从比特币看经济理论中的计算问题*

——基于计算主义的思考

谢志刚

摘要:新兴数字货币比特币涉及的核心问题是“共识机制”,这从技术和社会经济两个层面对现有理论提出了挑战。在技术和算法层面,比特币通过区块链等多项创新初步解决了所谓的“拜占庭将军问题”,成为诸多数字货币模仿的对象;在社会经济层面,比特币作为货币的社会认同仍然存在着极大的不确定性。当前的研究较多关注技术层面问题,而忽视了比特币在社会经济层面对现有经济理论的挑战。本文借鉴计算主义的基本视角,将“认知”与“决策”视为“计算”,力图将两个层面的比特币“同意的计算”问题纳入一个较为统一的分析框架之中。作为对“强计算主义”的一种调和,本文尝试将计算划分为“简单范式”的“计算 I”与“复杂范式”的“计算 II”两种基本类型,并据此对比特币做出初步分析。计算视角的剖析进一步反思了主流经济学范式的机械最大化理性的局限,揭示了一种经济学“计算复杂范式”的可能。

关键词:比特币　计算主义　共识机制　经济计算

一、问题的提出:“同意的计算”

近年来,数字加密货币比特币(Bitcoin)的出现引发了决策层和理论界的广泛关注。作为一种源自互联网“密码朋克”运动的比特币,没有实物和主权背书,是一种数字形态的“虚拟实体”,但影响迅速扩大,很快催生了一系列新型数字货币的出现。受比特币的启发和推动,大量以比特币为基础的各种改进、创新版本的数字货币不断涌现,甚至推动了各国央行对于法定货币数字化的实验以及对法定货币体系的改革。比特币现象可以视为经济学中自发秩序的一个典型案例,且有其自身的独特性,由此带来了一系列的疑问。比特币的实质是什么?比特币是否具有货币属性?比特币是否具有价值?比特币的共识计算与货币信用是否存在内在关联?比特币与法定货币的关系如何?可以说,比特币对传统货币理论提出了数字时代的新问题。如何对它进行深层次的解释,将是经济学和现有货币理论面临的一个挑战。本文并不试图对比特币进行全面的考察,而是着眼于比特币的共识计算角度,将其与对经济学决策和计算的反思联系起来进行初步探讨。

比特币系统(Nakamoto,2008)将“区块链”(blockchain)、“不对称加密算法”和“点对点协议”等一连串的技术创新重新组合到了一起,更重要的是,其利用互联网的“信息生态”以自组织的方式产生了一种准货币体系,对当前的货币体系构成了熊彼特意义上的“创造性破坏”①。从计算主义(computationalism)视角来看,可以说比特币在技术和社会经济两个层面是依赖于“计算”的系统。

* 谢志刚,中国社会科学院经济研究所,邮政编码:100836,电子邮箱:xiezg@cass.org.cn。本文受国家社会科学基金重点课题“经济思想史的知识社会学研究”(14AZD109)、中国社会科学院基础研究学者项目“从黄金到比特币:基于知识理论视角的货币经济学研究”的资助。感谢匿名审稿人的批评和修改意见,文责自负。

①熊彼特(1947)将“不断地从内部使这个经济结构革命化,不断地破坏旧结构,不断地创造新结构”的这个过程称作“创造性破坏的过程”(process of creative destruction)。

比特币充分利用了互联网的数字技术,同时也依赖于互联网"去中心化"特征,以一种自组织的方式产生货币供给,从两个层面展示了计算的意义。借助经济学公共选择学派代表人物布坎南和塔洛克(1960)讨论公共选择的著名术语,比特币在两个层面上需要达成"同意的计算":其一是在技术层面上,需要通过去中心化的区块链等技术实现一致的、不可篡改的"总账本"或者"交易记录";其二是在公共选择和社会知识学层面,达成其货币功能和价值的共识。基于计算主义视角的分析表明,计算的两个层面既有不同通约的对立性质,也存在着某种统一性。

经济学对比特币共识机制的分析主要从两个方面展开。首先,在对主流经济学范式的批判和反思基础上,现有文献引入了复杂理论框架对其展开分析。Pilkington(2017)基于复杂理论框架批判了主流经济学的理性、均衡和自利三个基础假设,并阐释比特币是在无数个体互动与信息技术进步基础之上的宏观经济现象的涌现,而传统的货币政策和银行监管等理论都将被比特币的出现所动摇。Santos(2017, 2019)在"区块链经济学"研究之中阐释了比特币和区块链技术的逻辑基础,并以复杂系统理论框架进行考察。他的研究指出,区块链之中的工作量证明(proof of work, PoW)共识算法是非复杂系统的,而权益证明(proof of stake,PoS)共识算法则是典型的复杂系统,不能被简单的计算系统所兼容,进一步分析指出资本投机等行为在加密货币市场的作用可能导致非线性的"混沌"现象产生。Abadi & Brunnermeier(2018)从经济效率的分析角度也指出存在着"区块链三重难题"(blockchain trilemma),即区块链技术并不能兼顾理想的交易记录系统应当同时具有的"准确"、"去中心化"和"成本效率"三个特征。为保证记录的正确和去中心化,区块链系统不得不消耗大量计算资源,从而在效率上不如中心化记账系统。

其次,现有文献从"共识机制"的知识社会学角度展开了研究,这较为广泛地包括社会传播学、制度经济学、叙事经济学等新兴分析框架。Teigland et al(2013)基于奥地利学派经济学分析了比特币作为一种对抗法币系统的机制,他们指出比特币共同体(Bitcoin community)虽然遭遇欺诈、黑客攻击和法律限制等不利因素,但采用制度企业家的视角分析可以发现,其组织结构和网络结构具有自我修复能力和较强的生命力。Yelowitz & Wilson(2015)通过谷歌搜索数据甄别了比特币支持者的社会身份和行为机制,指出其支持者主要包括技术爱好者、投机者、自由主义者和网络犯罪分子。罗伯特·希勒(2017)提出"叙事经济学"的视角,并将比特币作为一个叙事研究的典型案例来考察。他分析了"投机与泡沫""无政府主义"和"对不平等的恐惧"等叙事模式对与比特币的社会认同的影响,并倡导经济学与传播学、传染病学等更多理论和方法的结合。这实际上是一种知识社会学与经济学的交叉视角。韩裕光等(2015)使用演化博弈论分析框架研究了比特币传播动力和传播速度的相互关系,他们的分析显示,交易收益、投机收益和心理效用均与比特币扩散速度成正相关关系,是比特币扩散的主要动力,其扩散是市场选择的结果。

从更一般的角度来看,近年来对于比特币的研究不断拓展深入,但较多的文献聚焦于计算机科学、算法领域,分别对比特币的区块链技术、加密算法进行研究,以及对其可能的改进和应用进行实用性的探讨。而在经济、金融领域对比特币的货币属性的分析,通常着重于从货币的本质、功能等方面辨析比特币的经济属性,或者分析其区块链等技术在金融领域拓展的可能等。赵磊(2018)、张成岗(2018)讨论了区块链技术的"信任""去中心化"和"共识"的关联问题。程炼(2020)通过对数字货币特征、运行逻辑和文化象征意义等角度的探讨,认为数字货币的出现并没有对现有主流货币理论构成实质性挑战,但揭示了反思传统概念框架并拓展其分析视角的必要性。这些研究意识到"共识机制"是比特币的核心问题,但没有就区块链计算与经济理论决策的逻辑关联做出分析。

这两方面的文献大多没有意识到可以从一种更为综合和统一的计算理论视角将比特币的社会共识问题的两个层面结合起来进行考察。"计算"这个通常被认为与数学相关的概念已经泛化到了人类的整个知识领域,并上升为一种极为普遍的科学概念和哲学概念。"计算主义"成为认识事物、研究问题的一种新视角、新观念和新方法。此外,"计算"也是一个极富争议的问题。如认知学者派利夏恩(1984)指出:"尽管经过了近 50 年的研究(从图灵关于可计算性的著名论文开始),对于什么

是计算的基本要素的问题至今还没有一致的看法。”从哲学、认知科学、计算机科学乃至于物理学等诸多领域,围绕着计算主义的争论也极为激烈。比特币的出现对于计算主义思路的发展和深化具有一定启发性。

经济学对“计算”问题最为著名的讨论是 20 世纪 30 年代“社会主义经济计算论战”①。论战在兰格等以新古典经济学为理论基础的计划主义者和米塞斯、哈耶克等奥地利学派自由市场主义者之间展开。米塞斯认为,计划经济取消了货币,必将导致无法进行理性的经济计算。在米塞斯对市场的理解中,企业家基于当前生产要素价格与预期未来消费品价格的经济计算基础上,制定他们的生产计划。米塞斯所说的“经济计算”,正是将这些预期的未来收入与当前的支出都以共同货币单位表示,并相互比较。而兰格等人认为,计划经济可以借助于一般均衡模型的思路计算出均衡价格。随着计算机技术发展以及“大数据”和“人工智能”时代的到来,从信息科学角度对计划经济进行论证的“计算机计划主义”(谢志刚,2018)观点兴起,经济学、数学和计算机科学等多个领域的理论相互结合将“经济计算大论战”延续到了现代(科特尔等,2007)。经济学是研究资源配置效率的理论体系,在一定程度上也是关于“决策”和“计算”的科学。因此,“经济计算”与认知、逻辑和数学意义上的“计算”并非仅仅是表层的关联。对于“计算”与“决策”的不同认识,也决定了不同经济学理论逻辑的基本脉络。

比特币更具革新意义的是在经济和社会层面的货币自组织实验。从经济学角度来看,货币与计算似乎有着自然的联系。货币的价值尺度功能使得其成为价值计算的基准,从而广泛地出现在个人、企业和国家的各种交易、决策、统计等计算之中。比特币在经济和社会两个层面上深刻体现出“同意的计算”含义。我们可以将比特币作为反思自发经济现象和计算主义的一个典型案例。计算主义可以对比特币的货币本质、发展路径等考察提供一种独特的视角。

本文将计算划分为两个基本类型:计算 I 与计算 II。对比特币“同意的计算”的分析在这两个计算层面分别展开。在计算 I 层面,加密和通信算法解决了所谓的“拜占庭将军问题”,而在计算 II 层面,比特币“同意的计算”仍然在进行之中,社会对比特币的接受程度仍然具有较大不确定性,该层面计算也体现了复杂系统的涌现特征。研究表明,比特币这样的数字加密货币依赖的不仅仅是机器计算,更是一种社会计算,而其真正的计算难题仍然在于计算 II 层面。“同意的计算”中的计算Ⅱ要解释的核心问题可以具体化为两个次级问题:第一,比特币被接受为一种货币,是集体选择的结果。对它的认可和接受过程,可以借助社会认知图式来给予解释。这种解释实质上也是行为经济学的一种解读。第二,人们的共同认可和赞同使比特币具有了使用价值,在此即可用于商品交换,从而作为这种集体选择基础的内生性偏好,自然地解决了个人理性选择走向集体理性的偏好“加总”问题。

二、计算的两层含义:计算 I 与计算 II

首先需要指出,虽然存在哲学家、逻辑学家和数学家对于计算的各种定义,但总体而言,计算仍然是一个尚未定论、开放的概念。在计算机和数理逻辑领域,艾伦·图灵(A. M. Tuling)和阿朗佐·丘奇(A. Church)的计算定义已经成为标准,但随着认知科学、人工智能和计算主义等研究领域的进展,计算概念实际上泛化开来,反而变得模糊。从经济研究的角度看,计算与决策本应当是经济学的核心主题,毕竟主流经济学的核心假设——理性人假设实际上就是不断进行成本收益计算的“计算人”。然而,在传统的经济研究领域,计算与决策实际是一个较为边缘化的主题,主流经济范式直接将决策与计算简化为机械的最大化。而对于真实的计算与决策分析,主要是在被视为非经济学或者非主流的如企业管理理论、奥地利学派、行为经济学和认知经济学等领域展开,故而有必要简单阐释计算概念的演进发展过程。

计算通常被认为就是数的加减乘除,或者再宽泛一些,则包括了更为复杂的方程的求解、函数的

①“社会主义经济计算论战”(the socialist calculation debate)也常被翻译为“社会主义经济核算论战”。在计算理论之中,图灵与邱奇分别使用了 computable 与 calculable 等术语,被证明为等价概念。在本文中,calculation、computation 也视为无差别,都翻译为“计算”。

微分和积分等。在计算机出现之后,一种有趣的定义就是“计算就是计算机所做的事情”(程炼,2007),而在此之前,计算可能被理解为“数学家所做的事情”①。17 世纪英国哲学家霍布斯提出了“推理即计算”,数学家莱布尼茨则将这种“推理计算”设想为一种“普遍文字”的形式系统,由此“哲学家之间已经不需要争吵”,要解决分歧只需要“让我们算算吧”(李建会,2012)。直到 20 世纪 30 年代,由于哥德尔(K. Godel)、丘奇和图灵等数学家的开创性工作,人们才逐渐形成一种较为清晰的关于计算的共识。这就是计算的“丘奇-图灵论题”(Church-Turing thesis),而“图灵计算”也成为相关研究领域的基础术语。

英国数学家图灵提出了一种抽象的“计算机器”(computing machines),并用它来定义“可计算的数”(computable numbers),开创了计算理论(可计算性)的新研究领域(Turing,1936)。图灵的论文针对的是数学家希尔伯特在国际数学大会上所提出的 23 个“数学问题”之中的“判定问题”——即对数学命题是否存在一般的、形式化的方法自动予以判定(李文林、袁向东,1981)。图灵机将数学运算分解为一系列简单重复的符号操作,实际上是以一种机械的方式展现人类如何进行数学运算——“一位持有铅笔、纸和一串明确指令的人类计算者,可以被看作是一种图灵机”②。简单而言,图灵机是一种符号处理系统的设想。如果一个问题的实例输入并在有限的变换后停止运行,这一个问题就被认为是“图灵可计算”的。图灵机对输入符号串所做的“接受还是拒绝”的判定就是“计算”,也被称为“图灵计算”。

几乎与图灵同时,美国数学家丘奇也提出了“可计算数”(calculable numbers)问题。图灵本人很快就证明了“图灵计算”与邱奇“计算”概念的等价性。在另外一条线索上,哥德尔首先在 1931 年提出了原始递归函数的概念。所谓原始递归函数,是由初始函数出发,经过有限次的使用代入与原始递归式而做出的函数。在这些研究的基础上,人们很快发现并证明了哥德尔的一般递归函数、邱奇的 λ 可计算函数与图灵机都是等价的关于计算的数理模型。这些最终促成了著名的“丘奇-图灵论题”:凡是可计算的函数都是一般递归函数,能够被通用图灵机执行。

“丘奇-图灵论题”被称作“论题”而不是“定理”,这是因为虽然丘奇和图灵对其计算进行了严格的定义,但“可计算”或者“计算”终究是一个依赖于人类主观界定的概念。正如心智哲学家塞尔(Searle,1996)指出,“计算是一个只相对于有意识的观察者和解释者才存在的抽象的数学过程”。换句话说,“你永远不能在自然中发现独立于人类解释的计算过程,因为你所可能发现的任何理过程都只相对于某种解释才是计算的”③。按照塞尔的说法,“你可以赋予任何东西一个计算解释”。计算主义的极端形式将整个宇宙的发展、演化都视为“计算”,这很容易被视为仅仅是术语名词之争。然而人类对“认知”“计算”这些主题的探究,不得不依赖于人类的认知本身,因此,“名实之争”难以避免。这也意味着“计算”问题实际上不可能局限于数学或者逻辑领域。

随着图灵计算概念的确立和计算机的发展,计算主义首先在人工智能、认知科学领域兴起,并产生了大量争论。认知计算主义的基本思想是,心理状态、心理活动和心理过程是计算,简单而言“认知就是计算”。围绕着“计算”的争论,最核心的问题在于“智能”是否就是“计算”? 争论的两极,支持者是“强人工智能”和“计算主义”学派,他们认为所有的智能,包括人类的心智,归根结底都是某种计算;而反对计算主义和人工智能的学者则坚持认为,智能特别是人类的心智并不能归结为“机械的计算”。由此基础逻辑的对立,衍生出大量的相关问题,如“心的功能是否计算”“决策是否等同于计算”等。

本文主要以认知的“计算主义”的视角来看待智能、心智和决策,但更重要的是力图将计算做出“两分法”,以反映围绕着计算主义和人工智能等领域的种种对立,同时也反映了对经济人理性理解的差异。这种划分较为鲜明地区别了主流经济学的机械式最大化的理性人决策和计算范式与非主

①数学家通常不认为自己的工作只是在“计算”,而棋手常常将自己的思考称作“计算”,这取决于对计算概念的理解。

②转引自佩措尔德(2008)。

③转引自唐热风(1998)。

流的真实决策和计算范式的不同。

计算主义倾向的学者通常主张认知等同于计算，或者至少可以还原为计算，他们对计算的层次理解较为常见地区分出“形式系统”和“意义系统”两部分。派利夏恩(1984)把计算机系统区分为三个层次：物理层、句法层和语义层。侯世达(1997)主张智能系统即计算系统，其系统内可以划分为低层的形式系统和高层的意义系统。现代心理、行为科学常见地将认知系统做出两分法。典型的如心理学家斯坦诺维奇(Stanovich，2005)从认知的角度将人类心智系统划分为系统 I 和系统 II。其中，系统 I 是“自发式系统”，而系统 II 是“描述分析系统”。

在此，本文重点考察人工智能学者王培(Wang，2007)将计算所面临的问题解决划分为“类型 I”和“类型 II”。在计算理论中，一个“问题”通常是一个“集合”，其中包含很多“实例”，而其“解决”必须是一个能处理所有这些实例的方法。一个“解法”必须是确定的和可重复的。这种问题的解决被称作“类型 I”，其所对应的“计算”即“算法”(algorithm，也可称为计算过程 computational procedure，或者图灵机)。而在数学之外的“实证科学和日常生活之中，一个问题通常并非定义良好的集合(well-defined set)，而是必须单独处理的具体实例”。这种类型问题解决输出的属性常常要根据经历和处境而定，过程具有不可预知性和不可重复性。这被称为“类型 II”。王培(Wang，2007)认为，对于“类型 II”的问题的解决，即便是依赖于计算机算法，如神经网络这样的人工智能学习系统，但由于其对于“经验”的依赖性和不可重复性，其过程不能被视为“图灵计算”。

由于计算问题的复杂性，不能简单地以“形式”“意义”或者“机械的”“复杂的”这样的单一视角的维度来划分计算。参考斯坦诺维奇(Stanovich，2005)、王培(Wang，2007)等研究，本文将“计算”划分为“计算 I”和“计算 II”。首先对计算 I 与计算 II 的基本过程做出一个形式化的近似表达。定义图灵映射：

$$f:U \to Y \tag{1}$$

即图灵映射为从输入向量 $U(u_1,u_2,\cdots,u_n)$ 到输出向量 $Y(y_1,y_2,\cdots,y_m)$ 的图灵计算过程。图灵映射是数学函数和映射的拓展形式，包含了函数和映射的递归过程。

定义计算 I 为：

$$f:X,U \to \dot{X},Y \tag{2}$$

其中，f 为从状态向量 $X(x_1,x_2,\cdots,x_i)$ 和输入向量 $U(u_1,u_2,\cdots,u_n)$ 到状态向量 $\dot{X}(\dot{x}_1,\dot{x}_2,\cdots,\dot{x}_i)$ 和输出向量 $Y(y_1,y_2,\cdots,y_m)$ 的图灵映射。这实际上是合并了状态方程和输出方程的状态空间模型的简化一般形式(胡寿松，2007)。将计算 I 视为一种控制系统，那么图灵映射是系统的外部模型，即只描述了输入和输出，而此处计算 I 的状态空间模型则包括了内部模型，即输入向量和状态向量决定新的状态向量，新的状态向量拟合输出向量。状态向量在状态空间模型之中一般随时间变化而写成微分或者差分形式，在此代表不同计算阶段的差异。

定义计算 II 为：

$$F:X,U \to \dot{X},Y,F'(X') \tag{3}$$

其中，F 首先类似于 f，为从状态向量 $X(x_1,x_2,\cdots,x_i)$ 和输入向量 $U(u_1,u_2,\cdots,u_n)$ 到状态向量 $\dot{X}(\dot{x}_1,\dot{x}_2,\cdots,\dot{x}_i)$ 和输出向量 $Y(y_1,y_2,\cdots,y_m)$ 的图灵映射。但与映射 f 不同之处在于，其输出结果还包括了对自身框架的改变，即 $F'(X')$，并且 X' 为新的状态向量 $(x_1,x_2,\cdots,x_j)$。这里映射框架的改变不仅是映射 F 的更新，也包括了状态向量维度 $i \to j$ 的变化。这是计算 II 的一个阶段过程描述，其整体计算可以成为 $F \to F' \to F'' \cdots \to F^k$ 这样一个反复迭代过程。此外，可以把 F 看作包括从输入向量到输出向量的映射 f，以及从基于输出向量与预期结果的误差试错而反向更新映射框架的反向映射 $\hat{f}$ 这样两个图灵映射的复合：

$$F\begin{cases} f:X,U \to \dot{X},Y \\ \hat{f}:\dot{X},Y \to F'(X') \end{cases} \tag{4}$$

从形式化的标准来看,这里只是对计算划分的一种视角描述,而非严格定义。在严格的状态空间模型之中,状态向量是描述系统状态的最小变量集。而在计算 II 的描述之中,只有在完全信息条件下,假设 $X(x_1,x_2,\cdots,x_i)$包括了所有的潜在状态变量,才符合标准状态空间模型。生成的 $F'(X')$ 的反向映射 f 在其层次足够"复杂",其映射逻辑无法"理解"或者"解释"的情况下,才成为复杂系统的"涌现"。这种形式化的近似描述,除了能够在此更为准确地对计算做出两分法之外,更重要的是,正如后面的图式理论分析将显示的,能够把计算理念与意识、认知等难以形式化的理论结合起来获得更为规范的解析。

考虑塞尔对于"计算"主观性的论述,对于计算 I 与计算 II 的划分很难给出一个严格形式化的界定。这正如"复杂""可解释性"和"可理解"等界定类似,这些对于系统描述的属性在很大程度上并非系统自身的"客观性质",而是依赖于人类这样的主体观察者,具有强烈的主观主义认知倾向。以神经网络和深度学习为代表的人工智能技术的突飞猛进,正代表了人工智能和计算领域从计算 I 向计算 II 系统的飞跃。人工智能从物理符号系统转向联结主义的路线,在获得技术功能上大幅进步的同时,"计算"的"复杂性"和"不可解释"也越来越成为其遭受诟病的根源。机器学习被认为遭遇了"可重复性危机"(reproducibility crisis),甚至被批评为"炼金术"(alchemy)①。这指的不仅仅是某个算法和计算过程是无法解释的"黑箱",而是指整个"机器学习"的系统和方法都成为"黑箱"。我们并不理解我们创造出来的东西。哪怕通过神经网络得出了我们预期的结果,我们仍不知道它是如何工作的,只是知道初始条件、参数,但并不能用一种形式化科学以我们能理解的形式将它表示出来。

这种表达的关键在于将计算 I 视为一种静态的映射系统,而计算 II 则是一个动态演化系统。计算 II 处于一个动态演化的过程之中,完整的计算 II 过程实际上包括了在不断的输入输出之中,映射本身的不断演进 $F\rightarrow F'\rightarrow F''\cdots\rightarrow F^k$,并得到最终结果。

从图灵计算的一般递归性质来看,计算 I 与计算 II 都存在着计算的递归或者自指部分。但对于计算 II 而言,这种动态递归的性质更为重要。对于计算 I 而言,递归仅仅是数据的再次输入,而在计算 II 之中,结果的递归意味着对系统控制的反馈,从而使得计算 II 过程表现为反复迭代的动态试错。通常由于复杂系统之中普遍存在的"自然选择"压力,使得计算 II 的映射 F 不断发生迭代优化。

由于其计算框架实际上在不断地演化变动,其计算与计算 I 的确定性大不相同。计算 I 是静态的,对于输入 X 输出 Y 是确定的,并且这种计算的每一个步骤都是可逆的。而对于计算 II,即便输入同样的 X,其输出结果包括 Y 和 F 都完全可能不同。计算 II 是历史条件依赖的(嵌入历史的),具有时间的箭头,不可逆转。从复杂系统的视角来理解,计算 I 是微观层面上的运行,而计算 II 则是在更高的、宏观层面的"涌现"。这正如物理学在微观层面可以观测到的分子个体的布朗运动,可以是完全遵循牛顿力学法则,是完全可逆的,然而在宏观层面的热力学意义上,这种热运动是不可逆的熵增过程。按照计算主义的观点,这甚至不是类比,而是物理系统之中的运动就被视为计算本身。从一般性来看,计算 I 与计算 II 都是状态向量的操纵和映射,但计算 II 可以称为一种状态向量的"动态演变"。

另一方面,从计算 II 的 $F\rightarrow F'\rightarrow F''\cdots\rightarrow F^k$ 整个生命周期统一来看,则仍然可以将其视为静态计算 I。如果将计算 II 的动态可变框架 F 抽取出元映射 f,而将可变框架的结构和参数视为元映射的输入向量,那么整体看,从起始状态到最终结果,计算 II 又可以被视为计算 I。正如王培(Wang,2007)所承认的,"类型 II"这样依赖于初始条件和经验(数据)积累的学习系统,只需要将其算法从最初始状况开始考虑,其整个"生命周期"仍可以被看成是一个图灵计算过程,其输入由它历史上所经历的所有输入符号串衔接组成。由此,在任意局部而言,计算 II 都是非决定论的、不可逆、不可重复的,但总体而言却可以是决定论的、可逆的、可重复的,只需要完全一致的历史输入,计算结果将是确

①参见 Hutson(2018),"AI researchers allege that machine learning is alchemy", https://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy;Leopold(2019),"Machine learning takes heat for science's reproducibility crisis", https://www.hpcwire.com/2019/02/19/machine-learning-reproducability-crisis-science/。

定的。从这个角度来看,计算 II 的可变框架可以理解为一种可变算法 F,而可变算法 F 有着一个元算法 f。元算法 f 的逻辑和结构是可理解的、可解释的,但在其历史演算过程之中的 F 其逻辑结构成为不可理解、不可解释的计算 II。

计算 I 与计算 II 的划分,既符合计算主义的基本立场,也是一种调和主义。首先,按照计算主义的立场,计算 I 与计算 II 都属于图灵计算。计算 I 与计算 II 可以理解为计算的两种类型,也可以理解为计算的两个不同层次,它们在本质上是可以统一的。其次,计算 I 与计算 II 的统一又是不可通约和不可还原的。从计算思想的数理逻辑源头来看,罗素和怀特海的《数学原理》、希尔伯特纲领(Hilbert's Programme)代表了试图将数学分析和推理统一化、机械化的思想,而在此就意味着将一切计算归结为计算 I 的尝试。哥德尔不可能定理对这种思路给予致命一击。哥德尔本人实际上坚持反对将认知归结为计算的计算主义立场。计算 I 与计算 II 的不可通约性也可以在一定程度上反映类似哥德尔的这种反计算主义态度。以这样的计算两分法,可以看到关于心智、智能和决策等很多方面的对立都可以划归到其中。

表 1　计算 I 与计算 II 的对比

计算 I	计算 II
典型:算术、形式逻辑	典型:模式识别、模糊计算
线性系统	复杂系统
连贯的	非连贯的
机械的、简单的	直觉的、创造性的
逻辑的	涌现的
可模拟	不可模拟(只能完整模拟)
可预测、决定论的	不可预测、非决定论的
推理	类比
存在最优解	自然选择条件下的满意解
理念性的	实践性的
经济计算 I:会计核算、最优化、一般均衡	经济计算 II:企业家决策、创新、市场过程

对于“经济计算”,它不简单归属于其中一个类型,而是相应存在着两个部分的子系统:“经济计算 I”与“经济计算 II”。其中,“经济计算 I”正是主流经济学理性经济人的求解“最优解”的过程。这是在既定框架约束条件下的机械计算过程。而“经济计算 II”则对应着奥地利学派、行为经济学等理论所主张的人类个体面对不确定性的复杂决策。从这个计算两分法的视角很容易看到,社会主义经济核算大论战的双方对于“经济计算”的理解正好符合这个两分法。兰格等人主张的计划经济计算的可行性,指的是包括“会计核算”“投入产出”和“线性规划”这样的机械计算。而哈耶克等人主张的“理性计算”不可行则是指依赖于企业家直觉判断的个体决策在计划经济之中不复存在。这并不说市场体系之中的企业家或者个体不需要计算价格成本这样的相对简单机械的“计算 I”,而是说他们面临和承担(计算 I 意义上)“不可计算”的不确定性时做出的决策和行动才是真正重要的“选择”,并且对于资源配置起到优化作用。可以说,主流经济学范式及其形式化的取向是典型的经济计算 I 的理论进路,而奥地利学派、制度学派乃至于新兴的演化经济学、行为经济学等很大程度上注重于经济计算 II 的理解和研究。经济学家托马斯·萨金特曾针对机器学习的人工智能热潮表示:“人工智能其实就是统计学”①,这似乎模糊地体现了主流经济学的计算观点。类似于“生物学终究不过是物理学”或者“计算 II 本质上仍然是计算 I”这样的命题,虽然并不能算错误,但在很大程度上忽略了不同层次的不可通约性。

①参见李峥:《该怎么消除人工智能与生俱来的“歧视”》,《新京报》2018 年 10 月 16 日,http://www.bjnews.com.cn/opinion/2018/10/16/511431.html。

奥地利学派对于主流经济学理性范式的批评能够从另外一个角度揭示这种经济计算划分的意义。现代奥地利学派指出新古典经济学实际上是一个严格限定在给定的“目标－手段”(ends-means)的框架做出系统性分配性行为的分析架构(福斯、克莱因,2012)。当然目标是可以变化的,但是旧目标集被新目标集所取代是完全在经济选择这一行为之外的,这种机械的选择是在给定框架下的必然结果。经济当事人在做出决定时实际上已没有选择权,其他选择已经被给定的这一框架所排斥。这种机械的既定框架的决策可以理解为一种计算,但只能是计算 I 类型。奥地利学派关注的是主观主义视角下真正的具有不确定性的预期和决策,这是计算 II 意义上的计算和决策。

斯坦诺维奇(Stanovich,2005)的心智功能划分极具启发性,但与本文计算 I 与计算 II 的顺序正好颠倒,即自发式系统 I 对应于本文的计算系统 II,而分析系统 II 对应于计算系统 I。这种错位原因在于对理性、决策或者计算的基础与拓展的层次理解不同。斯坦诺维奇(Stanovich,2005)、卡尼曼(2012)等对理性决策的看法与主流经济学较为一致,即强调所谓“最大化”为理性决策;本文的计算划分则与赫伯特·西蒙(1982)、吉仁泽(2000)、哈耶克等强调的有限理性、演化理性的思路更具一致性。

三、“拜占庭将军问题”中的共识计算 I

比特币作为一种计算机网络信息系统,其重要创新在于通过区块链等技术和算法上解决了“拜占庭将军问题”。通过对拜占庭将军问题及其解决的分析,可以看到,其主要是比特币系统“同意的计算”在计算 I 层面的问题和解决。“拜占庭将军问题”(Byzantine generals problem)是计算机科学家兰波特等(Lamport et al,1982)提出的一种分布式对等网络中的容错问题①。这个“拜占庭将军问题”的逻辑及其解决,是比特币系统运行的核心机制。

“拜占庭将军问题”基本内容如下:假设有几支拜占庭军队正在一个敌城外扎营,每支军队由一个将军指挥。将军之间只能通过信使传递信息。观察完敌情后,他们必须制定一份共同的行动计划。然而,有些将军可能是背叛者(traitor),他们会尽力阻止那些忠诚的将军达成一致的计划。将军们必须有一个算法来保证如下条件:

条件 A:所有忠诚的将军必须达成相同的行动计划。忠诚的将军将会做该算法要求他们做的事情,但是叛变的将军可以做任何他们想做的事情。

无论叛变的将军会做什么,算法必须要保证条件 A。诚实的将军不能仅仅达成一致,他们还应该达成一个合理的行动计划(agree upon a reasonable plan)。所以我们也要确保:

条件 B:当少数人是叛军的时候,他们无法导致忠诚的将军接收到糟糕的计划。

兰波特等(Lamport et al,1982)指出,计划的“合理”(reasonable)或者说“糟糕”(bad)难以形式化确定。假设 $v(i)$ 代表第 i 个将军发送的信息。每个将军使用某种方法来根据这些信息 $v(1),v(2),\cdots,v(n)$ 拟定作战计划(n 代表将军的总数)。通过让所有的将军使用同一种方法就可以满足条件 A,通过使用一种稳健方法(a robust method)也可以满足条件 B。比如,现在需要决定是进攻还是撤退,$v(i)$ 代表第 i 个将军关于进攻还是撤退的意见,最终的决定可以通过在他们之间进行一个多数人决定的投票来实现。在这种情况下,只有当持两种意见的忠诚将军数目几乎相同时,少数的叛变将军才能影响最终的结果。但是在这种情况下,无论是进攻还是撤退都算不上是糟糕的方案,也就是满足了条件 B。简单地说,正确的做法就是每个忠诚的将军都正确地表达了自己的意思,不会因为叛徒让别的将军认为忠诚的将军是叛徒而不采用他传达的消息。

“拜占庭将军问题”简化成了:所有忠诚的将军都能够让别的将军接收到自己的真实意图,并最终一致行动;而形式化的要求就是,“一致性”与“正确性”。经过这样的设定,兰波特等把问题转化成

①兰波特等(Lamport et al,1982)注意到“把问题以故事的形式表达出来更能引起人们的关注”,杜撰了这样一个“拜占庭将军”的案例。

为:在存在背叛者的情况下,如何形成一致的序列$\{v(1),v(2),\cdots,v(n)\}$,就是如何达成“共同信念”。由于背叛者 i 可能给其他人发出不同的信号 $v(i)$,所以必须采用某种方法对每个人所接受到的信号组进行统一。

首先考虑一个“中心化”的“拜占庭将军问题”。依照兰波特等的基本设定,但增加一个假设存在着“拜占庭皇帝”,可以对诸多“拜占庭将军”进行中心化的行动协调。如图 1 左图所显示的,皇帝(0)对各将军进行统一的行动协调。这里隐含着:皇帝(0)必然是对自己忠诚的,他不会隐瞒或者歪曲与将军之间的信息和命令传达。

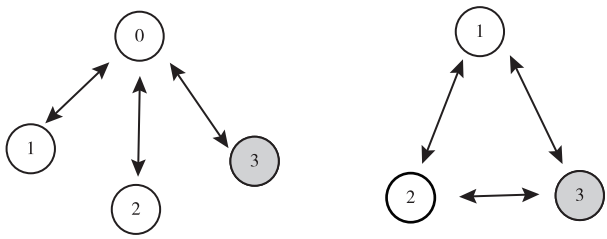

图 1 “中心化”与“去中心化”的拜占庭将军问题

在中心化的模型之中,多个个体的行动计划的“一致性”和“正确性”问题很容易解决,或者说几乎不存在问题。就这里的设定而言,将军将各自的判断和决策传递到中心即皇帝(0)这里。然后,皇帝根据一个比如“多数决定”的原则,统一众将军的意见,如果多数人主张“进攻”,那么便向各位将军下达“进攻”命令,否则便命令“撤退”。即便存在着背叛的将军,如图中带阴影的将军(3),但只要背叛者数量没有超过半数,便始终能够保持兰波特等所定义的整体行动的“一致性”和“正确性”。其中,“一致性”由皇帝自身的忠诚所保证,而“正确性”是一种放宽的限制条件,也就是假设多数将军的观察和决定就是“正确”的行动方案。

然后对比“去中心化”的“拜占庭将军问题”。假设不存在皇帝(0)作为行动计划的协调者,各位将军完全处于对等状态,考察这种情况下如何实现行动计划的协调。如图 2 中所显示,各位将军之间都存在着通信渠道,可以相互沟通验证信息和行动计划。然而“拜占庭将军问题”的困难立刻就显现出来了。仅考虑图 2 中对于将军(1)的行动建议的识别问题。

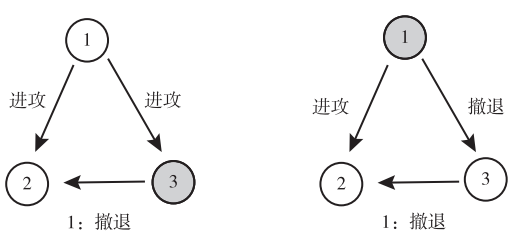

图 2 拜占庭将军问题的困难所在

注:根据 Lamport et al(1982)修改。

假设如图 2 所显示,仅有三位将军,并且存在一个背叛者。对于将军(2)而言,他完全无法鉴别出图中左右所显示的两种情况中哪一种是真实情况。因为将军(2)在两种情况下所获得的信息完全是一样的,即将军(1)提议“进攻”,而将军(3)却告诉将军(2):“将军(1)提议撤退”。这有可能是将军(3)是背叛者,篡改了将军(1)的信息,也可能由于将军(1)是背叛者而发出了不一致的信息。这里仅对于诸多将军之一的信息进行甄别就无法完成,从而要形成对全部将军的信息和行动计划进行统一便更成为不可能,也就是说最终无法达成满足“一致性”和“正确性”的“共同信念”。这就是兰波特等(Lamport et al,1982)所说的:“它的难解之处是通过一个令人惊讶的事实而体现出来的。”根据兰波特等的论证,“如果将军们只能发送口头消息,除非有超过三分之二的将军是忠诚的,否则该问题无解”。

当然,“拜占庭将军问题”的解是可以通过对于通信的条件设定而获得的。它的解就是所谓的“共识算法”或者“共识方法”,通过算法手段让各参与方对某个确定的结果达成一致的方案。寻求共识机制的概念和方法是一个数学领域长期以来的研究方向,尤其是在计算机领域针对分布式共识机制取得许多新进展。就“拜占庭将军问题”而言,存在着许多解决方案,如“实用拜占庭容错算法”(practical Byzantine fault tolerance)和“工作量证明”(PoW)、“权益证明”(PoS)[①]等等。新兴的去中心化“比特币”就是设计了极具实用性的“工作量证明”方法并依赖于此而运行的。

比特币系统将一连串的创新重新组合到了一起。这一连串的创新主要包括“互联网的产生”“不对称加密算法”“点对点协议”和中本聪所独创的“工作量证明”等。其中,“不对称加密”指的是使用公钥和私钥技术可以获得一个信息的“数字签名”,而试图伪造这个签名的计算极为困难,但用公钥验证签名的计算却极为简便。不对称算法解决了“拜占庭将军问题”之中的“口头信息”问题,使得信息传递难以被篡改和伪造。“点对点协议”是一个去中心化的信息发布系统,已经在互联网被成熟的BitTorrent系统用于相对小的用户子集之间共享许多文件,比特币利用它来在分布式结点之间传输和保存所有的交易记录,也就是所谓的“区块链”(blockchain)。

一笔比特币系统的交易就是生成一个交易记录,并且分布式地保存到诸多结点之中。由于比特币的交易记录全部都是公开的,因此,支付方是否拥有足够的比特币,完成这笔交易,这是可以验证的。一个典型的交易是用户发起一次比特币交易,需要向各结点提出交易金额、持有比特币的来源哈希编码(Hash)、本次交易双方的地址、支付方的公钥和支付方的私钥生成的数字签名。结点对上述信息逐一验证是否属实。确认交易的真实性以后,交易还未完成。交易数据必须写入数据库后才算成立,对方才能真正收到钱。比特币使用了区块链作为数据库。首先,所有的交易数据都会传送到矿工(miner)那里,矿工负责集合若干数量的交易,并把这些交易写入一个区块链之中,然后计算这个区块的哈希值。计算哈希值的过程叫作采矿,这需要大量的计算。矿工之间也在竞争,谁先算出哈希值,谁就能第一个添加新区块进入区块链,从而享受这个区块的奖励或者交易手续费用。

比特币对于“拜占庭将军问题”的解决方案是,它为发送信息加入了成本,这降低了信息传递的速率,并加入了一个随机元素以保证在一个时间只有一个“将军”可以发布信息。比特币加入的成本是基于计算一个随机哈希算法“工作量证明”。哈希算法是一种常用的计算机算法,在比特币系统之中哈希算法对一组数据,也就是“区块”进行计算获得一个与其他区块相区别的识别字符串。虽然哈希算法本身效率很高,但比特币系统特意设置了动态“难度系数”,使得符合要求的哈希值极为罕见而不得不依赖于大量的计算和时间消耗。首先计算出来并且符合要求的哈希值结果可以被比特币系统接受,成为“工作量证明”。

一个计算出下一个有效哈希值的结点相当于一个“拜占庭将军”,他把所有之前的信息放到一起,附上它自己的计算结果,加上签名发布给其他将军。其他将军接收到并验证通过了这个将军的信息,就会停止自己的信息整合,使用新的信息更新他们的已知信息,再次在此基础上对于新的交易重新开始计算。哈希计算竞赛从一个新的开始点重新开始,如此这般循环往复,网络持续同步。比特币不存放在数字钱包或其他别的地方,而是只存在于记录了全部历史交易的区块链上面。

由此,“工作量证明”“不对称加密”加上“点对点协议”,使得一个不可信网络变成了一个可信的网络,进而使得所有参与者可以达成“共同信念”,如攻击计划,一系列的交易、政治投票系统,或者任何其他的需要分布式协议的地方。这正是一种数字意义上的布坎南“同意的计算”。比特币对于“拜占庭将军”问题的解决,虽然应用了区块链、加密等诸多复杂的技术创新,并且产生并依赖于大量的数据,但其各个计算阶段都是“简单的”“机械的”计算Ⅰ。整个比特币系统在实际运行之中会不断扩充计算节点、交易数据,但其基本计算框架本身是既定的和不变的。所以,其核心算法可以视为计算

①权益证明依赖于比特币持有数量和时间为要素的所有权,Santos(2019)对比分析了PoW与PoS两种共识算法,认为后者计算属于复杂系统。

I 的映射 $f:X\rightarrow Y$。事实上，比特币的基本计算 f 作为一种基础算法框架，其源代码被公布在互联网上，成为诸多新衍生数字货币的基础。然而这些新的衍生或者改进的数字货币系统通常是由其他软件工程师进行修改。比特币的算法本身并不涉及对自身计算框架的改进。因此，在此意义上，比特币“同意的计算”是静态的计算 I，而非动态的计算 II。如果将比特币及其衍生的数字货币系统都考虑进来，这实际上加入了其他数学家、软件工程师的计算，这样的复合系统将变成一种更为复杂广泛的“社会计算”，从而成为计算 II 系统。然而，比特币系统在经济社会的应用和发展本身就蕴含着一种更为直接的“社会计算”，这就是比特币作为货币的社会共识的形成。

四、比特币的社会共识计算 II

比特币作为一种无形的虚拟数字货币而能够被相当多的人所接受，这对传统经济学货币理论提出了挑战。主流的货币理论通常对货币本质持“商品论”与“信用论”两种观点。信用货币论无法解释比特币的货币属性，而商品货币论则在一定程度上与比特币相容。特别是按照奥地利学派货币理论和主观价值论，比特币虽然是一种数字产品，但仍然可以具有其用户所赋予的“主观价值”，进而具有了成为商品货币的可能。更重要的是，尽管个人主义主观价值论提出了商品货币的形成原理，但对于法定货币、信用货币和比特币这样的新型货币，还未能给出具体的动态过程分析，这需要引入知识社会学与公共选择理论的框架。

塞尔也曾注意到货币的“社会建构”特征：“我从钱包里拿出一张纸币，反过来看，正面看，发现它不过是一张由某种纤维丝所制成的纸，面上印着某种颜色和符号。是什么使这片纸成为货币？塞尔的回答是，只有当且仅当人们都把这片纸认作是货币时，它才是货币。”[①]现代知识社会学的兴起，就是一种类似于塞尔的“社会建构”理论，但主要论题在于对人类知识的社会建构方面。知识社会学是一个方兴未艾的领域，尚未有统一的范式。德国社会学家哲学家曼海姆和舍勒是知识社会学的代表人物，而马克思的“社会存在决定社会意识”则被视为知识社会学思想的早期萌芽。简单而言，知识社会学关注的是人类的知识如何依靠社会而形成。在知识社会学意义上，货币不过是一种“集体意向性”，一种“社会共同信念”，或者说一种“集体幻觉”。

在此，首先将比特币成为货币被社会共识所接受的过程称为“同意的计算”。布坎南所谓的“同意的计算”(the calculus of consent)，指的是“解释或描述用以协调冲突的利益的手段”，布坎南指出：“在某种真正的意义上，经济学理论也是一种集体选择理论，且因此而为我们提供了一种解释：对于独立的个人利益，如何通过交易或交换的机制来加以协调”(布坎南、塔洛克，1960)。

“拜占庭将军问题”及其解决处理的是信息系统容错问题，看似较为表层的技术问题，但将其中将军的忠诚理解为公共选择之中的真实偏好的隐藏和显示，容易看出其实质仍然是公共选择问题。对“拜占庭将军问题”的设定稍做修改，便可以更深入一步地进入到公共选择的种种“不可能”问题之中。在此假设“拜占庭将军”面临着三种选择：A 进攻、B 驻守或者 C 撤退。其中，每一个将军都基于自己的某种判断得出了一个行动偏好序列。

表 2　拜占庭将军的“公共选择”问题

将军(1)	A 进攻	B 驻守	C 撤退
将军(2)	B 驻守	C 撤退	A 进攻
将军(3)	C 撤退	B 驻守	A 进攻

这实际上是“孔多塞投票悖论”的一个转化形式。如图 3 所示，在存在“拜占庭皇帝”并且“独裁”的情况下，也就是皇帝具有自己的偏好序列并且直接下达指令，那么将军们很容易达成统一行动。如果要按照“民主集中制”的原则形成统一决策，那么这样的偏好序列将导致悖论。由于将军(1)和

①转引自韦森(2004)。

(2)都认为 B 好于 C,根据少数服从多数原则,那么将军们的"社会共识"也应认为 B 好于 C;同样,将军(2)和(3)都认为 C 好于 A,共识也应认为 C 好于 A。由此共识应当认为 B 好于 A。但是,实际上将军(1)和(3)都认为 A 好于 B,出现矛盾无法达成共识。

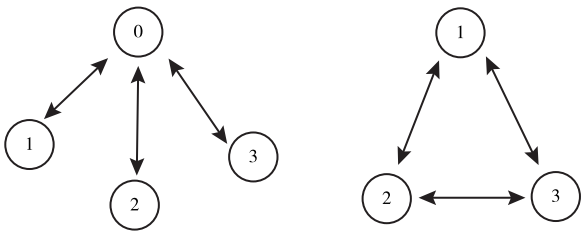

图 3 "拜占庭将军问题"的"公共选择"版本

阿罗不可能定理将"孔多塞投票悖论"这样的问题做出了形式化和一般化的解析。阿罗指出,在完备性、传递性和非独裁性等几个条件下,将个体偏好序列加总为社会偏好序列是不可能的(缪勒,1979)。方法论个人主义是经济学的基础方法之一,其核心主张在于,集体行为或者社会现象的解释应该根据个体的物理、心理状态、行动及个体间的互动来解释,强方法论个人主义则进一步要求一切集体或者社会现象可以并且应当还原为个体层面。阿罗的公共选择理论、伯格森－萨缪尔森社会福利函数等都是这种方法论个人主义的反映,但其中存在着一个重要的争议便是对个体的"加总"问题。福利经济学认为社会福利是个人福利或效用的某种函数,阿罗公共选择理论也采用类似的思路,"任何给定环境下的社会选择都是个人偏好的加总(aggregation)问题"(阿罗,1951)。布坎南虽然也秉承方法论个人主义,但受奥地利学派经济学的影响,他反对阿罗这种对个体的"加总"思路,反对将社会福利函数视为"集体理性"。

事实上,按照布坎南的方法论个人主义立场,他很可能反对将"同意的计算"做出知识社会学的解释。布坎南指出:"如果接受一种有机体的概念,那么集体抉择理论就被大大地简化了",这虽然"仍有可能进行有益的讨论",但"'个体'这个术语在真正的有机体概念之中没有多少位置"(布坎南、塔洛克,1960)。基于对"德国政治哲学家"的反对,布坎南认为:"只有关于社会的某种有机体概念,才能够设想出一种神秘的普遍意志的出现,这种普遍意志是被独立地演绎出来的,不依那种由独立的个人做出的政治选择来控制的决策过程为转移"。由此,他反对"德国政治哲学家"的有机体概念和立场,"不仅仅是否认国家作为某种超个人的实体而存在",而是"在一开始我们就得拒绝对集体做出任何有机体的解释"(布坎南、塔洛克,1960)。布坎南对于方法论个人主义的强调有其革新意义,经济学意义上的公共选择理论正是建立在对于传统政治哲学和民主理论的方法论批判和突破基础之上。布坎南对于"社会"持有一种本体论取消主义立场,即否认将"社会"以"集体理性"这样的概念视为"实在"。

布坎南认为如果把集体理性定义为最大化地达到某个设定社会目标的行动,而理性又纯粹地是根据个体行为定义的,那么更好的社会选择所体现的是"在个体行为中具有更多的理性,而不是说该选择具有更多的集体理性"(Buchanan,1954)。正如布坎南所指出的,"社会选择理论家曾经尝试界定'社会选择函数'或'政府选择函数'可能具有的合理属性。但是,所有这些努力都败于这样一个中心矛盾,该矛盾涉及一个单个人的偏好或价值排序转向为了反映由人组成的群体或共同体的潜在选择基础而设计的排序。许多社会理论分析的焦点一直是著名的阿罗不可能定理。这一定理说明,基于个人评价的社会选择函数不可能存在,除非违反从一开始就强加在这个函数上的一个或多个合理条件或者属性"(布坎南,1991)。基于此,布坎南反对伯格森－萨缪尔森社会福利函数这样的对个体价值偏好序列的社会加总的方法,他甚至认为阿罗的公共选择理论也隐藏着类似的反个人主义立场。在相关讨论之中,阿罗使用"社会福利函数"指称从个人排序得到社会排序或社会选择的过程,后来使用了"集体理性"和"议程"(constitution)等术语(Buchanan,1954),并认为布坎南与自己不存

在实质分歧，而只是术语之争。

从个体理性到集体理性的“加总”问题，在计算的角度来看就是计算 I 与计算 II 的通约性问题。伯格森－萨缪尔森社会福利函数是对个体价值偏好序列的社会加总，此种社会福利函数如果成立的话，其计算属于“计算 I”。然而，阿罗不可能定理则清晰地表明了，真实的社会共识的计算则属于“计算 II”。计算 I 意义上的、绝对的、客观的社会共识是不存在的，然而在计算 II 意义上的“同意的计算”则是可能，甚至就是现实。人们对于“社会”“民族”“国家”和“货币”等许多概念总是能够在很大程度上达成共识。人类社会的运转也必须依赖这样许多的“同意的计算”才能够得以维系。

“拜占庭将军问题”的“公共选择”版本揭示了，不存在绝对的、客观的“公共选择方案”。如果以计算 I 的方式来求得其确定解和最优解，在一定条件下是不可能的，从而按照计算 I，则此类问题是不可解的，因而也是不可计算的。事实上，阿罗不可能定理、森的不可能定理、赫维茨显示偏好不可能定理等一系列不可能定理都是计算 I。这些不可能定理，即便经过各种修补和增加约束条件，都表明了公共选择问题或者说“同意的计算”在计算 I 的层面上是无解的，至少可以说，不存在通用的、标准的、一致的解法。

布坎南实际上是具有强烈奥地利学派主观主义方法论倾向的经济学者。他批判主流经济学范式“在严格意义上说，这种理论根本不是关于选择的理论。个体在其中不会做出选择，他们的行为只是对所处环境中所发生的可客观度量的变化做出某种反应，而这种反应性行为是可以预见到的”(布坎南，1969)。布坎南认为企业家行为所展现的不是在“偏好－约束”框架已定条件下的被动选择，而是一种“创造性选择”。正如沙克尔所言：“决策，从字面意义上讲意味着一种断裂……过去与未来之间的断裂”①。最大化的逻辑是计算 I，既定框架和约束下实际上“别无选择”(have no choice)，只需要按照算法机械执行即可；而真正重要的“决策”(decision)实质上是“决断”，这种“断裂”，正是计算 II 的非连贯性。

另一方面，布坎南在对主流经济学范式的批评的同时，其公共选择理论自身也在很大程度上局限于计算 I。布坎南的方法论个人主义只看到并承认微观层面的计算 I，而忽略了问题的多层级性和复杂系统的涌现等。对于社会选择学派强调方法论个人主义的视角并没有错误，但不能由此忽略视角的多样性和层级性。知识社会学所指出的，“社会理性”是否存在或者是否是实体这样的问题并不重要，正如“国家”“社会”和“货币”等概念一样，即便是基于想象的“虚构”共识，仍然在人类社会之中发挥着重要作用。事实上，以布坎南等为代表的公共选择学派整个关于“同意的计算”都是计算 I。依照布坎南的“一致同意原则”，公共选择的“同意的计算”实际上无解，这也是公共选择学派的理论困境所在。正如布坎南所承认的，“一旦假定个人利益是同一的，经济学理论的主体部分就湮灭了”，“经济学理论解释的是，人们为什么通过交易来合作：他们这样做就因为他们是不同的”(布坎南、塔洛克，1960)。在调和的立场上，可以说，公共选择同意的计算，不仅仅是计算 I，更是计算 II，并且只有在计算 II 的意义上才能真正进行计算并获得解决。布坎南所提出的公共选择问题是计算 II 层面的，但其方法论却仍然是计算 I 层面的。

将计算与知识社会学观点联系起来看，“社会共同信念”或者“共同知识”的观点容易产生静态化的理解。人类的知识体系不能够静态化为百科全书，知识体系本身重要的还在于基础的方法论和视角等动态内容。可以使用“认知图式”这样的术语来表达关于比特币的社会共同知识，强调其中的动态计算含义。在 1952 年，哈耶克出版了心理学著作《感觉的秩序》，以“适应分类系统”“地图”和“模式”等核心概念表达了一种关于人类动态知识结构的理解和解释，并构成了一种复杂系统的“哈耶克认知图式”(谢志刚，2016)。这种图式理论是哈耶克社会自发秩序的认知理论基础。哈耶克心智理论中“分类”构成了其认知图式的基础单元，而这些基础单元的相互连接形成的“地图”(map)和“模式”(model)就是心智秩序(Hayek，1952)。哈耶克使用了“地图”的比喻来描述心智系统的知识结构，他指出，“地图”是一个“准恒定联结系统”(semi-permanent connexions)，它“表征的并非当前环

①转引自布坎南(1969)。

境,而是生命体在其全部过去的经历”。与之相对照,“模式”的理解则是:“任何时刻从既有的半恒定网络之中被发现(traced)的脉冲刺激组合方式可以被视为一种对于特定环境的‘模式’(model),在其中有机体感受到当下的自身存在,并且能够将此环境纳入其全部行动的考量”(Hayek,1952)。简单而言,“地图”就是认知主体既有的知识总和,同时也是一个知识结构;而“模式”则是从“地图”之中所选取的与当前环境条件的局部匹配。

认知图式可以理解为包含计算框架在内的“状态向量”。由此,比特币同意的计算 II 层面的结果是其社会认知图式,也就是在社会经济无数个体参与之下所形成的动态的知识格局,其不仅仅包括计算所获得的静态的信息,还包括了动态演化的计算框架。哈耶克的地图实质上是一个“关联网络”。如图 4 所示,在哈耶克图式之中,A 与 Y 之间存在某种因果关系,A 是 Y 的产生原因或者条件,但不一定是充分或者必要的条件。这是建立在主体认知之上的一种信念逻辑链条。“模式”是在“地图”基础上对当前外部条件的匹配。假设当前面临的决策条件与 A、B、C 和 Y 等单元相关,那么如图虚线框所示的关联网络的一个局部:$\{A,B,C\}\rightarrow Y$ 便被提取出来。认知主体以此“模式”作为对当前环境的辨识和决策的行动依据。

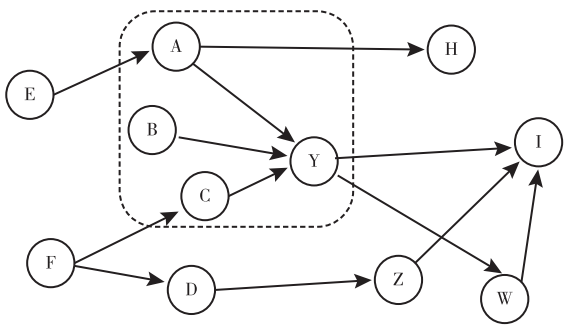

图 4　地图与模式:$\{A,B,C\}\rightarrow Y$

静态地看,在哈耶克认知图式之中,每一个节点都可以视为一个状态向量 $X(x_1,x_2,\cdots,x_i)$,地图则是全部的状态向量以及状态向量之间关联的集合。模式即计算 I 的一个映射 $f:X,U\rightarrow \dot{X},Y$,其中,输入向量 U 和输出向量 Y 根据具体条件由特定的状态向量所替代。因此,一个简单的、机械的选择或者决策过程,也就是从认知地图之中找到相应的模式,进行计算 I 层次的映射过程。

更为重要的是,哈耶克认知图式是动态结构。“地图”之所以是“半恒定”结构,关键原因在于“地图”和“模式”存在互动关系。这里不仅仅是“地图>模式>行动”,即“地图”提取“模式”,进而指导“行动”的单向路径,而是还包括了“行动>模式>地图”的反方向作用。行动所产生的体验感觉,往往并不能与原有的地图和完美地匹配产生模式,而是会出现偏差。尤其是较为新颖体验的模式匹配很有可能存在很大困难,甚至干脆无法匹配。种种情况都意味着需要依据不完全匹配模式及其偏差对原有地图进行调整修正,极端情况下甚至需要进行较大范围的地图重构,这在哈耶克认知图式的语境之中就是图式创新或图式转换。在简化意义上,“地图>模式>行动”的过程是决策过程,而“行动>模式>地图”则是学习认知过程。由此,哈耶克认知图式动态的解释实际上是计算 II,其“模式”的更为准确的模型是状态向量的 F 映射,即包括了 f 与 $\hat{f}$ 的双向映射动态过程。

比特币在罗伯特·希勒(2017)所提出的“叙事经济学”框架之中成为首要的和重点的知识社会学分析案例。叙事经济学主要是“研究其他人对重大经济事件的讲述,即像病毒般传播的流行叙事”。希勒(2017)在比特币叙事案例分析之中,分析了“投机与泡沫”“无政府主义”和“对不平等的恐惧”等叙事模式。他特别指出主流经济学极大地忽略了“从经济结果到叙事变化的反向因果关系”。这种所谓的“反向因果关系”正是计算 II 的核心特征之一。在此可以将希勒的比特币叙事与计算的认知框架结合起来,阐释比特币的社会图式过程。参考希勒,可以列举出一系列与比特币相关的认知决策图式的状态向量:

表 3 比特币决策相关状态向量

A	预算约束
B	投机需求
C	投资需求
D	货币认知
E	身份认同
F	好奇与探索
G	无政府主义倾向
H	个体的比特币社会认同判断

事实上,不同个体关于比特币的各种状态向量维度和权重都完全不可能一样。各个状态向量之间也可能存在着子状态向量的交叉重叠和形成模式。这些状态向量相互连接构成了个体关于比特币的认知图式,而个体图式进一步合成关于比特币的社会认知图式。

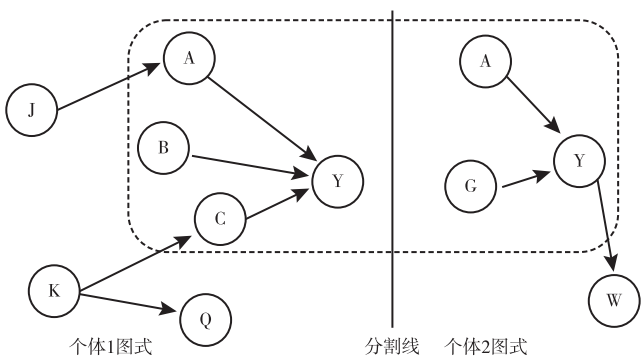

图 5 比特币社会认知图式的个体合成

图 5 显示了一个简单的比特币社会认知图式。它由两个个体认知图式合成而来。分割线的左右两边分别为两个个体关于比特币的认知和决策的图式。其中,个体 1 主要基于 A、B 和 C 的状态向量驱动而产生 Y,即“持有一定数量的比特币”的决策。在此,Y 可以视为计算 II 的输出向量,而 A、B 和 C 等状态向量都可能作为触发向量,也就是输入向量而产生 Y 决策。可以用计算模型表达为:$f:A,B,C\rightarrow Y$。而对于个体 2 而言,他完全可能拥有一个类似的或者完全不同的产生类似输出向量 Y 的认知决策模型,即 $f:A,G\rightarrow Y$。虚线框表示为两个个体关于比特币认知图式的“合成”,也就是个体 1 与个体 2 所构成的“集体”的“社会认知图式”。这里以简化的方式展示了社会认知图式的基本逻辑。如果接受个体认知图式作为一种“实在”(reality)的话,那么社会认知图式正是这种合成的社会实在,而不是主流经济理论类似社会福利函数这样的“加总”简化假设。

考察一个比特币决策模式:设想预算约束作为输入向量的模式触发。个体在获得某个新的财务预算条件下,如突然获得一笔收入可以用于投资,“获得比特币”成为一个灵感的涌现。此时,可以以预算约束 A 为输入向量,而其他向量均可视为状态向量,进入一个对比特币的“计算”过程。通过对上述 A、B、C 和 D 等一系列向量的评估和综合计算,最后得到一个输出向量 Y,决定占有一定数量的比特币。这种决策模式显而易见不会是标准模型。不同个体会基于不同认知图式和不同的状态向量形成特定模式。作为触发器的输入向量也完全可以是上述各种状态向量的任意一个。如对于货币本质的认识改变、身份认同和从众心理(“朋友们都在讨论和购买比特币”)等因素都有可能触发“获得比特币”的决策计算模式。

社会计算和认知图式的视角有助于理解比特币的深层逻辑。主流实证经济学的基本思路是寻求一个简化的因果逻辑,并找出状态向量 A、B、C、D…等之中具有决定性的因子与比特币的持有量等向量进行“拟合”。这种方法具有较强的可操作性和实用性。但是如果比特币的这种社会计算属

于计算 II 的话,那么将很难找到一个简化的回归模型。并且由于计算 II 的动态、复杂等特性,即便由于大量具有类似决策图式的个体存在而在某些阶段似乎可以得到结果显著的拟合,但也完全可能由于微小的扰动而导致认知图式的大规模分化或者变形。对于属于计算 II 的复杂的人类决策而言,经济学的实证方法很难发挥出作用。事实上,与物理学家等比较,经济学家的预测通常不具有公信力。正如罗伯特·希勒(2017)所指出的:"研究数据的经济学家在创建抽象理论模型和分析短期经济数据方面有着出色的表现。他们可以准确地预测未来几个季度的宏观经济变化,但在过去半个世纪里,他们的年度预测总体上是毫无价值的。"希勒因而倡导将叙事传播学、传染病模型等定性、定量方法引入经济分析之中。对比特币的社会计算和社会认知图式的分析也只是基本思路和概念的初步梳理。计算和认知模型启示着包括人工智能、知识图谱等更为广阔的现代科学研究理论和方法,都可能应用到对比特币的分析之中。

五、结论与展望

虽然比特币在数字技术等方面体现出了巨大的创新,但其本质上的创新和革命意义仍然不在于技术层面,而是来自其作为货币的"去中心化"与"同意的计算"方面。因此,一味强调其所谓"区块链"等技术革新,而将其仅仅视为一种"数字化货币"的看法实际上并没有抓住其货币的基础逻辑创新。围绕比特币仍然存在着大量的争议,包括比特币的价值和价格极不稳定、区块链的"数字挖矿"计算的算力浪费、比特币总额的绝对限定等。这诸多的问题仍然有待深入实践和研究。在此,计算主义能够初步提出某些启发性的视角。面对比特币价格起伏极大、难以作为流通货币这样的问题,逻辑上很容易设想通过改进比特币的供给系统,使得其具有某种程度的适应性,自动调节数量及价格。从计算主义的视角特别是计算的两分法可以发现,这种设想实质上是将比特币系统的计算 I 与计算 II 联系起来,从而其可行性可以归结为一般系统计算 I 与计算 II 是否可以通约问题。随着人工智能、大数据的科学技术的进一步发展,不难设想比特币系统与某些机器学习、自适应系统的结合而产生更为复杂和稳健的数字货币体系。然而,这种具有"炼金术"性质的算法体系与依靠简明清晰经济机制和社会制度保障的金本位、国家法定货币甚至是哈耶克所谓的"非国家化"的货币体系相比较,哪些会在真实的"社会计算"之中胜出,还有待于进一步的研究和实践检验。

通过从计算主义视角分析比特币,本文得到一些重要的启示:

第一,比特币仍然是一个极具争议性的事物,对其是否具有货币本质、能否履行货币职能、如何进行相应的监管等重要问题,学界尚没有较为统一的意见。就比特币的货币性质争议来说,其需要被置于传统货币理论关于货币的本质、功能以及货币与金融、货币政策等一系列的理论框架之中,得到更进一步的审视和分析。比特币的自发兴起表明,实践常常超前于理论,其背后隐含着深刻的经济、社会和技术背景。比特币的出现是一种计算 II 层面复杂系统"涌现"现象,是布坎南所谓的"同意的计算",也是哈耶克所谓的"自发秩序"。比特币诞生并且获得社会关注之前,很难预测甚至想象到比特币会获得如此巨大的影响。另外,作为难以解析、无法预测的复杂系统,比特币是否能够成为货币,其发展前景尚不明朗。个体心智、社会认知图式、经济秩序在从个体到社会的不同的层面,体现出同样的计算 II 层面的自组织与复杂系统特征。对这些问题的深入理解也有待于公共选择、演化、计算和复杂系统、知识社会学等领域的分析框架的进一步发展和完善。由于其对金融市场和经济社会的巨大冲击性,决策和监管层面采取审慎的态度是有必要的,但理论界应当以更为开放的态度面对这类新事物。

第二,计算技术和人工智能等技术的发展将会给社会经济领域带来像比特币这样越来越多的自发秩序现象,从而对经济学的某些正统理论提出新的挑战。作为一种基于信息处理技术的数字货币,比特币典型地反映了货币的一般"计算"性质。比特币在很大程度上可以视为一种数字的"金本位"。作为一种对比,金本位时代的黄金成为货币,其解决计算 I 层面的"同意"问题依赖于黄金的物理属性——重金属的稳定性、容易分割和便于检验。而金本位在计算 II 层面的共识问题则与比特币

完全一致，在很大程度上属于一种自组织现象。计算主义的分析视角可以拓展到对包括金银这样的商品货币和法定货币为代表的信用货币等领域。传统经济学货币理论有可能在比特币和计算主义的视角下获得某种重构。

第三，比特币重要的技术创新——区块链技术解决的仅仅是较为表层的计算 I 问题。真正的公共选择困境仍然不是可以依靠区块链这样的计算 I 层次的方法能够处理的。计算的不同层面问题仍然需要在不同层面解决。以比特币为代表的数字货币所面临的真正挑战仍然在于经济社会制度层面的“社会计算”。这取决于比特币系统自身的算法机制和对货币本质、职能的理解和社会共识的形成和发展。

第四，比特币的产生不仅仅是作为一种新兴货币而对现有的社会经济系统产生影响，更重要的是其新技术的应用和经济社会实践对人类哲学、认知科学和经济学等诸多方面的理论、逻辑、方法和思想上的冲击。计算主义正是在信息计算、人工智能和类似比特币这样的数字系统的影响之下逐渐兴起的崭新的研究视角。计算主义本身正处于与认知科学、复杂系统科学、经济学等研究领域相互启发、交相呼应、发展深化的过程之中，其对于经济学、货币理论的影响也将逐渐深化。就经济学而言，比特币促使人们重新反思“货币”“计算”和“理性决策”等重要主题。经济人理性、选择和决策等核心理念，甚至是整个经济学的方法论和范式都有在这样的计算主义框架之下重新反思的可能。

参考文献：

阿罗，1951：《社会选择与个人价值》，上海世纪出版集团 2010 年中译版。

科特尔 等，2007，《计算、复杂性与计划——再谈社会主义核算论战》，《当代世界社会主义问题》第 2 期。

布坎南 塔洛克，1960：《同意的计算：立宪民主的逻辑基础》，中国社会科学出版社 2000 年中译版。

布坎南，1969：《成本与选择》，浙江大学出版社 2009 年中译版。

布坎南，1991：《宪政经济学》，中国社会科学出版社 2004 年中译版。

程炼，2007：《何谓计算主义？》，《科学文化评论》第 4 期。

程炼，2020：《数字货币：从经济到社会》，《社会科学战线》第 6 期。

丹尼尔·卡尼曼，2012：《思考，快与慢》，中信出版社（中译版）。

丹尼斯·缪勒，1979：《公共选择理论》，中国社会科学出版社 2010 年中译版。

福斯 克莱因，2012：《企业家的企业理论：研究企业的新视角》，中国社会科学出版社 2020 年中译版。

格雷戈里·蔡汀，2012：《证明达尔文：进化和生物创造性的一个数学理论》，人民邮电出版社 2014 年中译版。

韩裕光 孙伟 朱力，2015：《比特币的崛起：扩散速度与扩散动力》，《华东经济管理》第 3 期。

赫伯特·西蒙，1982：《现代决策理论的基石》，北京经济学院出版社 1989 年中译版。

侯世达，1997：《哥德尔、艾舍尔、巴赫：集异璧之大成》，商务印书馆。

胡寿松，2007：《自动控制原理》，科学出版社。

M. Sipser，1997：《计算理论导引》，机械工业出版社 2000 年中译版。

吉仁泽，2000：《适应性思维：现实世界中的理性》，上海教育出版社 2006 年中译版。

李建会 符征 张江，2012：《计算主义：一种新的世界观》，中国社会科学出版社。

李建会 赵小军 符征，2016：《计算主义及其理论难题研究》，中国社会科学出版社。

李文林 袁向东，1981：《希尔伯特数学问题及其解决简况》，《数学的实践与认识》第 3 期。

罗伯特·希勒，2017：《叙事经济学》，中信出版社 2020 年中译版。

唐热风，1998：《心的本质是计算吗？》，《自然辩证法研究》第 4 期。

派利夏恩，1984：《计算与认知：认知科学的基础》，中国人民大学出版社 2007 年中译版。

佩措尔德，2008：《图灵的秘密：他的生平、思想及论文解读》，人民邮电出版社 2012 年中译版。

塞尔，1996：《社会实在的建构》，世纪出版集团、上海人民出版社 2008 年中译版。

韦森，2004：《货币、货币哲学与货币数量论》，《中国社会科学》第 4 期。

谢志刚，2010：《对均衡的再认识——由社会主义经济核算大论战所引发的反思》，《社会科学战线》第 4 期。

谢志刚，2016：《资本的秩序与哈耶克认知图式》，《学术研究》第 12 期。

谢志刚，2018：《哈耶克知识问题中的信息与知识论》，《人文杂志》第 6 期。

熊彼特，1947：《资本主义，社会主义与民主》，商务印书馆 1999 年中译版。

张成岗,2018:《区块链时代:技术发展、社会变革及风险挑战》,《人民论坛·学术前沿》第12期。
赵磊,2018:《信任、共识与去中心化——区块链的运行机制及其监管逻辑》,《银行家》第5期。
Abadi, J. & M. Brunnermeier(2018), "Blockchain economics", CEPR Discussion Paper, No. DP13420.
Buchanan, J. M. (1954), "Individual choice in voting and market", *Journal of Political Economy* 62(4):334－343.
Hayek, F. A. (1952), *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, University of Chicago Press.
Lamport, L. et al(1982), "The Byzantine generals problem", *ACM Transactions on Programming Languages and System* 4(3):382－401.
Nakamoto, S. (2008), "Bitcoin: A peer-to-peer electronic cash system", https://bitcoin.org/bitcoin.pdf.
Pilkington, M. (2017). "Bitcoin through the lenses of complexity theory", In: R. Martin & J. Pollard(eds), *Handbook on the Geographies of Money and Finance*, Edward Elgar.
Santos, R. P. D. (2017), "On the philosophy of Bitcoin/Blockchain technology: Is it a chaotic, complex system?", *Metaphilosophy* 48(5):620－633.
Santos, R. P. D. (2019), "Consensus algorithms: A matter of complexity? －Markets, communications networks, and algorithmic reality", In: M. Swan et al(eds), *Blockchain Economics: Implications of Distributed Ledgers*, World Scientific.
Stanovich, K. E. (2005), *The Robot's Rebellion: Finding Meaning in the Age of Darwin*, University of Chicago Press.
Wolfram, S. (2002), *A New Kind of Science*, Wolfram Media Inc.
Teigland, R. et al(2013), "Breaking out of the bank in Europe－Exploring collective emergent institutional entrepreneurship through Bitcoin", https://papers.ssrn.com/sol3/papers.cfm? abstract_id=2263707.
Turing, A. M. (1936), "On computable numbers, with an application to the entscheidungs problem", *Proceedings of the London Mathematical Society* 42:230－265.
Wang, P. (2007), "Three fundamental misconceptions of artificial intelligence", *Journal of Experimental & Theoretical Artificial Intelligence* 19(3):249－268.
Yelowitz, A. & M. Wilson(2015), "Characteristics of Bitcoin users: An analysis of Google search data", *Applied Economics Letters* 22(13):1030－1036.

The Calculus of Consent in Bitcoin

— An Economic Study from the Perspective of Cognitive Computationalism

XIE Zhigang
(Chinese Academy of Social Sciences, Beijing, China)

Abstract: The "calculus of consent" problem in the emerging digital currency Bitcoin system poses theoretical challenges from both technical and socio-economic levels. At the technical and algorithmic level, Bitcoin has initially solved the so-called "Byzantine generals" problem through a number of innovations such as the blockchain, and has become the target of many digital currency imitations; at the socio-economic level, the social identity of Bitcoin as currency still remains unsettled. From the perspective of computationalism, this study regards "cognition" and "decision" as "calculation", and tries to put the consent problem of Bitcoin into a more unified analytical framework. As a reconciliation of "strong computationalism", this study attempts to divide computing into two basic types: "calculation I" of the "simple paradigm" and "calculation II" of the "complex paradigm". The analysis from the computational perspective further reflects on the limitations of the mechanical maximization rationality of the mainstream economics paradigm, and reveals the possibility of a "computational complexity paradigm" of economics.

Keywords: Bitcoin; Computationalism; Calculus of Consent; Economic Calculation

(责任编辑:何伟)
(校对:刘洪愧)