

一种使能技术：让计算系统具备反思能力

从多学科的角度看标签化冯·诺伊曼体系结构的必要性

刘宇航¹ 刘宗惠²

¹中国科学院计算技术研究所

²武汉理工大学

关键词：计算机系 自动控制系

问题的提出

在计算系统中，数据访问或服务请求的高度并发性 (concurrency)，往往伴随着令人难以接受的性能或延迟的不确定性 (uncertainty)。此现象在数据中心这样的大规模分布式系统中十分常见，在多核或众核处理器组成的单个计算节点中也经常发生。

部分与整体之间既是对立的，也是统一的。当数据中心作为一个整体时，可以看作是由多个计算节点互连构成的计算机网络；在作为组成单元的每个计算节点内部，也是一种类型的网络。所以，一方面，数据中心是网络，计算节点也是网络；另一方面，数据中心是系统，计算节点也是系统。传统的计算机节点之间的网络上的数据包都是带有标签的，而多核或众核处理器内部的网络上的数据包是否也应该带有标签，成为了一个值得研究的问题。

中科院计算所从工业界数据中心存在的问题出发，为多核或众核处理器芯片提出了“标签化冯·诺伊曼体系结构”(Labeled Von Neumann Architecture)^[4-6]，在工程上取得了抑制性能不确定性的显著效果。“标签化冯·诺伊曼体系结构”在经典冯·诺伊曼体系结构的基础上有了改进和增强，主要体现在三个方面：一是区分 (Distinguish) 不同数据访问的处理器来源；二是隔离 (Isolate) 不同的处理器；三是根据

轻重缓急优先化 (Prioritize) 不同的处理器。这三种能力被简称为 DIP，或者说，DIP 是计算系统的一种使能技术。

如图 1 所示，我们对系统 S 进行了观察和控制 (区分、隔离、优先化)，这时的系统成为了一个新的系统 S'，并且这个过程可以一直进行下去，依次生成 S'' 乃至更多更强的系统。我们发现，这种结构具有必然性，背后的深刻逻辑值得深思，并且与经济学原理、递归、数理逻辑学中的悖论、作为数学基础的元数学、哲学中的反思、系统科学中的熵、物理学中的守恒定律都有联系。

本文将从多学科的角度探讨标签化冯·诺伊曼体系结构的必要性，在写作方法上借鉴了艾伦·图灵 (Alan Turing) 在 1950 年发表的《计算机器与智能》(Computing Machinery and Intelligence) 一文^[1]。图

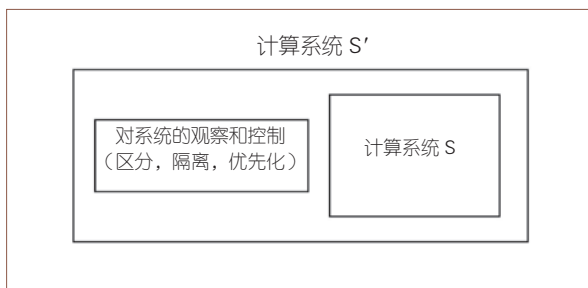


图 1 标签化冯·诺伊曼体系结构

灵在该文中介绍、分析并反驳了来自9个方面的对立观点，涉及了神学、物理、化学、生物、信息论、数理逻辑等多个学科，但都紧扣“机器是否能够思考”这一主题。可以看出，图灵能够游刃有余地驾驭多学科的知识。

目前从多学科角度探讨问题的文章还不多。今年国家制定了《加强“从0到1”基础研究工作方案》，旨在充分发挥基础研究对科技创新源头的供给和引领作用，解决我国的基础研究缺少“从0到1”原创性成果的问题。我们希望本文能起到抛砖引玉和启迪思维的作用。

从生理学和经济学的角度看

图灵在《计算机与智能》一文中提到了海伦·凯勒(Helen Keller)。海伦·凯勒是一位自幼年就失去听力和视力的人，但仍然具有学习的能力，并取得了远远超出常人的卓越成就，这一现象引起了图灵的重视。海伦·凯勒在《假如给我三天光明》一文中说过这样一段话：“有视觉的人，他们的眼睛不久便习惯了周围事物的常规，他们实际上仅仅注意令人惊奇的和壮观的事物。然而，即使他们观看最壮丽的奇观，眼睛都是懒洋洋的。”这一段话蕴含的原理是：对人来说，注意力、好奇心、兴趣是一种稀缺资源，这可以从生理学和经济学两个角度以及它们的结合点分别解析。

从生理学的角度看，如果一个人长时间处于对很多事物充满好奇的状态，这个人很快就会疲劳。更具体地说，当越来越多的信息涌入他的头脑后，大脑内的多巴胺等化学物质将快速地消耗殆尽，也就是说，他之前赖以兴奋好奇的物质基础没有了或减少了，神经系统会处于一种抑制的状态。

从经济学的角度看，当代经济学家薛兆丰教授撰写的《经济学讲义》^[10] 在其中的第9和10讲中提到：经济学大厦建立在“稀缺”（不是“人本理性”，也不是“人本自私”）的基础上才最稳固。由于资源总是稀缺的，当人们在利用有限资源的时候，就不得不对资源的用途进行选择；而每当要做选择时，

都必须采取某种选择的标准；一旦确定了选择标准，就意味着区别对待，而区别对待就是歧视。所以，稀缺、选择、区别对待和歧视这四个概念，其实是一体的，只要有一个，就会同时有其他三个存在。

若把上面两个角度结合起来，可以得出以下结论：由于用以支撑好奇心的多巴胺、生物电等物质资源是稀缺的，人们必须对这些资源的用途做出选择；而每当要做选择时，都必须采取某种选择的标准，就像海伦·凯勒所说的，“眼睛实际上仅仅注意令人惊奇的和壮观的事物”，而眼睛对于其他事物就歧视了，即“眼睛对这些事物是懒洋洋的”。

“区别对待”也是“标签化冯·诺伊曼体系结构”的思想：由于数据中心的资源（包括处理器、存储、输入输出等）是有限的，面对高并发的用户请求时，有限的资源不能同时满足这些并发的请求，从而出现资源稀缺的状况，计算机系统就必须对这些资源的用途做出选择；而每当要做选择时，都必须采取某种选择的标准，优先满足更重要、更紧急任务的资源需求，此时计算机系统对于其他任务就采取了“歧视”的态度，即计算机系统对这些任务是以低优先级的方式响应的。

从数理逻辑的角度看

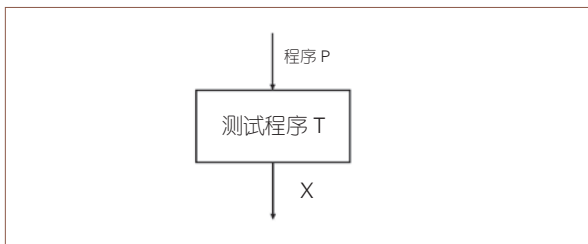
图灵在《计算机与智能》一文中提到机器无法回答以下类型的问题：“考虑有以下特点的机器……这台机器会不会对任何问题做出‘是’的回答？这里省略的是对某台标准形式机器的描述。如果所描述的机器与那台被提问的机器具有某些相对简单的联系，那么我们就可以知道，答案不是错了，就是没有答案。”

罗素悖论(Russell's paradox)是由罗素发现的一个集合论悖论，其基本思想是：对于任意一个集合A，A要么是自身的元素，即 $A \in A$ ；A要么不是自身的元素，即 $A \notin A$ 。根据康托尔集合论(Cantorian set theory)的概括原则，可将所有不是自身元素的集合构成一个集合S，即 $S = \{x: x \notin x\}$ 。也就是说，罗素构造了一个集合S：S由一切不属于自身的集

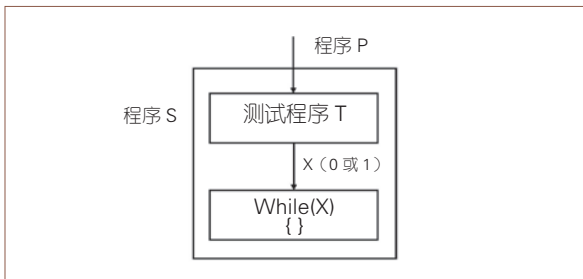
合所组成。那么 S 是否属于 S 呢？根据排中律¹，一个元素或者属于某个集合，或者不属于某个集合。因此，对于一个给定集合，是否属于它自己这样的问题是有意义的。然而，如何回答这个看似合理的问题却令人陷入两难境地。因为如果 S 属于 S，根据 S 的定义，S 就不属于 S；反之，如果 S 不属于 S，同样根据定义，S 就属于 S，无论如何都是矛盾的。

著名的“理发师悖论”也存在自己无法论述自己的矛盾。有一位理发师说：“我将为也只为本城所有不给自己理发的人理发。”那么，他给不给自己理发？如果不给自己理发，他就属于“不给自己理发的人”，他就要给自己理发；而如果他给自己理发，他又属于“给自己理发的人”，他就不该给自己理发。类似地，还有书目悖论：一个图书馆编纂了一本书名词典，它需要列出这个图书馆里所有不列出自己书名的书。那么，它究竟列不列出自己的书名？

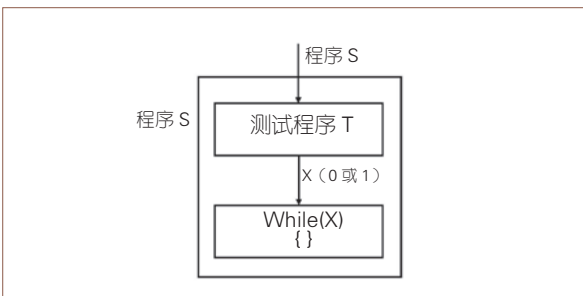
接下来，我们来举一个计算机无法计算或判定的例子——停机问题。对于任意的图灵机和输入，是否存在一个算法，可以判定图灵机在接收初始输入后可达停机状态？若能找到这种算法，停机问题就可解；若无法找到该算法，则不可解，停机问题的答案就是否定的。这个问题可以用反证法证明，先假设存在这样的测试程序，然后再构造一个程序，该测试程序不能测试。假设存在测试程序 T，当输入程序 P 能终止时，输出 x=1，当输入程序 P 不能终止时，输出 x=0。



我们总能构造这样一个测试程序 S，当 P 终止时，S 不终止；当 P 不终止时，S 终止。



这样如果我们把 S 当作输入（S 本身的输入），就会得出一个悖论：当 S 能终止时，S 不终止；当 S 不能终止时，S 终止。



从编码学的角度看

标签化的思想对于复杂系统具有基本意义，也就是说，标签对于涉及数量众多的对象的复杂系统既是必须的，也是可能的。欧内斯特·内格尔 (Ernest Nagel) 和詹姆士·R·纽曼 (James R. Newman) 在《哥德尔证明》(Godel's Proof)^[12] 中对哥德尔不完备定理的原文进行了解析，在中译本的第 56 页，有这样一段话：

首先，哥德尔表明给每一个原始符号、每一个公式（或符号串）以及每一个证明（公式的有限序列）都指定一个独一无二的数是可能的。这个数可以看作是一种区别用的标签，被称为符号、公式或证明的“哥德尔数”。

注意到这段话中提到的“独一无二”“区别用的标签”。在哥德尔证明中，哥德尔数是一种区分逻

¹ 排中律 (law of excluded middle) 是形式逻辑的基本规律之一，指同一个思维过程中，两个相互矛盾的思想不能同假，也不能同真，必有一真一假。

辑系统中众多对象的工具，没有哥德尔数，整个形式化证明就无法展开。同样，人们的身份证号、学生的学号以及超级市场中商品的条形码等，都是起到了标签的作用。

从元数据和元数学的角度看

元数据(metadata)的意思是描述数据的数据(data about data)，是对数据的观察、分析和归纳。形而上学(metaphysics)的含义是指超越物理学之上的，即“元物理”。《易经·系辞》中就有“形而上者谓之道，形而下者谓之器”。“元数学”(metamathematics)是一种将数学作为人类意识和文化客体的科学思维或知识，即一种用来研究数学和数学哲学的数学。将元数据进一步递归，可以有元元数据，并以此类推。具体地，为了表示事物的运动规律，有速度、加速度、加加速度；为了表示事物的不确定性，有熵、超熵、超超熵^[9]，李德毅院士据此提出了基于熵的“不确定性人工智能”理论。信息、知识、预测、洞察、智能、智慧，是逐级递升的元数据。

系统因为自己能力的局限，往往在讨论系统自身的时候，就出现了悖论那种纠缠不清的问题。“当局者迷，旁观者清”“不识庐山真面目，只缘身在此山中”，“剧中人”往往没有“剧作者”那样全面地掌握信息，因而不能像“剧作者”那样清醒和自觉。所以，我们需要在系统外部对系统做一些观测，这就是“元××”的思想，例如对于数学的外部观测就是“元数学”，对计算机数据访问的外部观测就是“元数据”等，这也是标签化冯·诺伊曼体系结构DIP的三种能力从“元数据”的角度来说十分有必要的理由。

“元××”这个可递归的重要概念在从数据、信息、知识直到智慧的逐级不断抽象过程中具有重要作用。根据丘奇(Church)论题，递归函数就是可计算函数，递归在定义自己时又使用了自己。内格

尔和纽曼在《哥德尔证明》^[12](在中译本的第70页)有这样一段话：

正如我们将要看到的那样，将一个字符串的哥德尔数带入这个字符串本身(然后取结果的哥德尔数)，这个看似绕圈子的概念是哥德尔关键性的想法之一，他也付出了极大的努力来使读者确信这个函数是直接可计算的，因而是原始递归的，并在对应引理的适用范围之内。我们将用记号‘sub(x,17,x)’来代表新哥德尔数，它是老哥德尔数x的函数。尽管说起来有些饶舌，我们还是可以准确地讲出这个数是什么：它是取一个哥德尔数是x的公式，其中凡是有变量y出现的地方均用x的数字替换而得到的新公式的哥德尔数。

我们可以看到，当命题自己讨论自己的时候，这个命题不可证，也就是既不可证实，也不可证伪。为了让系统具有强大的能力，系统需要来自外部的观察和控制，也是标签化冯·诺伊曼体系结构所遵循的思想。

从系统科学的角度看

系统的封闭性是系统能力有限性的根源。例如，合外力、合力矩为0，这都是系统封闭性的体现。在合外力为0时，动量是守恒的；在合力矩为0时，角动量是守恒的。标签化资源管理、性能优化都会产生“合外力”或“合力矩”，让系统远离平衡态，这也是普里高津耗散系统²的思想。

我们以尺规作图来说明封闭系统能力的局限性。尺规作图起源于古希腊的数学课题，是指仅使用无刻度的直尺和圆规作图，并且只准许使用有限次，来解决不同的平面几何作图问题。尺规作图不能问题就是不可能用尺规作图完成的作图问题，其中最著名的是被称为几何三大问题的古典难题：(1)三等分角问题：三等分一个任意角；(2)倍立方问题：作一个立方体，使它的体积是已知立方体体积

² 伊里亚·普利高津(Ilya Prigogine)于1969年提出耗散结构理论：在远离平衡的非线性区形成的新的稳定的宏观有序结构，由于需要不断与外界交换物质或能力才能维持，因此称之为耗散结构。

的两倍；(3)化圆为方问题：作一个正方形，使它的面积等于已知圆的面积。2400年前的古希腊已提出了这些问题，直至1837年，法国数学家万芝尔(Wantzel)才首先证明“三等分角”和“倍立方”为尺规作图不能问题。1882年，当德国数学家林德曼(Lindemann)证明了 π 是超越数后，“化圆为方”也被证明为尺规作图不能问题。钱学森也在1990年撰文，特别强调了系统开放性的重要性^[11]。

高熵对应的是“具有很多微观态的宏观态”，低熵对应的是“具有很少微观态的宏观态”。如“曲高和寡”“学如逆水行舟，不进则退”“由俭入奢易，由奢入俭难”这些词句，都蕴含着这样的思想。克劳修斯熵、玻尔兹曼熵、香农熵、云计算熵等，虽然形式不同，但是道理都是相通的。在数据中心中，有限的资源像一个公共的池子(有时被称为资源池)一样被很多请求使用。在没有任何控制的情况下，这些请求相互争用资源，对应着很多微观态，因此属于高熵；而我们期望的是达到一种“请求的轻重缓急得到充分兼顾、各得其所”的宏观状态，这种宏观状态对应的微观态是很少的，因此属于低熵。

系统的内与外有很重要的分别。系统的整体具有系统的各个部分所不具有的性质，是系统的整体性；系统的内部与外部的区别是系统的边界性。系统的边界性隶属于系统的整体性。系统的整体性是系统的灵魂。当系统的整体性对内或对外有一个消失的时候，系统就消亡了。人是一个系统，计算机同样是一个系统。人的整体性对内表现为，人具有各个器官单独存在时所不具有的性质；对外表现为，人与环境有清晰的界限。

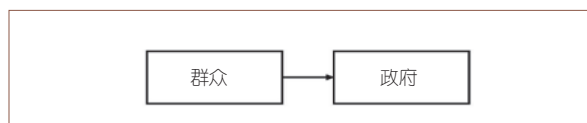
从社会学的角度看

“社会计算”是社会学与计算机科学的交叉学

科，我们需要研究客观规律，促进国家治理体系和治理能力现代化。在社会学中，“历史周期律”是一个著名的论题。1945年7月，民主人士黄炎培到延安考察，向毛主席提出了一个问题，原话是这样说的^[2]：“……真所谓‘其兴也勃焉’，‘其亡也忽焉’，一人，一家，一团体，一地方，乃至一国，不少单位都没有能跳出这周期律的支配力……有的因为历时长久，自然地惰性发作，由少数演为多数，到风气养成，虽有大力，无法扭转，并且无法补救³。也有为了区域一步步扩大了，它的扩大，有的出于自然发展，有的为功业欲所驱使，强求发展，到干部人才渐见竭蹶，艰于应付的时候，环境倒越加复杂起来了⁴。控制力不免趋于薄弱了。一部历史，‘政怠宦成’的也有，‘人亡政息’的也有，‘求荣取辱’的也有。总之没有能跳出这周期律⁵……希望找出一条新路，来跳出这周期律的支配……”

毛主席是这样回答的：“我们已经找到新路，我们能跳出这周期律。这条新路，就是民主。只有让人民来监督政府，政府才不敢松懈。只有人人起来负责，才不会人亡政息。”

政府是一个系统，群众作为政府服务的对象，处于系统的外部，但是仅仅靠系统内部的自律是不够的，还需要“人民来监督政府”，也就是来自系统外部的合外力或合外力矩。政府和政权必须遵循全心全意为人民服务的宗旨，保持同人民的血肉联系，并接受人民监督。



从哲学的角度看

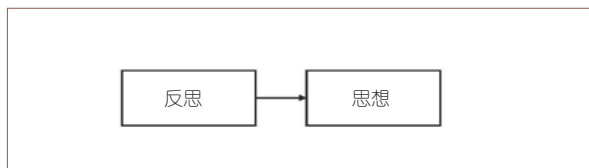
哲学的思维方式是反思^[3](reflection)。反思不是

³ 这里所描述的是人类社会的高熵情景，类似地，数据中心在高并发时请求的尾延迟剧增，熵亦剧增。

⁴ 这里说的是恶性循环，本质是正反馈。证明分布式系统领域著名的CAP定理的那篇文章^[7]第2页提到，系统最需要良好服务的时候，往往也是最容易出问题的时候，因为此时请求的并发度高，资源争用剧烈，本质上也是正反馈。

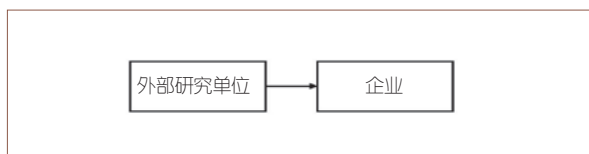
⁵ 这里的历史周期律在本质上是一种特定形式的系统局限性。

反复思考，而是思想以自身为对象反过来而思之。当代哲学学者孙正聿⁶在《哲学通论》^[1]第94页中指出，在人类思维的这种反思活动中，作为“感觉”和“直观”、“想象”和“意志”的全部“精神”活动，以及这些“精神”活动的全部“对象”，都处于“思维”的统摄下而成为“反思”的对象。如果说计算系统中正在进行的数据访问是计算系统思想的一部分，那么DIP就是对思想本身的思想，因此是一种计算系统本身的反思。



从管理学的角度看

2019年，华为公司出版了一本阐述熵减思想的书^[8]，系统地阐述了任正非管理思想核心理念。该书的第9页给出了华为的活力引擎模型。模型把企业比作一个系统，由于万物发展的自然走向是热力学第二定律的熵增，企业的自然走向就是组织懈怠、流程僵化、技术创新乏力、业务固定守成。为了实现企业的生存和发展，必须减熵，其中的一个有效办法是开放合作，这与DIP的思想是一致的。



结束语

如何构建具备更强能力的计算系统，是我们面临的一个重要问题。标签化冯·诺伊曼体系结构的目的是为了让传统的计算系统具有一定的反思能力。从多学科的角度看标签化冯·诺伊曼体系结构的必

要性，不仅有助于我们更好地理解 and 设计该体系结构，还有助于我们探讨递归、机器智能、悖论、元数学、哲学反思、系统熵之间的本质联系，从而助推多学科的融合。■



刘宇航

CCF 专业会员，CCCF 特邀译者、特邀专栏作家。中国科学院计算技术研究所副研究员，硕士生导师，主要研究方向为计算机体系结构、高性能计算、大数据、智能并发系统。liuyuhang@ict.ac.cn



刘宗惠

武汉理工大学计算机科学与技术学院本科生。主要研究方向为计算机系统结构和存储系统。

参考文献

- [1] Turing A M . Computing Machinery and Intelligence[J]. *Mind*, 1950, 59(236): 433-460.
- [2] 黄炎培. 八十年来[M]. 文史资料出版社, 1982年8月版, 148-149.
- [3] 孙正聿. 哲学通论[M]. 复旦大学出版社, 2019年4月第2版, 91-96.
- [4] Ma J, Sui X, Sun N, et al. Supporting Differentiated Services in Computers via Programmable Architecture for Resourcing-on-Demand (PARD)[J]. *ASPLOS' 15: Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems*. ACM Press, 2015: 131-143.
- [5] Bao Y, Wang S. Labeled von Neumann Architecture for Software-Defined Cloud[J]. *Journal of Computer Science and Technology*, 2017, 32(2): 219-223.
- [6] Xu Z, Li C. Low-entropy Cloud Computing Systems[J]. *SCIENTIA SINICA Informationis*. 2017,47(9): 1149-1163.
- [7] Gilbert S, Lynch N. Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services[J]. ACM

⁶ 孙正聿，首届国家级教学名师，曾任教育部哲学学科教学指导委员会主任、吉林大学学术委员会主任。

SIGACT News, 2002, 33(2): 51-59.

- [8] 华为大学. 熵减: 华为活力之源 [M]. 中信出版集团. 2019年8月第1版.
- [9] 李德毅, 刘常昱, 杜鹞. 不确定性人工智能 [M]. 软件学报, 2004, 15(11):1583-1594.
- [10] 薛兆丰. 经济学讲义 [M]. 中信出版集团. 2018年7月第1版.
- [11] 钱学森, 于景元, 戴汝为. 一个科学新领域——开放的复杂巨系统及其方法论 [M]. 自然杂志, 1990(1):3-10.
- [12] 欧内斯特 内格尔, 詹姆士 R 纽曼. 哥德尔证明 [M]. 中国人民大学出版社. 2008.