次成形。^[11]现代的个人主义者宣称,个人在先,社会在后,而社会不过是个人为了满足其欲望—利益诉求和实现其自身目的一价值的而形成的一

个有效的合作系统。然而,正如个人主义的批评者所注意到的,个人身份也是社会—文化演化的自然结果。具有血缘或姻缘关系的个人组成了家庭,而家庭则是社会的细胞。一个人不管在家庭内部,还是离开家庭进入社群和文化共同体的更大空间,都无时无刻不受到社会的浸染或塑造。个人即便费尽心思也难以改变他所身处的社会,而社会在轻易之间就“选拔”或“淘汰”了它需要或不需要的个人。在经济学家赫伯特·西蒙(Herbert Simon)看来,巨大的利益来自更具“驯顺性”(docility)的个体,因为这些个体会自动向其他受过文化教育的成员学习、复制他人有益的行为,从而不必付出昂贵的试错代价。^[12]通常我们也称这种人具有很强的社会性,因为他们能够更好地适应一个社会,付出的成本更小,但收益更大;而那些不愿意循规蹈矩的人,则要为此而付出足够大的代价。

国家至上的意识形态则反其道而行之,认为个人的政治身份只有基于国家才能得到合理辩护,离开了国家这个“大写的个人”^[13],那些小写的个人的政治诉求必定陷入缥缈的虚空之中。国家的统一性不仅确保了个人的生活意义,还为社会的稳定和可持续提供了最坚实的政治支撑。人们先假定存在一个统一的政治目标,所有人都应当为之而奋斗,而国家是实现这个统一目标的唯一合法组织者;脱离国家属性的所谓个人权利,是无源之水无根之木。虽说国家公权力容易沦为中饱私囊的工具,政府的无能和腐败很难得到根除,但这些不是集权专制国家的“特权”,因为崇尚个人自由平等的民主国家也常常无法避免自己系统性的低效无能和不可救药的平庸堕落。

然而,社会既不想迁就个体向国家索要的权利,也不愿屈从于国家对个人的塑造要求,因为社会本身的独特品格,既不依赖个人主义的理性建构,也不仰仗国家至上的意识形态,而是人类自然而然的演化结果。“社会并不是由理性的人类创造的,它是随着我们的本性一起进化的结果。和我们的身体是基因的产物一样,社会也是我们基因的产物。”^[14]社会不仅涉及基因—文化的共生演化,还涉及文化—制度的共生演化。譬如,我们的原始祖先生活在一个等级森严的社会当中,而不是生活在一个奉行平等主义的社会当中。等级森严的社会满足了一个社会的稳定性和可

持续性要求,尽管它并不符合现代政治中的平等主义规范,但并不影响它作为一个成熟的社会。这个社会中可能会有足够多的搭便车者(free-rider)^[15],也可能有足够多的地痞恶霸,但他们并不足以将社会得以维系所必需的那种最低限度的普遍合作和利他主义侵蚀殆尽。

国家至上对社会的规制是有目共睹的,但个人主义也不是省油的灯。在个人与国家之间,被各种各样的权力话语充斥着,基本上没有为社会留下多少活动的空间。譬如,作为一种颇有影响力的个人主义版本,约翰·罗尔斯表面上容许了整全性学说(comprehensive doctrine)的自由活动权利,也允许公民拥有自己的个人文化身份,但是这种容许不可能脱离政治自由主义的建构前提。规范性学说似乎是社会哲学的唯一范式,理性建构则是彼此冲突的各派理论之间心照不宣的方法论默契。几乎没有多少人愿意拨开个人主义的云雾或者国家至上的“铁幕”,还社会一个它原本拥有的独特品格。除了规范性学说,我们可以有描述性的社会哲学;除了理性建构的进路,我们还可以有博弈—演化的进路;除了从个人到国家或者从国家到个人的单一线性思维,我们还可以有社会维度的复合互动性思维,前者追求理由的力量和权力结构的稳固,而后者青睐因果的力量以及动态均衡的生成演化。一种秉持独特品格的社会哲学,不同于任何一种政治哲学,尽管它现实上不可避免地被烙上了这种那种的政治烙印,但并不影响它以去政治化为其追求目标。政治化是社会发展到一定程度的必然产物,但社会的生成演化并不以政治化为前提,对这种生成演化的探索也可以避免作出政治话语的方法论承诺。

对于什么叫“社会”,可谓众说纷纭,莫衷一是。譬如,德国社会学家斐迪南·滕尼斯(Ferdinand Tönnies)将社会理解为一种介于共同体与协会之间的东西:“人们说语言的、习俗的、信仰的共同体;但是(却说)职业的、旅行的、学术的协会。……在一般的意义上,也可以说有一个包括整个人类的共同体,像教会所希望的那样。然而,人类的社会被理解为相互独立的个人的一种纯粹的共存。”^[16]在他看来,共同体是持久的和真正的共同生活,社会只不过是一种暂时的和表面的共同生活。对社会的这种理解类似于我们今天对团体、协会、社区的理解,比如在现代社会里,

公民们分别属于不同的社区、教会、学校、俱乐部和行业协会,这些都是人们基于不同的兴趣、利益、目标和理念而建立起来的或多或少带有临时性特征的组织,在某些政治制度下,这些组织在不同程度上受到了国家和理论的双重制约,构成我们今天所说的“社会建构”的讨论对象。

相比之下,我们倾向于赞同哈耶克对社会的“自发秩序”式理解。他曾经写过一篇专文,题目就叫《何谓“社会的”,它是什么意思》。据他考证,“society”这个概念,其原初的学术含义是用来描述一种自发产生的人际关系秩序,以区别于特意设立的国家组织。当我们谈论“社会结构”——譬如语言和习俗,或同特意授予的权利相对的那种逐渐得到承认的权利——时,我们仍然是在原初的含义上使用该词,意指这些事情并不是某种个人意志的产物,而是无数的个人和世世代代的偶然行为的意外结果。^[17]据他观察,这个词后来逐步包含了“共同体的”(communal)这层含义,带有很强烈的人为控制和理性建构色彩,指“符合社会利益”,或者“按照社会意志”,作为把社会人格化的结果,使得人们根本无法认识到它是一种自发的秩序,“社会的”这个词普遍用来描述有意安排的行为目的。^[18]

社会原本就有属于其自身的话语体系。若从休谟和哈耶克意义上来理解“社会(的)”,那么我们就发现,在社会系统中,不管是个人向国家讨要的权利,还是国家对个人实施的权力,抑或道德和政治的起源,所有这些故事都应当得到重新讲述。从社会角度看,道德和政治的诞生都是自然而然的事件,人类理性也以某种形态在一定程度上被演化出来,但它们作为演化的产物,不应当一下子就被拔高到凌驾于它们的母体即整个社会之上。在这种拔高的过程中,人类理性起到了关键作用,因此运用社会的独特话语对理性的运用进行“划界”也是必要的。理性推进了社会的发展,但阻碍了人们对社会的认识,并且由于它对社会的过分干涉,而生成了一样那样的“社会病”。所以从某种意义上看,这种社会哲学的工作,乃是在某种意义上延续了康德的理性批判和维特根斯坦(Ludwig Wittgenstein)语言治疗的哲学抱负。

尽管许多人致力于从人类社会自然演化出美德,但许多人必定会大失所望,演化伦理学也不可能取代规范伦理学,因为演化所能产生的美