



# 算法统治世界

智能经济的隐形秩序

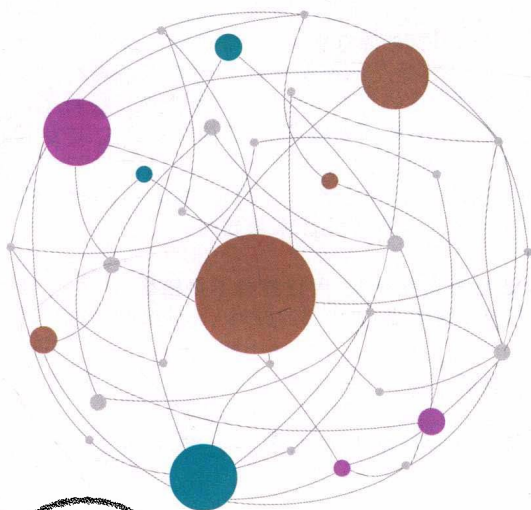
徐恪 李沁 著

The Invisible Order of  
the Intelligent Economy

畅销书《赛博新经济》作者最新力作



清华大学出版社



# 算法统治世界

智能经济的隐形秩序

徐恪 李沁 著

清华大学出版社  
北京

## 内 容 简 介

今天,互联网已经彻底改变了经济系统的运行方式,经济增长的决定性要素已经从物质资料的增加转变成成为信息的增长。但是,只有信息的快速增长是不够的,这些增长的信息还必须是“有序”的。只有“有序”才能使信息具有价值,能够为人所用,能够指导我们实现商业的新路径。这种包含在信息里的隐形秩序才是今天信息世界的真正价值所在。经济系统的运行确实是纷繁复杂的,但因为算法的存在,这一切变得有律可循,算法也成为新经济系统里那只“看不见的手”。那么算法究竟是如何保障信息的有序增长的?又是如何改变和控制我们的经济系统的?在这本书中,我们将赛博经济系统从上至下进行了梳理,把推荐算法、分配算法、匹配算法、动态定价算法、大数据处理算法、数据交易算法、隐私保护算法和区块链技术及相关算法在新经济组织中的运行做了深入浅出的阐述,从而为读者打开了新经济的现在与未来的大门。本书虽有专业的深度,但也适合一般读者阅读,同时这些创新背后的算法逻辑也将帮助企业更好地规划自己的商业模式和未来战略。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

算法统治世界:智能经济的隐形秩序/徐恪,李沁著. —北京:清华大学出版社,2017  
ISBN 978-7-302-48808-8

I. ①算… II. ①徐… ②李… III. ①计算机算法—应用—经济系统—研究 IV. ①F014.9

中国版本图书馆 CIP 数据核字(2017)第 265983 号

责任编辑:龙启铭

封面设计:傅瑞学

责任校对:李建庄

责任印制:宋 林

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:北京亿浓世纪彩色印刷有限公司

经 销:全国新华书店

开 本:170mm×230mm 印张:23.5 插页:1 字 数:442 千字

版 次:2017 年 12 月第 1 版 印 次:2017 年 12 月第 1 次印刷

印 数:1~3000

定 价:69.00 元

产品编号:076220-01

如今,互联网如同高速公路一样,成为了人类社会的信息基础设施,在经济领域催生出众多前所未有的商业新形态,并开始影响和改变整体社会经济系统的运行方式。当网络与经济系统不断融合,催生出了一一种全新的经济系统,我们将之称为“赛博经济系统”(Cyber Economic System),它是以广泛连接和海量数据为基础,以互联网技术为载体,以算法为内在驱动力的新型经济系统。

新经济系统已经到来,但我们还没来得及真正深入探寻其内在规律,特别是 2008 年的经济危机,很少有经济学家做出提前预测的现实,让经济学领域一时陷入低谷,纳西姆·塔勒布在《黑天鹅》一书中将原因归结为“因变化而需要知识的事物,通常是没有专家的”。的确,经济学面对的问题太复杂,复杂到几乎不可预测。虽然经济领域正变

得越来越复杂,但天气系统其实同样复杂,而借助超级计算机,我们至少可以预报近期的天气。类似的例子还有 AlphiGo,当 AlphiGo 战胜李世石时,人类棋手第一人柯洁痛苦地承认:“人类几千年甚至没人沾到围棋真理的边”,而 AlphiGo 为围棋真理打开了新的大门。这些例子在启发我们思考,是不是我们也没有沾到经济系统运行真理的边?也许我们需要一些新的思路。

麻省理工学院的物理学家塞萨尔·伊达尔戈在《增长的本质》一书中给出了一个有趣的想法:“经济是通过实体化信息,增强人类对于知识的实际应用的系统”,“经济增长的本质,是因为信息(量)的增长”。如果我们采用“增长”的概念来重新定义赛博新经济,那么赛博新经济是信息增长更快的经济。赛博世界为信息的增长提供了快速路径,这是由赛博的特征决定的。赛博的特征包括小世界、信息级联、幂律和长尾等,这些都是赛博世界所独有的,这也让赛博时代的信息增长速度超过以往任何时候。

不过,只有信息的快速增长,还是不够的。这些增长的信息还必须是“有序”的。只有“有序”才能使信息具有价值,能够为人所用,这是一种隐形的秩序。这种包含在信息里的隐形秩序,才是今天信息世界的真正价值所在。那么我们该如何保证信息按照有序的方式增长?经济系统的运行确实是纷繁复杂的,但因为算法的存在,使一切变得有律可循,从某种意义上说,算法已经成为新经济系统里那只“看不见的手”。

那么算法究竟是什么?又有哪些算法正在改变和控制着我们的经济系统?在这本书中,我们将赛博经济系统从上至下进行了梳理,把推荐算法、分配算法、匹配算法、动态定价算法、区块链技术及相关算法、大数据处理算法、数据交易算法和隐私保护算法如何推动新经济系统的运行做了详细阐述。当然,仅仅理解和掌握这些算法,并不能保证对经济系统运行做出准确预测,但当我们理解了这些算法和它们所代表的赛博经济的隐形秩序后,或许再次面对经济系统的不确定性时,会增加一份信心,相信也会帮助大家从中找到

解决问题的全新视角。

这里要多说一句,身为计算机科学工作者,从事经济学方面的研究似乎有不务正业之嫌,但最近经济学知名期刊 *Econometrica* 的主编、经济学家 Joel Sobel 在一次会议报告中指出,当前经济学中最有趣的研究是计算机科学家完成的,因为计算机科学家有场景、有平台并能真正通过实验验证理论。这给了我这样的研究者莫大的鼓励,我个人认为,未来经济学的重大突破很可能会来自计算机科学工作者和经济学工作者的合作研究,让我们共同努力,使这一天早日到来。

本书的研究得到了国家自然科学基金(61472212)支持,在此深表谢意。在本书的编写过程中,我们借鉴吸收了许多国内外专家学者的研究成果,在此也致以诚挚的谢意。本书部分内容曾经在“网络科学与策略机制”“互联网发展导论”和“互联网发展与创新经济”课程中进行过讲授,同学们也提出了非常宝贵的意见,在此一并致谢。教学的快乐也许就在于此吧。感谢我的研究生苏辉、李立、李彤、吴波、吕亮、付瑶、张欣欣、张宇超、杨帆、姚文兵,感谢我的合作者李沁女士。还要感谢清华大学社科学院的王勇教授,他的洞见和观察使我获益良多。

限于水平,书中不可避免存在欠缺之处,恳请读者批评指正。读者可以通过关注“赛博新经济”公众号和我们交流。

徐恪

2017 年 10 月

## 算法定义的新经济系统

5G 通信技术已经悄然出现,这个号称“人类历史上最复杂的通信系统”,预计将在 2020 年开始商用化;2016 年 3 月 15 日,当韩国棋手李世石在第 180 手投子认输的那一刹那,宣告了以 AlphaGo 为代表的人工智能技术开启了崭新的一页;2016 年 8 月,中国将世界上第一颗量子通信卫星“墨子号”送入太空轨道,并完成了人类历史上首次量子通信过程;中国在关键技术上的重大突破,有望使量子计算机在四五年内进入实用阶段……人类正在加速进入赛博时代。

如果把时间倒推三十多年<sup>①</sup>,那时的人们应该很难想象:用来打电话的手机能够完成这么多事情;素不相识人

---

<sup>①</sup> 1986 年,美国 NSFNET 建成,有人认为这是互联网真正诞生的标志性事件。

之间可以跨越时间和地域的限制相识相知；任何一个角落里发生的事情都可以在几分钟之内传遍世界；普通人也能一呼百应，影响千千万万的人。赛博时代，高度发达的互联网络在人与人之间、人与物之间，架起一道道看不见的连接线，将素不相识的人、貌似毫无关联的物，连接在了一起，缩短了时空的距离。

同时，这种无所不连的连接也打破了原有经济系统中的生产、消费、市场等环节，形成了覆盖更广、连接度更高、效率更快的新经济系统，即赛博新经济。赛博新经济自出生之日起就展现出极强的生命力，发展速度远超过去任何一个经济时代，这其中还涌现出许多充满活力的新机制、新模式。随着算法不断地渗透，信息增长的秩序和赛博新经济系统的运行秩序也不断地被它重新定义和改写，从某种程度上，我们甚至可以说，它决定了经济系统的不断进化。

## 信息与秩序

人们曾经以采集食物为生，而如今他们要重新以采集信息为生，尽管这件事看起来很不可思议。

——马歇尔·麦克卢汉，原创媒介理论家、思想家

小米董事长雷军曾说：“投身信息产业的怀抱快三十年了，我有时也在想：信息何以会具备如此强大的力量？它的力量来自哪里？我们又该如何驾驭这一力量？”作为见证中国互联网崛起的企业家，雷军依然对信息这个词有所困惑，更不要说普通人。

20世纪40年代末，信息的概念开始出现，并以迅雷不及掩耳之势扩展

到很多学术领域。信息既不是宏观的,也不是微观的,它可以刻在古巴比伦人用于记录的泥板上,也可以隐藏在一段生物 DNA 中。它几乎适用于所有学科领域,这种超然的性质使得它在人类的知识体系中具有非常重要的地位。

1948 年 7 月,时年 32 岁的克劳德·艾尔伍德·香农(Claude Elwood Shannon)在 *The Bell System Technical Journal* 上发表了论文 *A Mathematical Theory of Communication*(通信的数学理论)。这篇划时代论文的发表,标志着现代信息论研究的开端,截至 2017 年 6 月,这篇文章的 Google Scholar 引用量已经超过六万。在这篇论文里,香农带给人们一个自创的新词“比特”。香农对这个新词的解释是“测量信息的单位”,到现在,比特已经和千克、分、摄氏度等一样,成为人们日常生活中常见的量纲。在香农提出的信息理论中,信息熵是一个重要的概念,可以用来定量反映信息的不确定程度。例如,在自然语言处理中,中文的静态平均信息熵是 9.65 比特,而英文是 4.03 比特,中文的信息熵大于英文,说明中文的复杂程度更高,词义更丰富,但处理起来也更困难。

在信息理论提出后,为了让人们更好地理解这一理论,香农指出需要把信息和含义区分开,不能混淆在一起。这样的考虑是出于工程和哲学两个方面的原因。从工程技术的角度来讲,香农当时致力于制造一种能够使信息得到传递的机器,该机器无须了解所传递信息的真实意义。从哲学的角度来讲,“含义”与“信息”这两个词,其实是完全不同的概念。

首先,信息本身是无意义的。麻省理工学院的知名物理学家塞萨尔·伊达尔戈在其所著的《增长的本质》一书中肯定了“信息是无意义的”这个概念。伊达尔戈认为,对人类来说,要将信息和含义区分开是很困难的,因为人们会不自觉地将自己理解的含义注入信息,并认为这是理所当然的。例如对于“1111”这一串数字,喜欢网络购物的人想到的是“双十一”购物节;单身狗们

想到的可能是光棍节；而对于学计算机的人，可能想到的是二级制的数字15。从这个例子可以看到，对于信息“1111”来说，不论是购物节还是光棍节，都是人们在无意中强加上去的含义，这些意义并不是该信息本身的一部分，但人们无法控制这种解读信息的本能。有时候，信息和含义可能会完美地结合在一起，但这也并不意味着信息本身是有意义的。

其实，信息是没有实体的，但信息又是物质性的，它能够被物理性的表达呈现出来。信息可以理解为一种物理秩序，或是物理事物的组合方式。对于汉字来说，每一个汉字都可以看作笔画的组合，或者是笔画的排列秩序，人们对其中一些组合赋予或强加了意义，才形成了文字；英文单词也是一样，是英文字母的排列秩序。显然，这样的笔画或字母的秩序远远大于汉字或英文单词的数量，那些没有被赋予含义的组合仍然是不具备意义的信息。从这个角度出发，上面所讲的信息“1111”的例子就很好理解了：1111的不同含义其实是代表了信息“1111”的某一种特定秩序。

在赛博新经济时代，信息的重要作用日益凸显，并在社会经济增长方面迸发出巨大的能量。《增长的本质》一书指出：“经济是通过实体化信息，增强人类对于知识的实际应用的系统”“经济增长的本质，是因为信息（量）的增长”。这里，信息增长促进赛博经济增长，包含了两层含义：首先，信息能够在赛博世界中快速增长。赛博时代具有有别于以往任何时代（如农业时代、蒸汽机时代）的特征，这些特征包括小世界、信息级联、幂律和长尾等，这是赛博时代所独有的由赛博带来的特征<sup>①</sup>。在这些特征的作用下，赛博时代中信息的增长速度超过以往任何时候。赛博世界为信息增长提供了快速路径，并且这样的路径只存在于赛博世界中。其次，促进经济增长的信息是“有序”

---

<sup>①</sup> 更多关于赛博特征的内容，可参见《赛博新经济：“互联网+”的新经济时代》（清华大学出版社）。

的,即这些信息中包含了某种隐形的秩序。赛博时代中如果只有信息的快速增长是不够的,正如前面所讲,增长的信息可能是无序的,没有任何意义。物质的排列秩序是多种多样的,这些排列秩序里,有些是有意义的,有些则可能什么都不是。例如一团杂乱无章的纤维是纤维这种物质的一种排列秩序,而在另一种排列秩序下,这团纤维可以编织为毯子。显然,纤维的两种排列秩序都是信息,相对于杂乱无章的秩序,第二种排列比第一种更为“有序”,这种“有序”使信息具有了某种含义和价值,能够为人所用,这就是一种隐形的秩序。我们可以把这种信息称为“知识”。可见,这种包含在信息里的隐形的秩序才是信息的价值所在,才是这个世界所需要的,才能促进社会经济增长。

在赛博世界中,每时每刻都有大量新的信息产生,也有老旧信息消亡。对于人类社会经济系统来说,那些包含有秩序的信息(也就是知识)才是需要的,正是不断增长的知识才促进了社会、经济不断向前发展。赛博(互联网)的存在,只是信息能够快速增长的基础,这些混乱的信息需要一种力量去有序化,这股力量就是在赛博世界里不断渗透和不断演进的算法。从赛博的本质,到赛博系统中的各种上层应用,其背后都有算法的存在。算法保证了信息的有序增长。

此外,算法还在建立另外一种秩序——赛博新经济的运行规律。从经济系统底层信息、数据的产生、增长,到上层的具体应用实践,算法都在其中起着决定性的作用。信息的有序增长促进经济增长,算法决定着信息增长的秩序,同样也决定了赛博新经济系统的秩序。

## 新经济还需要另一只“看不见的手”

他这样做只是被一只看不见的手引导着，去促进一个并不是出自他本心的目的。

——亚当·斯密

随着互联网和移动互联网的发展，建立在其上的应用和服务，无论称为“互联网+”还是“+互联网”，我们都能看到它对于人类社会的巨大改变。三十年前人们很难想象，三十年后人们只需要简单地点几下手机，足不出户就可以处理大多数日常事务。今天，一个普通人的一天很可能如此度过：

早晨，当你一觉醒来，会使用手机控制灯光亮起、享受智能厨具做好的美味早餐，浏览今日头条推送的新闻事件和好友的微信留言。出门上班，通过在线查看实时交通状况并规划出行路径和方式，可以选择搭乘网约车，或是公共交通。在公司，按照日程安排处理公务、收发邮件、参加视频会议。到了中午，可以一边享受网上预订的外卖，一边追一追感兴趣的美剧，或是听听在线音乐。午休时间，可以打理下网络理财产品，或是逛逛电商的网站。下班后，可以尽情吃喝玩乐，在朋友圈秀自拍。

我们对上面描述的这一切已经习以为常，并将这一切理所当然归功于互联网。然而，“互联网”是一个浅层次的答案，不够准确，也不够本质。两百多年前，英国经济学家亚当·斯密（Adam Smith）在《国富论》中用“看不见的手”来描述市场机制在经济运行中的作用。与此相似，我们发现在赛博世界，在经济运行以及这些日常活动的背后，也存在着另一只看不见的手，在有条不紊地操纵着一切，维持着系统的正常运转。

而这只看不见的手,就是“算法”。日常行为的背后所涉及的一些算法如表 0.1 所示。

表 0.1 日常活动背后的算法

| 活动 /应用    | 涉及的算法        |
|-----------|--------------|
| 手机控制智能家居  | 无线通信与加解密     |
| 浏览推送新闻    | 推荐           |
| 视频会议      | 包调度、音视频编解码   |
| 网络约车      | 匹配、路径规划、动态定价 |
| 网络购物      | 推荐、区块链       |
| 在线音乐、影视平台 | 推荐、音视频编解码    |
| 拍照美化      | 图像处理         |
| 网络金融      | 分配、区块链、匹配    |
| 交通状况查询    | 大数据处理        |

可以说,几乎所有的经济运行流程,或是人们的活动,都是由算法在幕后重新建立一种秩序。就拿在电商网站购物来说,当你选择了购买某种商品,算法会判断这种商品剩余数量是否能满足你的购买量,如果可以满足,则会将你选购的商品放入购物车;如果剩余数量不足的话,会返回一个提示信息。在结算时,算法会根据商品单价和购买数量,计算你需要支付的金额,在这个过程中,算法也会自动查询你是不是有可用的优惠券或满足打折条件。如果你选择在线支付,算法会显示相应的支付界面。付款后,算法会计算出送货时间,并显示在你的手机或计算机屏幕上。一个简化的网络购物流程如图 0.1 所示。

通过上述购物流程,可以体会到算法在人们日常活动中所起的作用,必须按照算法规定的流程来操作,否则就不会成功。其实不只是人们的日常事务,就连赛博新经济的基础——互联网,保证其正常运行的核心因素也是算法。如果您是计算机网络领域的从业者,也许对此并不陌生。这类算法有个专门的名字,称为“网络协议”。网络协议就是为了便于计算机网络中的不同

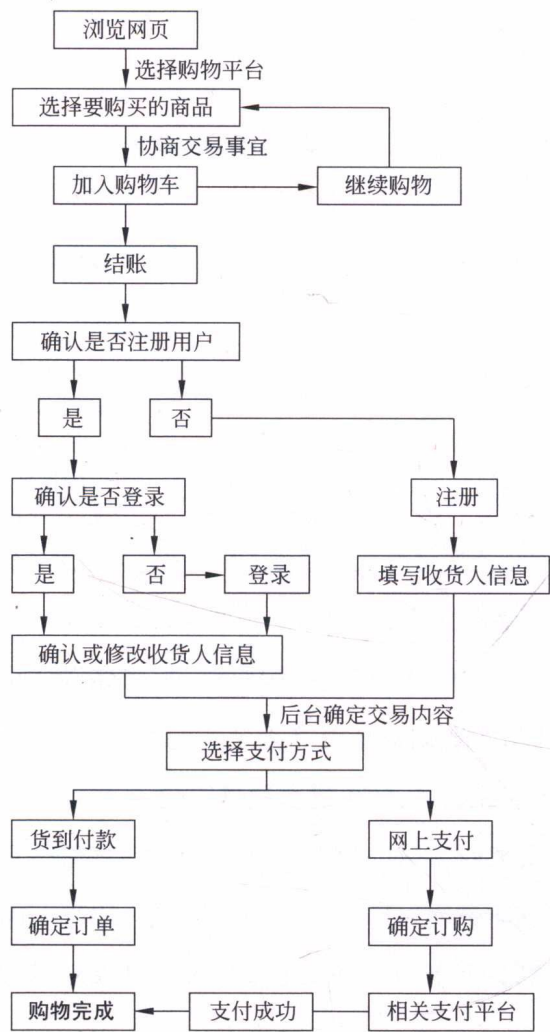


图 0.1 简化的网络购物流程图

计算机之间进行通信和数据交换而建立的所有计算机都要共同遵守的标准或约定。众所周知,互联网又称为网际网,是由各种不同的网络构成的,网络协议则是这些网络之间能够连通的保证。举个简单的例子,一位中国企业家与一位德国老板开会商谈合作事宜,如果中国企业家只懂中文,德国老板只懂德文,那他们没有办法正常交流;如果两人的秘书都懂英文,那么一方的秘

书可以先将自己上司要表达的意思转化为英文传递给对方秘书,对方秘书接收到对方信息后,再将英文转化为自己老板的母语,这样,交流就可以愉快地进行下去。在这个过程中,秘书和英语就起到了类似网络协议的作用。

一个标准的网络协议至少包括语法、语义和时序三个要素:

- 语法:规定了用户数据和控制信息的格式,包括数据出现的顺序;
- 语义:规定了各种控制信息的意义,说明通信双方应该怎么做;
- 时序:规定了事件的顺序,例如何时通信、先做什么、后做什么、传输速度等。

用上面中国企业家和德国老板开会的例子来类比的话,语法就是大家都能理解的英语语法,例如定语从句、状语从句等等;语义就是使用的英语单词和语句的意义;时序就是两位秘书事先商量好:谁先说、谁后说,先讨论什么内容、后讨论什么内容,语速是快还是慢等等。

网络协议有很多,其中最著名的莫过于 TCP/IP 协议。TCP/IP 是 Transmission Control Protocol/Internet Protocol 的简写,中文名称是传输控制协议/互联网协议。事实上,TCP/IP 协议是一个协议族;包括很多协议,例如 UDP 协议、ICMP 协议等等,TCP 协议和 IP 协议是其中最广为人知的两个协议。顾名思义,TCP 协议是为数据传输提供服务的协议,它通过一种称为“三次握手”的机制在终端之间建立连接,提供可靠的传输服务。“三次握手”机制得名的原因,是传输数据的终端之间要通过三次交互过程才能完成可靠连接的建立。有一个名为“我给你讲一个 TCP 的笑话吧”的段子生动地描绘这一过程。

路人甲:好想听一个笑话。

路人乙:你好,你想听 TCP 的笑话吗?

路人甲:嗯,我想听一个 TCP 的笑话。

路人乙:好的,我会给你讲一个 TCP 的笑话。

路人甲：好的，我会听一个 TCP 的笑话。

路人乙：你准备好听一个 TCP 的笑话吗？

路人甲：嗯，我准备好听一个 TCP 的笑话。

路人乙：Ok，那我要发 TCP 笑话了。大概有 10 秒，20 个字。

路人甲：嗯，我准备收你那个 10 秒钟长，20 个字的笑话了。

路人乙：抱歉，你的连接超时了。你好，你想听 TCP 的笑话吗？

好吧，这个笑话有点冷，我们需要知道的是网络中所有的“行为”，例如建立连接、发送/接收数据等等，都是由标准化组织制定的各种网络协议来规范的。这些网络协议不能直接搬到网络上用，还需要由码农们用机器语言把这些协议“翻译”成各种网络终端能够“看懂”的形式，也就是大家俗称的“程序”。在这些程序里，包含了各种各样不同功能的算法，正是这些算法保证了网络协议中所描述的功能能够实现，例如路由协议的实现依赖于最短路径算法，传输控制协议的实现依赖于排队算法等等。一般来说，网络协议的实现需要多个算法共同支持。有时候，网络协议本身就是一个算法，例如三次握手协议。很多时候，网络协议之间还存在复杂的依赖关系，这里就不多讲了。

是的，我们使用的每一个计算机软件、程序、APP，背后都是算法。也许我们常在新闻报道里见到“程序”“应用”“软件”“算法”等意义相近的词汇，这在媒体报道里很少加以区分。

不过，为了保证其严谨和准确性，我们在本书会使用算法，而不是“程序”或是“软件”。算法、程序、软件，是三个既有区别又有联系的词汇。算法是针对某一问题的解决方案的准确描述。例如计算机领域的排序算法描述的是，给定一系列数，按照由大到小或由小到大的顺序输出：如果输入是(5, 3, 9, 2, 6)，则输出是(2, 3, 5, 6, 9)或(9, 6, 5, 3, 2)。算法可以使用自然语言来描述，当然也可以使用计算机语言或是数学语言。算法一定要在有限的步骤中得到问题的结果，即满足有穷性。而程序就不一样了，程序不一定满足有穷性，

它可以进入无限循环的状态,例如计算机操作系统,只要不关机或者系统不遭到破坏,它就会一直运转,等待新的任务到达。此外,程序是使用编程语言写成的,用于实现某种特定目的的一组计算机指令。如果将一个算法使用编程语言进行描述,就成为一个程序,该程序是这个算法在计算机上的特定实现。软件,则是程序的有机集合体,软件中可以只有一个程序,也可以是多个程序的集合,程序从属于软件。例如手机上或计算机里最简单的计算器,就只包含数值计算程序;而 Office 办公软件则包含了 Word、PowerPoint、Excel 等多个程序。对比算法、程序、软件这三个词,可以看出算法的含义是最明确的,程序或软件只是算法的某种特定实现或这些实现的集合,而且程序或软件包含的范围更大,不如算法准确。算法描述更准确、更本质。

## 被算法唤醒的新经济

生产率不等于一切,但长期看它几乎意味着一切。

——保罗·克鲁格曼,经济学家,诺贝尔经济学奖获得者

从经济系统的角度看,赛博经济系统继承了传统经济系统的各个组成部分,并赋予了它们新的特点,形成了赛博下的生产系统、交换系统、消费系统和金融系统,如图 0.2 所示。

从上至下贯穿赛博经济系统的算法包括八类,分别是推荐算法、分配算法、匹配算法、动态定价算法、区块链技术及相关算法、大数据处理算法、数据交易算法、隐私保护算法。同时,每一类都包括针对不同具体问题的算法,以及数量不菲的侧重不同方面的变种。

由于经济系统不是孤立运行的,它必然会造成同一类算法可能既存于这