



对数字、数据和算法的若干哲学思考

李伯聪

中国科学院大学 跨学科工程研究中心, 北京 100049

摘要: 数学是形式科学。数字不是天然的客观存在, 数字的根本含义和特征在于它只突出表现对象的“(数)量”, 而不反映、不表现其“(性)质”。从算盘发展到电子计算机, 不但数字的载体发生了巨大变化, 而且在数制上也由十进制变成了二进制。数据本身也不是天然存在的, 而是通过不同方法从现实世界中采集、获取和记录而来的。电子计算机出现后, 对于中国古代数学特别重视算法的传统和体系, 可以有新认识和新评价。在数字时代, 出现和形成的一个重要的新特征和新场境是大数据、遍在算法、大算力、泛在计算的“四结合”。

关键词: 数字; 数据; 算法; 大数据

中图分类号: N91

文献标识码: A

文章编号: 1674-4969(2025)01-0032-10

引言

《大数据时代》的英文版出版于2013年,《大数据时代》^[1]的中文版也是2013年。这本书的英文版和中文版“同步”出版,这件事本身就是引人注目的“大数据时代”的具体“产物”和具体“表现”。这本书的推荐序二说:“大数据之所以可能成为一个‘时代’,在很大程度上是因为这是一个可以由社会各界广泛参与,八面出击,处处结果的社会运动,而不仅仅是少数专家学者的研究对象。”^{[1]V}正是由于大数据不仅仅是少数专家学者的研究对象,《大数据时代》一书有了广大的读者群,产生了广泛的社会影响。该

书的推荐序一说:“大数据和云计算是一个问题的两面,一个是问题,一个是解决问题的方法。”^{[1]I}与我国出现大数据热互为表里,不旋踵间,云计算和算法也成为了社会热点^[2-3],引起了许多关注^[4]。

在大数据、云计算成为热点的环境中,涌现了许多专业性、普及性、传播性的对数字、数据和算法进行科学技术、经济学、社会学、政治学分析的文章和著作。由于数字、数据和算法的发展不但涉及以上“学科领域”,而且它们也都成为了“工程生态”对象和“工程生态”问题,需要对其进行“工程生态视野”的分析和研究。“工程生态视野”的分析和研究与“具体学科视野”的分析和研究既有密切联系的方法

收稿日期: 2025-01-08; 修回日期: 2025-01-22

基金项目: 中国工程院战略研究与咨询项目“工程生态论的理论体系和实践研究”(2024-XZ-69)

作者简介: 李伯聪(1941—),男,教授,研究方向为工程哲学、工程社会学和工程史研究。E-mail: libocong@ucas.ac.cn

引用格式: 李伯聪. 对数字、数据和算法的若干哲学思考[J]. 工程研究——跨学科视野中的工程, 2025, 17(1): 32-41. DOI: 10.3724/j.issn.1674-4969.20250004. CSTR: 32282.14.JES.20250004

Li B C. Philosophical Reflections on Numbers, Data, and Algorithms[J]. Journal of Engineering Studies, 2025, 17(1): 32-41. DOI: 10.3724/j.issn.1674-4969.20250004. CSTR: 32282.14.JES.20250004

面又有“超越性”的方面。本文将在“工程生态立场和视野”这个“总立场和总视野”中又将“侧重哲学分析方法”对数字、数据和算法进行一些初步的讨论。

1 对“数字”和“数据”含义的若干认识和分析

1.1 对“数字”含义的若干认识和分析

1.1.1 人类和个人认识“数字”的历史和意义 学会计“数”是人类智力发展中的一件大事。但目前我们很难断定人类何时开始会计“数”，开始有“数字”概念。可以肯定的是古埃及、古巴比伦、古代中国在五六千年前——甚至更早时间——就会数“很大”的数字了。而直到20世纪，甚至还有某些原始部落的人只能数到3或4这样“很小”的数字，再大一点的数字他们就不会数了。这个事实，有些人也许会仅仅将其作为逸闻趣事，但细思起来，它的深刻意蕴是生动地启示人们：认识到和学会计“数”——能够数很大的“数”——是人类思维和社会发展中的一件大事。

虽然许多人也许会忽视“学会数数”在人类思想发展史上的重大意义，但大概无人会忽视个体认知成长过程中“识数”的重大意义。著名的瑞士儿童心理学家皮亚杰（Jean Piaget）更在儿童数学认知发展的理论研究中取得重要进展^[5-6]。可是由于“识数”“数数”这件事为人们所熟知，人们往往会忽视其中的意蕴和重要性，特别是可能忽视了“说话（语言）”和“数字（数学）”并不完全是一回事的深刻意蕴。

以上所说的关于“学会数数”在人类认识发展和个体认识发展中都具有很大重要性的事实告诉我们有必要进一步深入思考和研究与“数字”有关的许多深刻问题。

1.1.2 关于“数字”的“性质”的两个重要问题 在科学哲学和技术哲学中，发现和发明是两个不同的概念和过程。

由于发现和发明之间有深刻的区别，于是这就出现了与“数”有关的一个重要问题：“数”是“被人

发现出来”的还是“被人发明出来”的。

实际上，在数学哲学中，早就有学者关注了关于“数学”究竟是发现的对象还是发明的对象这个问题。有人持前一种观点，也有人持后一种观点。这里只讨论“数字”和“发现”与“发明”的关系。就“数字”而言，笔者认为“数字”是“发明”出来的，而不是“天然已经存在”而后“被人发现”的。作为例证，我们必须说作为自然界对象的煤矿是“被人发现”的，而作为人工物的蒸汽机是“被人发明”的。由于“3”这个“数字本身”——其他数字也一样——不是像“煤矿”那样的“天然存在”从而“可被发现”，我们需要认为“数”是“发明出来的”。

“数”的发明是人类思想史上的一件大事。在数学史上，在发明了“自然数”后，又发明了“负数”“实数”“虚数”等等。

在此应该特别关注的是“0”这个数字的意义和重要性。在《数，科学的语言》中，丹齐克说：“在没有发明一个表示空级的符号，表示无的符号，也即我们现代的零以前，任何进步都是不可能的。”^{[7]25}如果进一步考虑到在二进制中，只用“两个数字”——0和1——就可以表示所有的数字，则0这个数字的发明就更加意义重大了。著名的数学家和哲学家莱布尼兹在思考二进制的奇妙性质和哲学含义时更惊世骇俗地说：“用一，从无，可生万物。”^{[7]16}

与“数”有关的第二个重要问题是应该怎样认识“数字”的含义和本质特征。

从哲学角度看，“数字”的含义和本质特征在于它只突出表现“对象”的“（数）量”，而不反映、不表现“对象”的“（性）质”，因此学者们在断定数学的学科性质时认为数学是关于“形式”而不涉及“内容”的学科。

在数学和汉语语言使用中，常常需要把“数字（形式）”和“表现内容和性质”的“单位”结合起来，形成一个词组，例如15°C、15°F、15 V、15 t、15个飞行员（小学生）等等，才能形成“有内容”的词语，而单纯的15这个数字本身则是“不带内容”的“单纯”的“（数）量”。

1.1.3 对“信息”的“特性”的不同认识 由于数字的意义和数字的运用中，必然关联“信息”，这就使得

有必要讨论与“信息”含义有关的一个问题——应该怎样从“形式和内容的相互关系”的角度认识和处理“信息”的性质?

在“信息”含义的理解和解释过程中, 申农(Claude Elwood Shannon)是一个有划时代贡献的人物。申农在《通信的数学理论》中, 明确指出:“通信的基本问题是在通信的一端精确地或近似地复现另一端所挑选的信息。”申农承认“通常, 信息是有意义的”, 但他又强调“通信的语义方面的问题与工程问题是没有关系的。”^{[8]508}这就是说, 在申农的信息理论中只涉及信息的形式问题而不涉及信息的内容问题。必须承认和肯定, 这个不涉及“信息内容”的基本立场、观点和方法是成就申农理论贡献的关键所在, 而且应该承认在研究和分析通信工程问题时必须采取这个“只关注形式和信息量而不关注信息内容”的观点和方法。例如, 在通过“电报公司”“打中文电报”时, 只能按照“字数”收费, 而不能对“同样字数”的电报因“内容”不同而收取不同的费用。在信息理论其后的发展中, 许多学者关注了语义信息——也就是信息的内容——问题, 可是, 对语义信息的研究迄今仍然未能取得公认性的根本突破。

法国哲学家让-皮埃尔·迪皮伊(Jean-Pierre Dupuy)曾经乐观地说:“我们仿佛拥有了关于这个世界的越来越多的信息, 但这个世界在我们看来却缺乏意义。”“不可避免地, 意义将顽强回归。”^{[9]26}可是, 迪皮伊的“预言”没有成为现实, 直到今日, “人们还没有找到能够指导解决这个问题的理论和根本方法”^{[9]27}。

这就是说, 无论是对“数字”的解释还是对“信息”的解释中, “传统认识”都只在“形式”方面取得了“共识”而未能在如何将“形式”与“内容”解释结合起来这个问题上取得突破和形成新的“共识”。

1.2 对“数据”含义和意义的若干认识和分析

1.2.1 什么是“数据” 虽然无论从历史看还是从逻辑关系看, “数据”这个概念的产生都早于大数据, 但“数据”这个概念之被广泛关注和被赋予更重要、

更深刻的意义却是在大数据被广泛关注之后。这实在是一个耐人寻味的事情。

《数据要素五论》^[10]一书在讨论数据的定义时, 特意介绍了已经提出的关于数据的30个定义(如果更多的搜集资料, 会发现更多数量的关于“数据”的不同定义)。其最后的两个定义分别是:“数据是观察或测量结果的事实。信息^①是有意义的信息, 或是通过一定方式解释、组织, 进而产生意义的信息”“数据是以数字、事实、图片等形式反映自然与社会世界中各种现象的人类事物。信息是生物之间交流的任何对象。它和能源、材料是生命生存和进化的三大支柱”。这30个“定义”并非完全不同, 但也不是完全相同, 显示出人们对数据的含义虽然有一定的共识(否则不会普遍地使用这个术语), 但在具体认识上又意见纷纭。

1.2.2 “数据”的来源和性质 与“数字”和“信息”的传统解释往往偏重“形式”不同, 日常语言通常在使用“数据”这个词语时, 往往不但关注“形式”或“数字”的方面而且同时关注与“形式”或“数字”结合在一起的“内容”方面。

数据是从何而来的呢? 数据本身并不是天然存在的, 数据是通过不同手段、不同方法和不同技术从现实世界中采集、获取和记录而来的。这些数据可以是测量的结果、观察的记录、计算的输出等。因此, 数据是人为活动和采集工作的结果。

数据采集可以直接通过人的感觉器官和感知活动进行, 也可以通过其他方式(如传感器、问卷调查、试验等)获取原始数据。这些数据在采集后经过处理和分析, 可以用于决策、设计、预测、工程实践、运营、评估等多种目的。

在许多情况下, 采集特定的数据需要有相应的某种“感知器”(或曰“传感器”)。例如, 温度传感器可以感知和采集温度数据, 湿度传感器可以感知和采集湿度数据, 而摄像设备则可以感知和采集图像数据。这些感知器可以将物理世界中的信息转换为可以被理解和被计算机处理的数据。

从哲学角度看, 承认数据不是天然存在的而是人

① 目前人们对“信息”的定义的理解多种多样, 甚至有很大差异, 此处引文中对“信息”的理解就与前文谈到的一些观点大相径庭。

为和采集的结果，这是一个重要而影响深远的“工程哲学”和“数据哲学”论断。

以上谈到了数字、信息、数据的含义，对于它们的相互关系，一方面，必须承认三者相互渗透，常常难分难解；另一方面，又要承认三者的含义有很多区别，各有自身特点和用法，不能混为一谈。

1.2.3 大数据形成后“数据”的性质和意义的巨变

(1) 大数据的形成。辩证法认为，量变可以转化为质变。虽然“数据”早已“存在”，并且社会中的“数据”的“数量”也在不断增加，可是，在农业社会和工业社会数千年的发展进程中，“数据”的“数量变化”也仅仅是“量变性质”的变化而已。可是，在最近几十年中，“数据的量变”转化成为了“数据性质的质变”——这就迫使人们又提出了大数据这个新术语和新概念。

所谓大数据，其特点是5V: volume (大量)、velocity (高速)、variety (多样)、value (价值性)、veracity (真实性)^{[2]42}。可以说，这主要是从技术角度对“数据领域”出现大数据现象之后形成的新特点的概括和总结。

虽然技术专家可以和应该着重从“技术视野”认识和分析“大数据的5V”与“(传统的)数据的‘非5V’”的区别，但哲学家却需要着重从“哲学视野”认识和分析“大数据现象和大数据概念”与“(传统的)数据现象和数据概念”的区别。由于这里出现的区别不是“量的区别和变化”而是“质的区别和变化”，于是有些学者开始关注和研究“大数据哲学”^[2]。大数据哲学涉及的哲学问题和跨学科问题都很多，下面就略谈其中的两个重要问题。

(2) “数据”成为了哲学论域的新范畴和经济活动的新的生产要素。本文的分析不脱离有关的技术基础，但本文的分析和论述会更加着重从跨学科和哲学——特别是工程哲学——的角度分析有关问题。

上文谈到从“数据”到大数据意味着发生了“质变”，其语义不是说大数据不再“属于”“数据”，而是说大数据意味着“旧质的数据”变成了“新质的数据”。还请读者注意，本文在谈到“数据”时，有时

指前者，有时指后者，有时兼指二者。

无论从理论角度看还是从现实角度看，大数据形成后“数据”的社会性质、意义和影响都发生了惊人的巨变。

如果从跨学科视野和工程哲学角度看大数据的意义和影响，可以看出，大数据出现后，其最重要的意义之一就是它使现代人对“数据”有了新认识，使“数据”这个以往语言使用中的“普通概念(术语)”升华成为了科技、社会和哲学领域的一个具有全新含义的“新范畴”。

从哲学角度看，概念和范畴既有密切联系又有重要的区别。大体而言，范畴比概念的抽象程度更高；概念可以用于理解和解释具体事物，而范畴主要用于理解和解释整个知识体系及知识体系的结构。可以说，所有的范畴都是概念，但并非所有的概念都可以升华为范畴，就此而言，可以认为范畴是“概念体系”中的“基本概念”——特指概念体系中具有广泛、深刻的意义和影响的特别重要的概念^①。

在人类认识的发展进程中，人类的知识体系必然也要不断发展。在人类知识体系——包括哲学知识体系和各门具体学科的知识体系——演化的过程中，必然会形成新范畴。新范畴的形成，可有两种方式：一是运用提出全新术语的方法形成新范畴，一是通过对原先知识体系中的某些普通概念进行理论升华的方法形成新范畴。例如数据是原先就存在的一个普通概念(术语)，可是在新的知识体系中，它升华成为了一个(新)范畴。所谓(新)范畴不但意味着其内容、内涵有了新丰富、新发展，而且意味着其意义、功能、影响——特别是其在整个知识体系中的位置和影响——有了今非昔比的巨大变化。

在出现和形成大数据的新现象和新形势之后，(大)数据这个新范畴很快就成为了哲学界的一个研究热点。

而更加值得注意的是，在经济学领域，在生产要素的结构这个基本问题上，学者们又提出了一个新观点：数据成为了新的生产要素。

许多学者认为，以往的经济活动建立在四个生产

① 也应承认，在许多作者的用语中，重要“概念”和“范畴”的语义和语用没有重要区别。

要素的基础上: 土地、劳动力、资本和技术。现在, 数据成为了新型的和新兴的生产要素。数据要素与传统的四种生产要素相比, “其属性既有相似之处, 也有显著的差异, 并且还具有一些独特的属性。”^{[11]25}

已有许多论著关注和研究了“数据要素”的性质和特征。大体而言, “数据要素”的特性包括“非竞争性、部分排他性、虚拟性、外部性、流动性、非消耗性、非均质性、高敏感性、边际报酬递减、规模报酬递增等。”^{[11]26} 此外还有学者谈到数据的无形性、可复制性、共享性等, 它们也都是数据的不可忽视的重要性质。目前, 关于数据要素问题的研究论著几乎可谓出现井喷之势, 此不赘言。

2 数字、计算及其“载体”和数字时代

2.1 数字与数学

许多人认为, 数学是关于“数”和“形”的学科。数学的一个基本研究对象和内容就是要认识和研究“数”。

应该怎样认识“数”的“本质”和“数学”的“学科属性”呢?

许多人习惯于把数学看作“科学”的重要组成部分。现代各国的“科学院”也都把“数学家”和“数学学科”看作自身的重要组成部分之一。可是, 从“本性”和“本质”上看, 数学又与物理学、化学、生物学等“实证科学”有根本的区别。物理学和化学等学科需要通过“实验”进行“实证”, 而数学不需要并且不可能通过“实验”进行实证。欧几里德几何的平行线公理在非欧几何中不成立, 可是欧氏几何和非欧几何又都是正确的。于是, 科学哲学家普遍认为数学是形式科学。

从概念分析和概念比较角度看, “形式科学”是一个与“内容科学”相对立和对待的概念。于是, 百度·百科中就有了以下的解释: “数学是人类对事物的抽象结构与模式进行严格描述、推导的一种通用手段, 可以应用于现实世界的任何问题, 所有的数学对象本质上都是人为定义的。从这个意义上, 数学属于形式科学, 而不是自然科学。”^[12]

作为“形式科学”的数学缺乏物理学、化学、生物学那样的实证内容, 这是它的缺点; 另一方面, 正是由于它是形式科学, 加上由于形式与内容有对立统一关系, 这才使它“基于‘形式’的特性”而有可能广泛地运用于自然界和社会生活的方方面面的不同内容中, 这就使“数”和“数学”有了可以广泛运用到全部世界的各个论域和范围的可能。

2.2 数字、计算与数学

在数学的多种运用方式中, 计算是最重要的运用方式之一。在中国科学史和数学史上, 所谓数学, 在中国古代时期通常被称为“算学”。中国古代最著名的数学著作是《九章算术》^[13]。唐代李淳风又编纂了汉初到唐末一千年的数学发展中的十部著作统一定名为《算经十书》(包括《周髀算经》《九章算术》《海岛算经》《孙子算经》《张邱建算经》《五曹算经》《五经算数》《辑古算经》《缀术》和《夏侯阳算经》)^[14], 再次显示中国数学重视计算的特点。

中国古代数学体系和欧洲古代数学体系有共性的方面, 同时也各有自身的不同特点。西方的数学传统更重视数学的公理化, 而中国的数学传统更重视计算和算法问题。

数学步入现代发展阶段后, 西方数学界在计算与算法理论上取得了诸多重大的新突破^[15]。尤其是电子计算机诞生及其广泛应用, 更是创造了令世人惊叹不已的奇迹。对于电子计算机的发明及其创造的诸多奇迹, 一方面, 必须肯定其得益于数学中数字、计算及算法理论的进步; 另一方面, 也必须承认, 电子计算机的现实应用与发展, 又反过来推动了数学在数字、计算及算法理论方面的深入发展。

2.3 数字与计算的“载体”及其变化发展

为什么“形式性”的数字和计算可以利用“物质性”的电子计算机表现和实现呢? 这就需要从“数字”与计算的“载体”问题谈起了。

从符号学的角度看, 数字是一种特殊的符号(符号系统)。符号的表现需要通过一定类型的“载体”。饶有趣味并且意蕴深刻的是, 同样的符号(符号系统)可以有不同的载体(载体系统)。例如, 数字符

号的一种重要“表现形式”或“载体”是“声音”。在不同的语言中，同一个数字有不同的发音。例如数字6在汉语中发音是“liu”，在英语中是“six”。在德语、法语、西班牙语、南非的祖鲁语中，也各有不同的发音。后来，在文字发明后，数字又表现为“书面符号”，这就使“书面符号”成为了数字的新载体。与语音载体一致，在不同民族和社会生活中，数字的“书面符号”也很不同。例如，就书面符号而言，现代社会中流行的阿拉伯数字符号1、2、8、11、101，在罗马数字中写为I、II、VIII、XI、CI；在古代汉语中写为壹、贰、捌、拾壹、壹佰零壹（又：《史记》卷次写为“卷一百一”）。

更加重要的是，数字和计算还可以有其他类型和方式的“载体”。这方面的一个重要事例就是中国的算盘。在算盘这种计算工具中，算盘珠成为了数字的“载体”，运用珠算口诀可以进行许多复杂、高效的计算，在中国古代的商业活动中得到了广泛的应用。

在近代历史上，机械式计算机和计算尺也曾经有广泛的应用。在机械式计算机中用齿轮作为数字的“载体”，使用的数制是通常用的十进制。

值得特别注意的是，在发明电子计算机时，不再使用算盘珠或齿轮等机械载体代表数字，而是使用电脉冲这种电子方式代表数字，以电脉冲的有无——更准确地说是以阈值为界——分别代表数字1和数字0，这也意味着电子计算机使用的是二进制而不是通常使用的十进制。

从数学和电子技术相结合的视野看，一方面，如果不使用电脉冲这种电子方式作为数字载体，计算机就不会有高速运行的技术条件。维纳在《控制论》的“导言”中说，他很早就认识到为了加快运算速度，计算机就必须用电子管而不是用齿轮^[16]；另一方面，如果没有关于二进制的数学理论，电脉冲的有和无也仅仅是电子学的电脉冲，而不可能成为数字的表示方式（“载体”）。

在电子计算机和同时期的一系列新技术大显身手的环境条件下，许多人愈来愈深刻地意识到人类社会已经从工业社会转变到数字社会，从工业时代转变到数字时代。于是，讨论数字社会和数字时代的论著——例如《数字社会学》^[17]——也开始多了起来。

3 数字时代中大数据、遍在算法、大算力、泛在计算的“四结合”

在从工业时代到数字时代的变革和转型的过程中，必然出现和形成一些前所未有的新现象和新场境，而大数据、遍在算法、大算力、泛在计算的“四结合”就是其中的一个新现象和新场境。

3.1 作为传统概念和一般概念的算法、计算、算力

正像数据不是一个新概念一样，算法、计算和算力也不是新概念。虽然它们是早就提出并得到广泛使用的传统概念，但它们又是今后还要继续使用和含义会不断更新和与时俱进的概念。

什么是算法呢？百度·百科说：“算法（algorithm）是指解题方案的准确而完整的描述，是一系列解决问题的清晰指令”“算法中的指令描述的是一个计算，当其运行时能从一个初始状态和（可能为空的）初始输入开始，经过一系列有限而清晰定义的状态，最终产生输出并停止于一个终态。”“一个算法应该具有以下五个重要的特征：有穷性、确切性、输入项、输出项、可行性。”^[18]

算法这个中文名称出自《周髀算经》^[19]，而英文名称 algorithm 来自于9世纪波斯数学家 al-Khwarizmi（因为 al-Khwarizmi 在数学上提出了算法这个概念），而“algorism”的意思就是指阿拉伯数字的运算法则，在18世纪演变为“algorithm”。

中国著名数学家吴文俊认为^[20]，世界数学史的发展中有两大传统：一是以《几何原本》为代表的公理化、逻辑化、演绎化传统；一是以《九章算术》为代表的算法化、程序化、机械化传统。在相当长的一段时间里，众多数学史家与科学家均“独尊”公理化、逻辑化及演绎化的数学传统，而轻视乃至忽视了算法化、程序化、机械化传统的价值与意义。然而，随着现代电子计算机的横空出世，以其展现出的强大能力为事实依据，人们不得不重新审视并评价算法化与程序化的数学传统。

算法的计算对象是数据。从哲学角度看，应该怎样认识数据与算法的关系呢？可以将其类比为语言活

动中词汇和语法的关系。在语言活动中, 词汇是语言的基础“元素”, 语法是组织和使用词汇的规则, 它规定了词汇如何组合成句子、段落和篇章。对于语言表达和传达来说, 词汇和语法缺一不可。口头语言是说话, 书面语言是写文章。一方面, 没有词汇就不可能有口头语言和书面语言; 另一方面, 没有语法也不可能有口头语言和书面语言。只有词汇和语法的密切结合才能完成语言表达和传达的任务。同样地, 只有在数据和算法相互配合和密切结合的条件和情况下, 才能完成计算和解决数学问题的任务。

计算牵涉到数据、算法和可计算性问题^{[15]168,190}。有些问题是所谓的“简单问题”, 例如购买两件小商品时涉及的交易价格问题, 这类问题的数据相对直观, 解决问题所需的算法也较为简易。而另一类则是“复杂问题”, 比如要求预测明天天气情况的天气预报问题, 这些问题所涉及的数据和算法, 相较于前者, 展现出了无可比拟的复杂性。在认识和对比这两类不同的问题时, 算力概念就自然而然地浮出水面了。前者涉及的算力较小, 而后者就涉及更大和更复杂的算力了。

所谓计算, 既可指针对某个具体问题的直接计算——即具体实例(或称事例)的详细运算, 亦可泛指众多具体计算的总称。在这两种含义中, 均会涉及数据、算法和算力: 在“具体计算”的层面, 它关联的是“具体数据”“具体算法”及“具体算力”; 同时, 计算也可超越具体情境, 在“一般”或“统称”的语境下探讨“数据”“算法”与“算力”。

3.2 算法、计算、算力发展的新形势和新阶段

3.2.1 不同时代的算法、算力的能力、范围和特点 在人类历史发展进程中, 数据、算法、计算、算力的表现形式、内容意义和反映对象的范围大小、场境特征都在不断变化。以下仅简要讨论古代算法的一些问题。

《九章算术》是最重要、最著名的中国古代数学经典, 大约成书在公元前1世纪的西汉时期^[13]。与突出“公理化”和“数学证明”的欧洲数学传统不同, 《九章算术》这个书名已经突出强调中国数学传统的突出特点是研究“数学算法”和“数学计算”问题。更加值得注意的是, 《九章算术》专注于探讨并解答

了“九个问题类型(或称领域)”的数学计算问题。这九个类型依次为: 方田、粟米、衰分、少广、商功、均输、盈不足、方程、勾股。从它们的命名中不难窥见: 《九章算术》既重视“纯数学问题”(如以数学术语为核心命名的勾股)的求解, 又尤为关注那些“与现实问题紧密结合的数学问题”(如以数学方法应用于社会现实为背景命名的方田、粟米、均输)。值得注意的是, 后者在数量上占据多数; 并且, 全书的核心思想在于强调数学方法与社会现实问题的紧密联系, 而非单纯进行“纯数学”的“问题设定”或“算法探索”。

《九章算术》的体例是对于各类数学问题先给出几个具体的例题, 然后给出解答问题的一般性算法。虽然书中对具体例题只给出了“问题来源数据(数字)”和“答案数据(数字)”而没有给出“具体计算过程”, 但人们可以根据随后的“一般性算法”进行具体计算, 得出具体答案。由于具体实例的数字可大可小, 有些小数字的计算可以采取心算方式, 而巨大数字的计算就必须使用算筹或算盘计算, 因为所需要的算力超出了心算能力的界限。

3.2.2 算法、计算、算力发展的新阶段

(1) 泛在计算的概念、内容和意义。最近时期, 科技界和理论界不但对大数据的讨论热闹非凡, 而且对云计算的讨论也热闹非凡。而更加令人惊讶的是, 云计算的热度还没有“退潮”, 有人又提出了“边缘计算”“‘云—边缘—端’一体化”乃至“泛在计算”的概念。

在云计算技术中, “通常把数据传输到云计算中心加以处理, 不过随着5G与人工智能时代的到来, 数据量爆发式增长, 对网络时延、数据安全性、可控性等, 提出了极高的要求, 为满足新的需求, 边缘计算应运而生。”^{[21]144}边缘计算“将原有云计算中心的部分或全部技术任务迁移到数据源附近, 具备实时数据处理和分析、安全性高、隐私保护、可扩展性强、位置感知、低流量等优势。”^{[21]144}

与边缘计算几乎同步发展, 又有了“‘云—边缘—端’一体化”技术。“云计算从中心向边缘扩展, 边缘计算实现了新的突破, 技术架构出现变革, 技术架构由原来的云一端, 变成了中心云—边缘云—端设备

协调工作的模式。”^{[21]150}

从工程哲学角度看,这种云—边缘—端一体化的现象和趋势实质上是形成了“泛在计算”^①。有人认为,泛在计算即是信息空间与物理空间的深度融合。在这一融合的空间里,人们能够随时随地、无缝且透明地获取数字化的服务^{[21]155}。

(2) 大数据、遍在算法、大算力、泛在计算“四结合”的到来。这里无意对泛在计算进行更多的哲学分析和解释。这里只想指出另外一个要点:与泛在计算密切联系的还需要有另外一个新概念——遍在算法。

从哲学分析的角度审视,要解决计算问题并得出答案,关键在于找到既可信又可行的算法。这里的算法,不仅局限于纯数学问题的求解方法,更广泛地涵盖了各类现实问题在数学上的表现形式,诸如购物结算、天气预报、城市交通调度、交通工具路线的设计与选择,以及自动驾驶等领域的算法问题。

在《九章算术》的时代,仅有极少数问题拥有明确的解题算法。彼时,绝大多数问题尚缺乏系统的解答算法,我们不妨将那个时代称作“零星算法孤岛时代”。时至今日,得益于现代科学技术与工程的飞速发展,千千万万乃至难以计数的问题都已找到了各自的算法。据此,我们可以将现今的时代称为“遍在算法时代”如果说在“零星的算法孤岛”时代,仅有少数人的少数问题需要通过有限的算法进行计算,那么,在当今这个电子计算机、大数据、云计算、智能手机、5G通信及工业互联网蓬勃发展的时代,尤其是人工智能、机器人技术、自动驾驶和工业软件大行其道的当下,几乎每个人、每个企业都面临着——或至少“应当面临”——需要进行计算的现实问题。为此,他们需要相应的算法(可称之为“遍在算法”)和算力支持,以便将数据、算法、算力及计算能力融合起来,共同应对和解决各式各样的现实问题。

如果把当前的时代和《九章算术》的时代进行对比,实在是出现了天壤之别。可以认为,当前的时代正在进入大数据、遍在算法、大算力、泛在计算“四

结合”的新时代!

4 结论

本文通过对数字、数据、算法及其在工程生态中的作用的深入探讨,系统阐述了这些要素在数字时代背景下的新特征和新意义。

数学作为形式科学,其数字的根本含义在于突出表现对象的“量”而不反映其“质”,而数据则是通过不同方法从现实世界中采集、获取和记录而来的信息载体。在数字时代,大数据、遍在算法、大算力、泛在计算的“四结合”成为了重要的新特征和新场景。

而在数字、数据与算法的哲学含义及其相互关系方面,数字是人为发明的符号系统,用于量化表示对象;数据则是通过感知、测量等手段获取的事实记录,具有形式与内容的双重属性;算法则是解决问题的一系列清晰指令和运算法则,与数据密切配合完成计算任务。这些要素在工程生态中相互交织,共同推动着工程活动的智能化和高效化。大数据形成后数据的性质和意义发生的巨变。大数据不仅改变了数据的量变到质变的特征,还使数据成为了哲学论域的新范畴和经济活动的新生产要素。这种变化对工程生态产生了深远影响,推动了工程管理的创新和发展。

数字与计算均需依托于特定的“载体”,且“载体”的差异可能对“计算速度”及“算力”产生极大的“反作用”,这一点不容忽视。以“算盘”作为“十进制数字与计算”的载体,相较于以“晶体管电子计算机”作为“二进制数字与计算”的载体,两者之间存在根本性的差异。数字与计算的载体及其变化发展,从算盘到电子计算机,再到云计算和泛在计算,这些技术变革不仅提高了计算速度和效率,还拓展了计算的应用范围和场景。特别是泛在计算的概念,将信息空间与物理空间深度融合,为人们提供了随时随地、无缝透明的数字化服务。

综上所述,数字、数据、算法不但涉及许多哲学问题,而且涉及许多工程科学、工程技术、企业、行业、经济、社会、伦理、政治等方面的问题。正是由

① 20世纪80年代,迈克·威士(Mark Weiser)博士提出了“泛在计算”这个概念^{[21]154}。

于数字、数据、算法都是“强渗透性”和“标志性突出”的问题,这就使得它们成为了“工程生态论”研究的重要内容。本文仅涉及了对“数字、数据、算法的特征”及其与“工程生态其他因素和内容”相互关系的部分内容的分析和讨论,限于笔者的学养,难免肤浅,进一步的讨论当俟诸他日。

参考文献

- [1] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶. 大数据时代: 生活、工作与思维的大变革[M]. 盛杨燕, 周涛, 译. 杭州: 浙江人民出版社, 2013.
Mayer-Schönberger V, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work, and Think[M]. London: Hachette, 2013.
- [2] 董春雨, 薛永红. 大数据哲学: 从机器崛起到认识方法的变革[M]. 北京: 中国社会科学出版社, 2021.
Dong C Y, Xue Y H. Philosophy of Big Data: From the Rise of Machines to the Change of Cognition Methods[M]. Beijing: China Social Sciences Press, 2021.
- [3] 涂子沛. 数据之巅: 大数据革命, 历史、现实与未来[N]. 北京: 中信出版集团, 2021.
Xu Z P. Top of Data: Big Data Revolution, History, Reality and Future[J]. Beijing: CITIC Publishing Group, 2021.
- [4] 李伦. 数据伦理与算法伦理[M]. 北京: 科学出版社, 2019.
Li L. Data Ethics and Algorithm Ethics[M]. Beijing: Science Press, 2019.
- [5] 皮亚杰, 英海尔德. 儿童心理学[M]. 吴福元, 译. 北京: 商务印书馆, 1980.
Piaget J, Inhelder B. The Psychology of the Child[M]. New York: Basic Books, 1969.
- [6] 皮亚杰. 发生认识论原理[M]. 王宪钊, 等, 译. 北京: 商务印书馆, 1981.
Piaget J. The Principles of Genetic Epistemology[M]. New York: Basic Books, 1970.
- [7] 丹齐克. 数: 科学的语言[M]. 苏仲湘, 译. 北京: 商务印书馆, 1985.
Dantzig T. Number: The Language of Science[M]. New York: Macmillan, 1930.
- [8] 庞元正, 李建华. 系统论 控制论 信息论 经典文献选编[M]. 北京: 求实出版社, 1989.
Pang Y Z, Li J H. Selected Classic Documents of Information Theory of System Theory Cybernetics[M]. Beijing: Qiu Shi Publishing House, 1989.
- [9] 杨志刚. 分析信息: 香农、维特根斯坦、图灵和乔姆斯基对信息的两次分离[M]. 北京: 人民邮电出版社, 2021.
Yang Z G. Analyzing Information: Two Separation of Information by Shannon, Wittgenstein, Turing and Chomsky[M]. Beijing: Posts & Telecom Press, 2021.
- [10] 张平文, 邱泽奇. 数据要素五论: 信息、权属、价值、安全、交易[M]. 北京: 北京大学出版社, 2022.
Zhang P W, Qiu Z Q. Five Views On Data Elements: Information, Ownership, Value, Security and Transaction[M]. Beijing: Peking University Press, 2022.
- [11] 戎珂, 陆志鹏. 数据要素论[M]. 北京: 人民出版社, 2022.
Rong K, Lu Z P. The Theory of Data[M]. Beijing: People's Publishing House, 2022.
- [12] 百度·百科. 数学[EB/OL]. [2024-12-12]. <https://baike.baidu.com/item/%E6%95%B0%E5%AD%A6/107037>.
Baidu Baike. Mathematics[EB/OL]. [2024-12-12]. <https://baike.baidu.com/item/%E6%95%B0%E5%AD%A6/107037>.
- [13] 郭书春. 中国古典数学的重要典籍《九章算数》[N]. 中国社会科学报, 2024-03-14.
Guo S C. Jiuzhang Suanshu, an important classic of Chinese classical mathematics[N]. China Social Sciences Today, 2024-03-14.
- [14] 中华典藏. 算经十书[EB/OL]. [2024-12-12]. <https://www.zhonghuadiancang.com/shudan/suanjingshishu/>.
Chinese Classics Collection. Suanjin Shishu[EB/OL]. [2024-12-12]. <https://www.zhonghuadiancang.com/shudan/suanjingshishu/>.
- [15] 吉尔·多维克. 计算进化史: 改变数学的命运[M]. 劳佳, 译. 北京: 人民邮电出版社, 2017.
Dowek G. L'histoire de l'évolution du calcul[M]. Paris: Le Pommier, 2011.
- [16] 维纳. 控制论[M]. 郝季仁, 译. 北京: 商务印书馆, 1961.
Wiener N. Cybernetics[M]. New York: John Wiley & Sons, 1948.
- [17] 凯普敦. 数字社会学[M]. 王明玉, 译. 上海: 上海人民出版社, 2023.
Lupton D. Digital Sociology[M]. New York: Routledge, 2020.
- [18] 百度·百科. 算法[EB/OL]. [2024-12-12]. <https://baike.baidu.com/item/%E7%AE%97%E6%B3%95/209025>.
Baidu Baike. Algorithm[EB/OL]. [2024-12-12]. <https://baike.baidu.com/item/%E7%AE%97%E6%B3%95/209025>.

- [19] 中华典藏. 周髀算经[EB/OL]. [2024-12-12]. <https://www.zhonghuadiancang.com/tianwendili/zhoubisuanjing/>.
Chinese Classics Collection. *Zhoubi Suanjing*[EB/OL]. [2024-12-12]. <https://www.zhonghuadiancang.com/tianwendili/zhoubisuanjing/>.
- [20] 吴文俊. 对中国传统数学的再认识[J]. 百科知识, 1980(7/8).
Wu W J. Re-understanding traditional Chinese mathematics[J]. *Encyclopedic Knowledge*, 1980(7/8).
- [21] 邢庆科. 算力: 数字经济的新引擎[M]. 北京: 北京大学出版社, 2022.
Xing Q K. *Computing Power*[M]. Beijing: Peking University Press, 2022.

Philosophical Reflections on Numbers, Data, and Algorithms

Li Bocong

Center for Engineering Studies, University of Chinese Academy of Sciences, Beijing 100049, China

Highlights

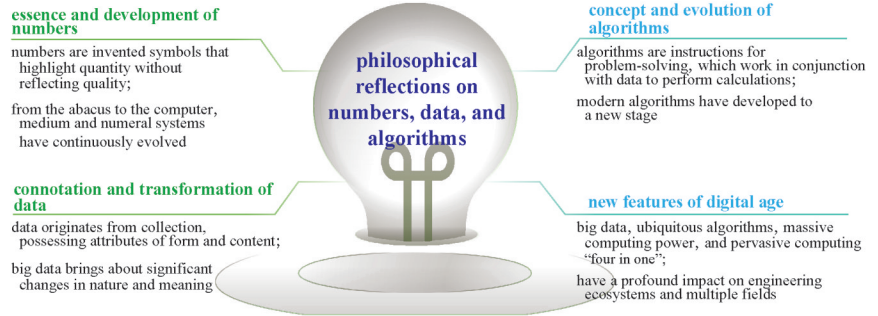
Point out numbers are human-invented to represent quantity, innovating understanding.

Analyze data's transformation into a crucial factor, highlighting big data's impact.

Re-evaluate the modern significance of the ancient algorithm tradition.

Summarize the "four combinations" in the digital age: big data, ubiquitous algorithms, massive computing power and pervasive computing, and their profound influence.

Graphical Abstract



Abstract: This paper explores the philosophical implications of numbers, data, and algorithms, and examines their roles within the field of engineering ecology. Numbers are recognized as a symbolic system devised by humans to quantify the "quantity" of objects, without inherently reflecting their "quality". Data, on the other hand, are seen as information carriers that are collected, acquired, and recorded from the real world through various methods, possessing both form and content attributes.

With the advent of electronic computers, the representation and manipulation of numbers and data have undergone significant transformations. From the abacus to the electronic computer, and subsequently to big data and cloud computing, these technological advancements have marked the dawn of the digital age.

In the digital era, the "four combinations" of big data, ubiquitous algorithms, immense computing power, and pervasive computing have emerged as key features. This paper analyzes the historical evolution of algorithms, computation, and computing power, noting that pervasive computing integrates information space with physical space, providing seamless and transparent digital services anytime, anywhere. Furthermore, the paper acknowledges the importance of algorithms and computation in ancient Chinese mathematics and discusses their relevance in contemporary society.

The paper also examines the interrelationships and roles of numbers, data, and algorithms within engineering ecology. The emergence of big data has not only facilitated the transition from quantitative to qualitative data analysis but has also elevated data to a new philosophical category and a crucial production factor in economic activities. These changes have had profound impacts on engineering ecology, driving innovation and development in engineering management.

In conclusion, the paper summarizes the new significance of numbers, data, and algorithms in the digital era, highlighting their relevance not only in philosophical discussions but also in engineering science, technology, economics, society, and other fields. These elements are intertwined, collectively contributing to social progress and development.

Keywords: numbers; data; algorithm; big data