**第17章　万物皆是图灵机？**

不管你对图灵的概念或原理有多深的了解，都不会对你构建一台真正的计算机有所帮助。数字计算机是由半导体管和其他一些如继电器和真空电子管等开关机制部件构建的。这些半导体管组装成逻辑门，从而实现简单的逻辑功能。寄存器和累加器等高层次的组件都是由这些逻辑门构成的。1

1 这些层次可参见本书作者所著的*Code: The Hidden Language of Computer Hardware and Software*（Microsoft Press，1999）。

图灵机是由什么构成的呢？图灵从来没有告诉过我们。图灵并没打算让他的机器成为实际计算机的蓝图。图灵机是对计算的一种简单抽象的模型，这种计算既可以由人完成，也可以由机器完成。图灵创造图灵机的初衷是为了一个特定的目的：证明对于一阶逻辑并不存在通用判定过程。只是后来，我们发现这种想象中的机器对理解计算理论也有重大的辅助作用。这个转变花了20年的时间，也就是在图灵机成为了我们目前称之为“计算机科学”这一学科的研究对象之后。

将图灵机应用于其他目的（而不是单纯为了证明判定性问题），需要在某种程度上重新改造图灵机。大多数图灵机会永不停息地计算0到1之间某个实数的数位。数学里的常见任务，也是计算机编程里的常见任务，就是函数的计算。一个函数需要一个或者几个数作为输入——称作函数的参数。基于输入，函数会输出计算的结果——称作函数的值。

一类重要的函数就是数论函数，之所以这么称呼是因为它的输入和输出仅限于自然数。图灵在其论文的第10节（本书第219页）中发明了一种技巧来计算数论函数，即打印由单个0间隔的连续1的串。第一个0前的连续1的个数表示参数为0的函数值，第二个0前（第一个0后）的连续1的个数表示参数为1的函数值，等等。

对图灵的数论函数持怀疑态度的一位数学家是斯蒂芬·科尔·克莱尼。克莱尼是邱奇在普林斯顿的学生，他在1934年取得博士学位，之后开始在威斯康辛麦迪逊大学教书。

克莱尼后来写道：“虽然我很向往图灵所构想的机器的非凡能力，但我依然对他用如此简单的方法将此应用到数论函数的计算中表示怀疑。在任何情况下，只有完全函数才能用这种方法计算。”2图灵的方法对部分函数不适用，这些函数只对自然数的部分子集成立。

2 斯蒂芬·科尔·克莱尼，“Origins of Recursive Function Theory”，*Annals of the History of Computing*，Vol. 3, NO.1（Jan.1981），61。

1941年开春，克莱尼在威斯康辛麦迪逊大学教授的一个数学基础研讨班上开始寻求一种不同的解决方法。克莱尼重新改进的图灵机在他1952年出版的经典图书《元数学引论》的第8节中占据了重要的位置。

克莱尼版本的图灵机仍然是读符号，写符号，沿纸带左右移动。不过，它只限于一种符号，即一个简单的竖线，称为tick或tally符号。机器只有这种符号和空格。自然数以由空格分割的一系列连续格上的tick符号表示。克莱尼的自然数以0开始，一个tick符号表示0，2个tick符号表示1，以此类推。克莱尼好像是第一个在文章中将图灵机纸带的例子作为插图的人。3

3 斯蒂芬·科尔·克莱尼，*Introduction to Metamathematics*（D. Van Nostrand，1952），358-360。

图灵机通常以一个空白纸带开始。而克莱尼改进后的机器则从一个已经编码了一个或几个隔着空格的连续tick串符号（作为函数输入）的纸带开始。克莱尼的机器稍后计算函数的值，并将数字编码到纸带上。克莱尼给出的第一个例子是计算后继函数（即计算被编码的数的下一个数）的值。它只是简单地将另外的tick符号打印在现有tick符号串的后面，因此很方便。

克莱尼的函数计算机器只需要在一段有限的时间内进行计算，当机器完成计算的时候就停止了。这种机器并没有特殊的“停止”或“停机”格局，但是有克莱尼所谓的“被动状态”，即已经不存在机器可以到达的位置。当机器被指令转移到不存在的格局时，“这个机器被称为停止了，我们称它停止时候的状态为终止状态或者输出”。4

4 克莱尼，*Introduction to Metamathematics*，358。

在图灵的概念里，一台好的机器——图灵称之为非循环机，即符合要求的机器，是永不停止的。经过克莱尼改造后，一台好的机器将在计算完函数后停止运行。一台陷入了无限循环而无法停止的克莱尼机是“不好”的机器。显然，克莱尼的机器更接近传统的数学观念，即函数接受输入并经过有限步骤输出结果。

就如第15章讨论的，到了1936年，已经存在3种形式的计算有效性直观表示，它们是：

* 图灵机；
* 1934年，哥德尔基于雅克·赫尔布兰德的建议而定义的递归函数，克莱尼做了进一步的发展；
* 邱奇及其学生（主要是克莱尼）发展的λ可定义函数。

这三种不同形式表示方法的等价性，一部分是由图灵在1936年关于可计算数的论文的附录中建立起来的，更严格的说明是在其1937年的论文“可计算性和可定义性”中。另外，斯蒂芬·克莱尼在1936年的论文“λ可定义性与递归性”中也有所说明。现在“递归函数”与“可计算函数”几乎表达同一意思。

斯蒂芬·克莱尼是第一个提出这些形式化表示方法如何直观表达可计算性的人。他在《元数学导论》一书中，第一次明确提出邱奇论题：“每一个有效可计算函数（或者有效可判定谓词）都是一般递归的。”克莱尼又在其后的两章中说：“图灵论题，即每一个被自然认为可计算的函数在他的定义下（即通过一台图灵机）也是可计算的，它实际上与邱奇论题等价……”5

5 克莱尼，*Introduction to Metamathematics*，300，376。

在其1967年出版的书中，克莱尼将两个论题结合在了一起：

图灵论题和邱奇论题是等价的。我们应该将它们统一称为“邱奇论题”，或者“邱奇-图灵论题”，以表明它和三种形式化表示方法之一的“图灵机”有关。6

6 克莱尼，*Mathematical Logic*（John Wiley & Sons，1967；Dover，2002），232。

自那以后，“邱奇-图灵论题”成为了最适当的术语。

《元数学导论》显然是一本面向数学家们的书。6年以后，另一本经典著作帮助我们跳出纯数学的视野，从计算机科学的角度阐述问题。

马丁·戴维斯于1928年生于纽约市。他在1950年获得了普林斯顿大学的博士，其博士论文是《递归的不可解性理论》。戴维斯的论文导师是邱奇，也是克莱尼（1934年）和图灵（1938年）的导师。

在伊利诺伊大学教书时，戴维斯开始称判定图灵机能否完成计算的问题为“停机问题”，也许最早是在1952年。7这个术语在1958年戴维斯出版了《可计算性和不可解性》一书后而广为人知。在这本书的前言里，戴维斯诡秘地写道：“虽然这一卷里真正新的东西很少，但是我对相应主题的安排和论述可能会让专家感到新颖，”随后他做了说明，“特别是，图灵机的概念是这本书行文论述的关键。”8

7 参见B. Jack Copeland，*The Essential Turing*（Oxford University Press，2004），40，脚注61。

8 马丁·戴维斯，*Computability and Unsolvability*（McGraw-Hill，1958，Dover，1982），vii-viii。

在克莱尼的《元数学导论》中，图灵机直到第321页才出现，在第13章之前也没有更深的提及；而在戴维斯的《可计算性和不可解性》中，图灵机在第1章的第一页就出现了。

和克莱尼一样，戴维斯把自然数表示为连续的tick符号，并用图灵机来计算函数。用图灵机计算加法、减法、乘法的例子出现在书的第12页。

虽然《可计算性和不可解性》表面上是数学图书，但是戴维斯意识到这本书“由于和某些哲学问题以及数字计算机理论相关，一些非数学家也可能有兴趣一读”。9

9 马丁·戴维斯，*Computability and Unsolvability*，vii。

为进一步强调不同，《可计算性和不可解性》作为McGraw-Hill出版公司“信息处理和计算机系列丛书”中的一部作品出版。即使在这套丛书中，这本书也是独一无二的。其他书都关注于计算机硬件和程序设计的“实用”主题。在1958年和1959年，该丛书出版了《模拟仿真：场问题的求解》、《高速率数据处理》、《数字计算机入门》、《数字计算机系统》和《数字计算机编程入门》。

马丁·戴维斯的《可计算性和不可解性》真正开创了将可计算性作为一门学科主题的研究，它后来成为计算机科学专业学生必修的一门课程。

在《可计算性和不可解性》的第70页，马丁·戴维斯引入了一个和图灵机密切相关的术语：

现在，令*Z*表示一个简单图灵机。关于*Z*，我们有如下判定问题：

对于一个给定的瞬间描述*α*，判定是否存在一个以*α*开始的对*Z*的计算。

也就是说，我们希望确定如果给定初始状态，那么*Z*会不会最终停止？我们将这个问题称为*Z*的停机问题10。

10 马丁·戴维斯，*Computability and Unsolvability*，70。

在第70页的最后，戴维斯构想了一个定理：“存在一种图灵机，其停机问题是递归无解的。”

戴维斯的书影响深远，一谈到停机问题就会联系到图灵机，尽管图灵原来设想的图灵机是永不停止的。

除了速度、存储能力、人机交互设备不尽相同外，现代计算机大体上类似。每台能模拟图灵机的计算机（这是最简单的需求）都是一台通用计算机。此外，一台通用计算机能够模拟任何其他通用计算机。

最初的一些很简单的计算机甚至还达不到图灵机的能力。显然，第一台至少能称为潜在通用计算机的是Z3，由康拉德·楚泽在1938年至1941年建造。11如果说是建造成功，第一台通用计算机应该算是查尔斯·巴贝奇在19世纪30年代建造的分析引擎，虽然它是由一堆齿轮而不是转换开关构成的。实际上，所有1944年以后生产的计算机都是通用计算机。

11 Raúl Rojas，“How to Make Zuse's Z3 a Universal Computer”，*IEEE Annals of the History of Computing*，Vol. 20，No. 3（1998），51-54。

通用计算机的一个关键特性是可编程性。必须存在某个将指令集引入计算机的方法，使计算机能响应这个指令集。现代计算机将这样的指令集作为字节存储在内存里，称为**机器码**。在楚泽的机器里，指令集的编码用35mm的电影胶带上的打孔来表示。巴贝奇的机器则使用打孔的卡片，与控制纺织机的卡片有点像。

一些早期的计算机只能用很不灵活的指令序列进行编程。一个通用计算机必须能够根据上几步的计算结果跳过一些指令序列。这个特性我们现在称为**条件分支**，这对于实现条件循环是必要的。

如果一门计算机编程语言能够模拟图灵机，它就经常被称为是**图灵完备**的。

最初在互联网上广泛使用的超文本标记语言（HTML）并不是用来为计算服务的，因此它显然不是图灵完备的。经常在HTML中使用的JavaScript就是图灵完备的。几乎所有现今使用的编程语言都是图灵完备的。任何图灵完备的编程语言都可以模拟任何其他图灵完备的编程语言。

图灵机不仅说明了进行有效计算的最基本需求，也说明了它的限制：没有任何一台现代计算机或一门编程语言能够比图灵机更强大，没有任何计算机或编程语言能够解决停机问题，没有任何计算机或编程语言能够判定其他计算机程序的未来运行状态。你不能使用更“先进”的编程语言或者不一样的机器来应付这种限制。你能做的，只能是加快计算机的工作速度。你可以拿出上千个处理器组成并行计算机集群，以进行大规模的并行计算，但是你不可能将无限带往我们生活的这个无助的有限世界里，哪怕一点点。

一些数学家不顾图灵机的限制，毅然钻入**超计算**领域，以试图让机器摆脱图灵的限制。图灵自己也部分推动了这项工作。他在自己1939年发表的晦涩的博士论文《基于序数的逻辑系统》（*Systems of Logic Based on Ordinals*）中称之为“神谕”，并写道：

假设我们拥有一种可以解决（不可判定）数论问题的不确定方法，那么可以称这种方法为神谕。对于这个神谕，我们除了知道它肯定不是一台机器外无法知道更多。在这个神谕的帮助下，我们可以构造一种全新的机器（称为o-机器），这个机器的某一个基本进程可以用来解决给定的数论问题。12

12 阿兰·图灵，“Systems of Logic Based on Ordinals”，*Proceedings of the London Mathematical Society*，Series 2，Vol. 45（1939），172-173。

也许我们每个人都希望生活中有这样的神谕，来帮助我们解决棘手的问题。探索超计算的研究学者基于神谕的概念，将其他特性引入图灵机以使其不再受到先前的限制。虽然出现了一些很有趣的数学构造，但是这样的超计算机并不切合实际，因为它违反了一些基本的物理定律，比如加速时间使得每步计算都是前一步计算的1/2时长。马丁·戴维斯形容超计算是谜一样的问题，并将它与三等分给定角和发明永动机这样的问题作比较。13

13 马丁·戴维斯，“The Myth of Hypercomputation”，Christof Teuscher，ed., *Alan Turing：Life and Legacy of a Great Thinker*（Springer, 2004），195–211。

在我看来，超计算的研究对于解答计算的普遍性是很有价值的。图灵设计出假想的机器，以刻画人类计算者在执行特定算法时进行的基本操作。他发现，图灵机具有一些固有的限制。那以后的几十年，我们建造了和图灵机等价的计算机，因此同样面临这些限制。我们还没有找到有效的方法来摆脱这些限制。

基于上述原因，计算的普遍性（在能力和限制上）对于任何数据处理活动而言似乎都是基本存在的。这些限制就像是热力学定律一样是自然界内在的规则。

如果图灵机的内在限制不能在遵守物理定律的前提下被超越，那么对于那些执行计算或逻辑运算的内在机制而言，这又暗示着什么呢？当我们从探索人的思维和宇宙自身的角度来考虑这两个最重要的（也许甚至有些令人烦恼的）“内在机制”时，这个问题变得最为深刻。

严格地说，图灵定理只涉及图灵机和机械算法的等价问题，并不一定意味着不存在超越图灵机计算能力的计算机器，这样的机器也未必一定违反某个物理定律。14

14 参见C. Jack Copeland，“The Church- Tuning Thesis”里面有准确的评论。Stanford Encyclopedia of Philosophy，<http://plato.stanford.edu/entries/church-turing>。

也许我们漏掉了什么，也许存在某种神奇的物理机制能够执行非常强大的计算操作，而这种机制无法在图灵机上模拟出来。图灵机真的有助于我们对人类思维和宇宙的理解吗？或者我们只是愚蠢地把一个非常复杂的问题简单化到图灵机的层次上？

图灵机在数学和计算领域外的遗留问题出现在图灵1936年论文发表之后的几年，源于沃伦·麦卡洛克（1898—1969）和沃特·匹茨（1923—1969）的一次偶遇。

在底特律，青年沃特·匹茨聪明好学，自学了拉丁语和希腊语、哲学和数学，家里人都认为他是一个怪才。15岁的时候，他跑到了芝加哥。因为无家可归，沃特·匹茨大多数时间游荡在公园里，在那里，他遇见了一位名叫伯特的老人，他们在哲学和数学上有着相同的兴趣。伯特建议他读一读由芝加哥大学教授鲁道夫·卡尔纳普（1891—1970）写的一本出版于1937年的书，可能是《语言的逻辑句法》。沃特·匹茨读了这本书后，就径直前往卡尔纳普的办公室，与他讨论自己在书中发现的几个问题。这个叫伯特的老人就是鼎鼎大名的伯特兰·罗素。15

15 这个故事来自Pamela McCorduc所著的*Machines Who Think A Personal Inquiry into the History and Prospects of Artificial Intelligence* [ 25th anniversary edition（A. K. Peters，2004）89 ]一书中，由麦卡洛克的前学生Manual Blum所述。它也出现在由麦卡洛克的夫人Rook McCulloch编辑的*Collected Works of Warren S. McCulloch* [ Intersystems Publications，1989， Vol. I， 31 ]一书Manual Blum的文章“Notes on McCorduc-Pitts'*A Logical Calculus of the Ideas Immanent in Nervous Activity*”中。故事的另一版本还详述于James A. Anderson和Edward Rosenfeld所编*Talking Nets: An Oral History of Neural Networks* [ MIT Press，1998，218 ]中对Michael A. Arbib的访谈。

如果你不相信这个故事，那么下面的故事似乎更可信。伯特兰·罗素在芝加哥大学教书的时候，有一次散步从杰克逊花园经过，发现有个年轻人正在看卡尔纳普的书。于是，罗素和年轻人开始交谈起来，并把匹茨带到了卡尔纳普的办公室。16

16 采访自Jack D. Cowan *Talking Nets*，104。

还有一则故事。当匹茨12岁、还生活在底特律的时候，有一次被几个流氓追赶，躲进了图书馆。图书馆关门后，匹茨被困在了里面。他决定读怀特海和罗素的《数学原理》作消遣。这么一读就是3天，然后他给罗素写了一封信，指出了书里的一些错误。当罗素写信邀请他去剑桥时，匹茨决定成为一个数学家。17

17 采访自Jerome Y. Lettvin，*Talking Nets*，2。参见Jerome Y. Lettvin，“Warren and Walter”，*Collected Works of Warren S. McCulloch*，Vol. II，514-529，匹茨在图书馆开放时间用了一星期阅读《数学原理》。

有一点是可以确定的，匹茨在1938年上过罗素在芝加哥大学教的课，同年他也到过卡尔纳普的办公室。卡尔纳普对这个年轻人印象深刻，想给他一份学生助教的工作，但是他并不知道匹茨的名字，所以无从找到他。18

18 Neil R. Smalheiser“Walter Pitts”，*Perspectives in Biology and Medicine*，Vol.43，No. 2 （Winter 2000），218。

匹茨是个“害羞、内向的孩子，他戴着眼镜，牙齿不齐，有拨弄头发的习惯，有轻微的神经颤抖症，走路时还经常会碰到东西”。19（在这以后的第二次世界大战期间，匹茨被征兵局划入4F级别20，并被认为处于“精神病发作前期”。但是匹茨后来参加了曼哈顿计划，并获得接触最高机密的许可。）21在匹茨访问了卡尔纳普的办公室几乎一年后，卡尔纳普终于找到了匹茨。匹茨开始跟着他学习逻辑学，同时在芝加哥大学上课，其中包括了红胡子的乌克兰人尼古拉斯·拉谢甫斯基（1899—1972）主持的研讨班。

19 Smalheiser，“Walter Pitts”， 22。

20 4F级别指“精神或身体上不适合服兵役”的一类人群。——译者注

21 同19。

拉谢甫斯基在基辅大学获得了理论物理的博士学位，1924年移民到美国。他对应用数学模型解决生物生长发育过程中的问题很有兴趣。在这个领域中，之前其他科学家都仅仅依赖于经验性的研究，并没有相关的科学实验方法。到了1934年，拉谢甫斯基提出了一个名字来形容他所研究的工作：“数学生物物理学”。1935年，他成为芝加哥大学第一名数学生物物理助教。1938年，一本名为《数学生物物理学》的书出版了，随后在1939年，诞生了名为《数学生物物理学报》的期刊，主要发表拉谢甫斯基及其理论追随者的一系列论文。22

22 Tara H. Abraham，“Nicholas Rashevsky's Mathematical Biophysics”，*Journal of the History of Biology*，Vol.37，No.2（Summer 2004），333–385。

1942年和1943年期间，匹茨接连在《数学生物物理学报》上发表了3篇论文。沃伦·麦卡洛克就是在这时听说了匹茨的这些研究工作。

沃伦·麦卡洛克在新泽西长大，一开始在哈弗福德学院，这是宾夕法尼亚州的一座“教友派”学校。当麦卡洛克1917年进入大学不久，哲学家鲁弗·琼斯（1863—1948）（他在这期间帮助创建了美国朋友服务委员会）就问了麦卡洛克一个问题：“你将会成为什么样的人？……你将会做什么事?”麦卡洛克说不知道，但是他说：“我想回答的是另一个问题：人可以认知的数是什么？可以认知数的人又是什么？”鲁弗·琼斯对此只能回应道：“朋友，看来你的人生要在忙碌中度过了。”23

23 沃伦·麦卡洛克，“What is a Number，That a Man May Know it, and a Man, That he May Know a Number ?”，*Collected Words of Warren S. McCulloch*，Vol. IV，1226。

后来，麦卡洛克去了耶鲁大学学习哲学和心理学，于1927年在纽约的内外科医学院获得硕士学位，并在贝尔维尤医院从事治疗严重脑损伤病人的工作。后来，他又去了洛克兰德州立医院从事治疗精神病患者的工作。241934年，麦卡洛克回到了耶鲁。他的同事里有杜赛尔·德·巴瑞内（1885—1940），巴瑞内是使用化学成分马钱子碱作用在猫的大脑并观察猫的反应，从而探明大脑各部分对应功能这一实验方法的先驱。1941年，麦卡洛克搬到了伊利诺伊州，在伊利诺伊神经精神病学院工作。

24 麦卡洛克的很多传记信息来自Michael A. Arbib，“Warren McCulloch's Search for the Logic of the Nervous System”，*Perspectives in Biology and Medicine*，Vol.43，No.2（Winter 2000），193–216。

麦卡洛克是个传奇人物。“有点像摩西：有一撮长长的胡子，浓密的眉毛，眼睛发出奇异的光芒。很多时候，他看上去就像个疯子。他的眼睛是灰色的，当它们变得明亮而闪烁的时候，就像是一副眼镜。”25麦卡洛克很喜欢社交，“每天晚上都要喝一瓶苏格兰威士忌作为他和别人闲聊时的情绪催化剂”，26他也是个很会讲故事的人。（有关沃特·匹茨和伯特兰·罗素在杰克逊公园相遇的故事最早来源于麦卡洛克。）麦卡洛克会写诗，还经常论述他的哲学观点，到处炫耀自己的博学多才。

25 采访自Jack D. Cowan，*Talking Nets*，102。

26 Arbib，“Warren McCulloch's Search for the Logic of the Nervous System”，202。

沃伦·麦卡洛克和沃特·匹茨事最初是怎么相识的，至今也不清楚。但是他们很快就情投意合，匹茨甚至搬来和麦卡洛克一起住。麦卡洛克一直想尝试建立一套大脑工作机理的形式化理论，而匹茨对于数学逻辑的通晓正是麦卡洛克所需要的。他们在麦卡洛克厨房的桌子上写出了名为《神经活动中内在意识的逻辑演算》的论文，1943年发表在拉谢甫斯基的《数学生物物理学报》上。麦卡洛克的女儿塔菲为他们第一次合作的这篇论文画了图解。27

27 Arbib，“Warren McCulloch's Search for the Logic of the Nervous System，”199。

得益于19世纪下半叶的研究，科学家们已经知道神经系统是由叫做**神经元**的细胞构成的，这些神经元似乎像网络一样连接在一起。20世纪的进一步研究指出，这些神经元就像开关一样，当刺激达到一定阈值时开关就被触发。28

28 Tara H. Abraham，“（Physio）logical Circuits：The Intellectual Origins of the McCulloch-Pitts Neural Networks，” *Journal of the History of the Behavioral Sciences*，Vol.38，No.1（Winter 2002），19。

对于麦卡洛克和匹茨来说，这些神经元就像逻辑开关，于是他们用卡尔纳普的标注方法对神经元进行命题逻辑意义上的建模。一个重要的、传统逻辑中并不存在的要素是，神经元逻辑存在输入与输出的延时特性。基于这样的延时，神经元就被组成为环状结构，从而信号可以在网络内保持一定时间的有效期。麦卡洛克和匹茨的论文为这个模型定义了几个公理，并进而证明了几个定理。

《神经活动中内在意识的逻辑演算》并没有参考前人太多的成果。这篇论文注明的参考文献只有卡尔纳普的《语言的逻辑句法》、希尔伯特和阿克曼的《数理逻辑原理》，以及怀特海和罗素的《数学原理》。在这篇19页论文的第15页，麦卡洛克和匹茨在做总结的时候透露了一些他们所参考的更广泛文献来源：

首先，如果每个网络接有纸带，它的扫描头与传入神经连接，并且有适合的传出神经进行必须的操作，那么它就只能计算图灵机能够计算的那些数；其次，图灵机能够计算的数也能够被这样的网络所计算……这就从心理学的角度，论证了图灵关于可计算性的定义，以及与其等价的邱奇的λ可定义、克莱尼的一般递归：如果任何一个数能够被有机体计算出来，那么它也可以被这些等价定义计算出来，反之亦然。29

29 麦卡洛克和匹茨，“A Logical Calculus in the Ideas Immanent in Nervous Activity”，*Bulletin of Mathematical Biophysics*，Vol. 5（1943），129。也可参见 *Collected Works of Warren S. McCulloch*，Vol. I，357。

几年后的1948年，麦卡洛克将大脑与图灵机的联系表述得更加清楚。他解释说，他正努力寻找一种方法来发展神经生理学理论。

直到我看了图灵的论文，才发现找到了正确的途径，也感谢匹茨那些必要的逻辑演算的帮助。我们认为，我们正在做的（我想我们获得了相当的成功）是将大脑视为一台图灵机；除非大脑出错，机能失常，否则大脑所进行的就是这台图灵机进行的功能……令人高兴的是，一些非常简单的假设就足以说明神经系统可以计算任何可计算的数。大脑就是一种仪器，一台图灵机（如果你喜欢这个词）。30

30 麦卡洛克Lloyd A. Jeffress，ed.，*Cerebral Mechanisms in Behavior: The Hixon Symposium*（John Wiley & Sons, 1951），32–33。

一直到后来（1955年），麦卡洛克依然坚持他的这个观点：“匹茨和我已经说明大脑就是图灵机。任何图灵机都可以由神经元构成。”31虽然现在的科学技术还不足以将这种等价性转换为实际用途。

31 麦卡洛克，“Mysterium Iniquitatis of Sinful Man Aspiring into the Place of God”，*The Scientific Monthly*，Vol. 80，No.1（Jan.1955），36。也见*Collected Works of Warren S. McCulloch*，Vol. III，985。

一个理论上的问题是，你能设计出一台机器做大脑所能做的事情吗？答案是：如果你能用有限的而且不含糊的方法说明大脑能够做的事情，我们就可以设计出这样一台图灵机。匹茨和我证明了如何完成这样的构造。但是你能说出大脑能够做的事情吗？32

32 同上，38。

要不是引起了两位20世纪计算领域的巨匠诺伯特·维纳和冯·诺依曼的注意，麦卡洛克和匹茨的论文可能会因无人知晓数学生物物理学而失去光彩。

继约翰·斯图亚特·穆勒后，诺伯特·维纳是又一个被那种臭名昭著的天才家庭教育模式培养长大的例子。这两个人后来都曾在回忆录中写过小时候被满脸胡须的父亲塑造为天才儿童的体验。而对于维纳，小时候留下的伤疤终其一生都难以愈合。多年来，他要和未经诊断的抑郁症抗争，与其天才般的研究热情相伴的是难以名状的暴躁情绪和产生自杀倾向的绝望状态。

11岁时，维纳就进入了塔夫斯大学，14岁时拿到了数学学士学位，18岁成为哈佛大学历史上最年轻的博士毕业生。维纳的父亲一直对新闻界说，他的儿子“并不是神童”，事实上“很懒惰”33。他的父母还对外界隐瞒了一个事实：维纳在15岁之前一直不知道自己是犹太人。

33 Flo Conway and Jim Siegelman，*Dark Hero of the Information Age: In Search of Norbert Wiener, the Father of Cybernetics*（Basic Books，2005），21。

离开哈佛后，维纳在剑桥随罗素学习数理逻辑，随哈代学习数论。第一次世界大战前夕，他又来到了哥廷根，随希尔伯特学习微分方程。后来，他又去了哥伦比亚大学，随约翰·杜威学习哲学。1919年，他成为麻省理工学院的教员。

在战争期间，维纳是当时新兴的通信工程和模拟计算研究领域的先锋。他参加了范内瓦·布什在MIT组织的模拟信号计算的研究项目，而且似乎这对克劳德·香农发展通信理论产生了影响。在二次世界大战期间，维纳的工作项目是研究防空火力系统。这些系统加入了比原有技术更复杂的预测部分，目标是预测飞行器为躲避导弹可能采取的路线。维纳对反馈的概念异常有兴趣，反馈就是从一个正在进行的进程中不断得到运行信息而反过来 修正这个进程。

维纳并没有参加1942年5月13日在贝克曼酒店召开的第一次物理学、生物学和人类学大会，这次会议由小约西亚·梅西基金会赞助，旨在拓展各学科间的交流。麦卡洛克参加了那次会议。参加那次会议的还有人类学家格列高里·贝特森和玛格丽特·米德夫妇。维纳参加了战后的第一次梅西基金会的会议，那次会议的主题是“生物和社会科学中的反馈机制和循环因果系统”。34参加这次会议的还有匹茨和冯·诺依曼。这样的会议让每个人都有机会吸收他人的研究成果，并探索各自研究领域是否存在一些契合的研究目标。

34 Conway and Siegelman，*Dark Hero*，155。

1947年，维纳写了一本书，总结了这些会议上讨论的一些研究工作。他想用一个新词来描述包含了机器、生物和社会结构领域中的各种通信和反馈的研究工作。他选择了一个希腊词cybernetics，原意是“舵手”，因为舵手的本质也是利用反馈来修正航线的偏移。维纳的书出版于1948年，书名最后定为《控制论：关于在动物和机器中控制和通信的科学》（*Cybernetics: Control and Communication in the Animal and the Machine*）。

美国《时代》杂志评价道：“很少有一本书能够在很多不同的科学领域激起强烈的反响。而《控制论》就是这样的一本。”35今天再读起来，《控制论》是一本古怪而小巧的书，里面有很多篇幅的数学公式和令人目眩的、不太切合实际的议论。在引言里，维纳对那些启发过他的人们表示了感谢，包括麦卡洛克，“一个对研究脑皮层细胞的组织结构充满兴趣的人”；阿兰·图灵，“也许是第一个研究智能机器的逻辑可能性的人”；沃特·匹茨，“卡尔纳普在芝加哥的学生，和拉谢甫斯基及其生物物理研究团队一直保持着联系”。维纳也对计算机的先驱“哈佛的艾肯教授、高等研究院的冯·诺依曼教授、宾夕法尼亚大学负责ENIAC和EDVAC巨型机的戈登斯坦教授”表示了感谢。36

35 December 27, 1948 issue, 引自Conway and Siegelman，*Dark Hero*，182。

36 维纳，*Cybernetics: or Control and Communication in the Animal and the Machine* （John Wiley & Sons，1948；second edition，MIT Press，1961），12，13，14，15，页码出自第2版。

《控制论》的第5章是“计算机器和神经系统”。维纳对比了数字计算机的转换机制与麦卡洛克和匹茨的大脑模型：

众所周知，人和动物神经系统能够完成计算系统的工作。神经系统的一个显著特点是包含了适合做继电器的元素，这些元素称为神经元或神经细胞。虽然它们在电流的刺激下会呈现复杂的特性，但是其普遍的生物特性都遵循“全或无”的原则，即要么处于休息的状态，要么当“工作”的时候经历一系列变化，而这些变化与刺激的外部环境和强度无关。37

37 维纳，*Cybernetics*，120。

两章之后，维纳写道：“认识到大脑和计算机器实现之间的共同点，可能为精神病理学甚至是精神病学的发展提供新的有效方法。”38然而，维纳并不是一个狂热的纯技术狂。他也很关心这个新的科学技术对人类的影响。他写了《人有人的用处：控制论与社会》（*The Human Use of Human Beings: Cybernetics and Society*），作为1948年《控制论》的补充。

38 维纳，*Cybernetics*，144。

控制论成为很多领域的研究焦点，这种情况一直到了1951年，维纳突然不加解释地中断了与麦卡洛克和控制论研究团队的联系。这个研究团队的人员很大程度上是因为麦卡洛克的个人魅力而聚在一起的，其中包括匹茨，他的博士论文是在维纳的指导下完成的。关于他们关系分裂有几种解释。一种观点是麦卡洛克性格较为细腻，而精神有缺陷的维纳已经无法捕捉麦卡洛克言语表达中的细微差别。有些时候，维纳分辨不出麦卡洛克是在陈述一件事实还是在说一个猜想。39另一种观点是维纳的妻子出于嫉妒心，为了维护丈夫的声名，谎称麦卡洛克带领的团队中有人勾引他们的女儿。40

39 Arbib，“Warren McCulloch's Search for the Logic of the Nervous System，” 201-202。

40 Conway and Siegelman，*Dark Hero*，222–229。

作为一门统一的学科，失去了维纳和麦卡洛克联手的控制论遭受了很大的损失。在因这次决裂受到影响的人当中，匹茨可能受到的打击最大。他的精神完全垮掉了，他亲手毁掉了自己的研究成果和博士论文，开始了一段漫长的自我堕落的生活。“他不是简单地酗酒，这种人人都会的行为与他这样的高智商不匹配。他在实验室里自己合成了一种类似巴比妥类药物和鸦片的化学物质，伴着酒吞下去。”41匹茨于1969年死于慢性饮酒过多造成的食道静脉破裂，年仅46岁。

41 Smalheiser，“Walter Pitts，”223。

即便维纳和麦卡洛克之间没有分裂，也不能保证控制论会一直发展下去。在美国学术界，跨学科的概念并没有市场，专业性的研究才是取得成功的钥匙。虽然有很多复兴控制论的尝试，但是大多数只是在当今流行的词语中加入cyber前缀的文字游戏而已，就像cyborg（cybernetic organism的缩写，机械人），还有无处不在的cyberspace、cybercafe、cyberpunk和cybersex。即使是这些以cyber为前缀的词，近年来也逐渐被e开头的词取代了。

麦卡洛克和匹茨关于神经网络数学模型的论文给了冯·诺依曼很多启迪。冯·诺依曼参与了几个重大计算机项目的设计工作，包括计算机的雏形EDVAC（Electronic Discrete Variable Automatic Computer，电子离散变量自动计算机）。在EDVAC的第一份报告（1945年6月30日）中，冯·诺依曼这样描述计算机的开关机制：“每一台电子计算机都包含了起着中继作用的‘元素’，这些元素有着若干个离散平衡状态，并能保持在某个状态不变。”42通过引用麦卡洛克和匹茨的论文，冯·诺依曼写道：“值得一提的是，高级动物的神经元毫无疑问就是上述意义下的元素。”43

42 冯·诺依曼，*First Draft of a Report on the EDVAC*（Moore School of Electrical Engineering，1945），§4.1。

43 同上，§4.2。

次年，冯·诺依曼开始着手研究生物和机器之间的关系。他用了一个希腊单词来表现生物体的这种特性：**自动机**（automaton）。44在一封写给维纳的信中，冯·诺依曼惊叹于他们直接研究大脑这个自然界最复杂的人工自动机是多么具有雄心壮志的一项研究工作：

44 William Aspray，“The Scientific Conceptualization of Information: A Survey”，*Annals of the History of Computing*，Vol. 7，No. 2（April, 1985），133。

我们的思维，我指的是你、我和匹茨的思维，目前为止都主要关注在神经系统上，更确切地说，是人类的中枢神经系统。为了搞清楚自动机及其普遍的机制，我们选择了研究大脑这个普天之下最难的问题。事实上，如果没有这些大胆而艰苦的研究工作，现阶段的研究人员，至少是我，对自动机这个课题的认识就会比现在混乱得多。这里，我也要指出，图灵在非神经领域的研究也是如此的大胆卓越。45

45 Letter of November 29，1946，from Miklós Rédei，ed.，*John von Neumann: Selected Letters* （American Mathematical Society，2005），278。

对冯·诺依曼来说，自动机就是一台有着输入、输出和中间处理过程的设备。1948年9月，他在加州理工学院的“行为的大脑机制”研讨会上做了报告。他的演讲题为“自动机的一般逻辑理论”，包含了很多关于大脑与1948年时计算机在尺寸、速度、转换机制和能源消耗上的对比。他强调了开发一种新型逻辑的必要性，并开始思考后来成为他主要课题兴趣的问题：自复制自动机。46

46 冯·诺依曼，“The General and Logical Theory of Automata”in Lloyd A. Jeffress，*Cerebral Mechanisms in Behavior: The Hixon Symposium*（John Wiley & Sons，1951），1-41。

维纳对冯·诺依曼的这个想法开玩笑说：“我觉得你所说的（自动机）未来能够自我复制的能力很有意思……看来有机会可以写一个新的《金西报告》47（Kinsey report）了。”48 但对于冯·诺依曼来说，自复制自动机可不是一个玩笑。他一直在想是否有某种未知的规律可以阻止一台机器制造一台它的复制品。即使生物不是这样繁殖的（虽然DNA本身是自复制的），这样的问题也折射出了有趣的本体论的味道。

47《金西报告》原名《人类男性的性行为》，是世界性学史上的一部里程碑式的著作。在其广为流传后，人们更喜欢用作者的名字来称呼它。金西指作者阿尔弗雷德·金西，报告指它是一部性社会学意义上的调查报告和资料汇编。——译者注

48 Letter of August 10，1949，quoted in Steve J. Heims，*John von Neumann and Norbert Wiener: From Mathematics to the Technologies of Life and Death*（MIT Press，1980），212。

对自动机和图灵机的不断研究诞生了几部经典的开山之作，如通信理论的创始人香农和人工智能的先驱之一、也是Lisp语言的创造者约翰·麦卡锡合著的《自动机研究》（*Automata Studies*），由普林斯顿大学出版社在1956年出版。这本书包含了冯·诺依曼、克莱尼和人工智能先驱之一的马文·明斯基的几篇关于自动机的论文，也包含了香农和马丁·戴维斯关于图灵机的第一篇论文。

当20世纪50年代早期冷战逐渐升温的时候，维纳和冯·诺依曼发现他们各自站在了不同的政治立场上。维纳被美国在日本广岛和长崎投下的两颗原子弹的破坏力所震惊，他从此拒绝了从政府那里接受研究经费，他的文章也逐渐聚焦于现代科技的使用带来的战争和和平等社会问题。相反，冷战让冯·诺依曼产生了反共产主义的倾向，他成为核武器的坚定支持者。1955年，冯·诺依曼被诊断出骨癌，1956年开始住院治疗，次年与世长辞，终年53岁。冯·诺依曼的癌症很可能是他亲临原子弹试验场时受核辐射所致。49

49 Heims，*John von Neumann and Norbert Wiener*，369-371。

冯·诺依曼死后留下了一系列未完成的讲义，这些讲义汇集成《计算机和人脑》（*The Computer and the Brain*）一书，于1958年出版。这本书虽然不尽人意，但它还是给了我们很多冯·诺依曼可能想表达什么内容的一些诱人的暗示。基于冯·诺依曼那些关于自动机的未完成手稿，阿瑟·W. 巴克斯编辑并完成了一本名为《自复制自动机的理论》（*Theory of SelfReproducing Automata*）的书，于1966年出版。

在早期对自复制自动机的研究中，冯·诺依曼想象机器处在一个拥有很多备用零件的环境中，然后探讨这部机器是如何组装产生它们的复制品的。这样的自动机就是**运动自动机**（kinematic automata），基本上和我们平常所说的机器人是一个意思。

在与他的好友斯塔尼斯拉夫·乌拉姆（一位研究晶体增长的科学家）共同讨论后，冯·诺依曼决定先研究一个较为简单的模型——**元胞自动机**（cellular automata）。

元胞自动机是对于细胞结构的一个数学构造。元胞自动机可以以多种维度存在，但是实际的研究中基本上只考虑二维网格。每个网格中的细胞都会受隔壁网格中的细胞所影响，仿佛这些细胞连在了一个简单的网络中。经过持续的“移动”和“生成”，细胞根据特定规则变化不同的状态。元胞自动机的简单规则经常会产生复杂的行为。冯·诺依曼研究了拥有29个状态的元胞自动机，并且证明了这些自动机可以通过组装构成一个通用图灵机。50

50 William Aspray，*John von Neumann and the Origins of Modem Computing*（MIT Press，1990），203-204。

元胞自动机在20世纪70年代突然超脱了自己的学术圈。英国数学家约翰·霍顿·康威（1937—　）设计了一个他称为“生命游戏”（Game of Life）的简单元胞自动机。该自动机有一个简单规则：在一个类似方格纸的二维网格上，一个细胞要么是活的（方格被填充）要么是死的（方格未被填充）。在每一次后继的繁衍中，一个细胞根据它周围临近的8个细胞改变其自身的状态：如果一个活细胞被2个或者3个活细胞包围，它依然存活；如果被0个或1个活细胞包围，它会因孤独而死掉；被4个或4个以上活细胞包围，它也会因过度拥挤而死掉。一个被3个活细胞包围的死细胞则会因为一种“神奇”的繁衍形式而成活。

在《科学美国人》中，马丁·加德纳“数学游戏”专栏中的几个填充谜题让康威的游戏变得流行起来，511974年，《时代》杂志抱怨“价值几百万美元的计算机都把时间浪费在对这种游戏持续增长的狂热中了”。52当然，1974年并没有个人计算机，只有大型机。现在，这个游戏大多是在个人计算机中运行了，我想《时代》杂志此时不会再说这很狂热了吧。

51 这种游戏后来被马丁·加德纳收录，*Wheels, Life, and Other Mathematical Amusements*（W. H. Freeman，1983）。

52 1974年1月21日发行，引自William Poundstone，*The Recursive Universe：Cosmic Complexity and the Limits of Scientific Knowledge*（William Morrow，1985），24。

虽然只有简单的规则，但这台自动机却能呈现出一些非常复杂的模式，例如它可以呈现出不断繁殖后代的模式。虽然看上去不可能，但是图灵机确实可以通过这样的元胞自动机构造出来，这台自动机是图灵完备的。53

53 Paul Rendell，“A Turing Machine in Conway's Game of Life”，March 8，2001。<http://www.cs.ualberta.ca/~bulitko/F02/papers/tm_words.pdf>。也可见<http://rendell-attic.org/gol/tm.htm>。

另一位对元胞自动机的研究感兴趣的人是德国工程师康拉德·楚泽。楚泽比图灵早两年零一天出生。当图灵在写他关于可计算数的论文时，楚泽正在他父母柏林的公寓里制造计算机。

1969年，楚泽出版了一本74页的书，名为《计算空间》（*Rechnender Raum*），是“数字物理”领域最早的一部著作。数字物理是研究如何在可计算的框架内解释宇宙运动法则的学科。

传统上，物理定律是假定为连续的。距离、速度、质量和能量的度量似乎最适合用实数表示，用微分方程运算。但是一些量子理论的观点指出，宇宙内在的自然结构可能是离散的、数字的。现实世界里自然界的连续性可能只是一个假象。“宇宙到底是数字的，还是模拟的，抑或两者皆有呢？”楚泽问道，“而提出这个问题本身是不是合理呢？”54为了从数字角度探索物理定律，楚泽创造了可以被元胞自动机运算的“数字粒子”的概念。《计算空间》是用数字物理描述宇宙的一个实验性的尝试，但毫无疑问，它是一个大胆的创新。

54 康拉德·楚泽，*Calculating Space*，译自*Rechnender Raum*（MIT Technical Translation，1970），22，（德文本p16）。

初看之下，很难把宇宙认为是一部巨型计算机。如果我们忽略相对而言非常渺小的、蜗居在宇宙中至少一个星体上的生命形式，似乎宇宙并没有涉及很多可计算的运动。那么，宇宙真的只是有一堆石头飞来飞去的时空吗？

我们从更宽阔的视角来看这个问题。现今的宇宙模型指出，宇宙开创于137亿年前的大爆炸，地球形成于45亿年前，地球上的生命始于37亿年前，最早的灵长类动物大概出现在1千万年前，而现代人类大概只能追溯到2百万年前。显然，有一些东西一直在促使着这个世界趋向于复杂。大爆炸后的最初期，宇宙是完全均匀的——一种简单的象征，然后出现较为复杂的粒子，最终原子、分子开始形成发展。这是一个由简单到复杂的过程，大概是基于相对简单的宇宙法则，这与元胞自动机有几分神似。

宇宙的计算模型一般归功于通信理论的奠基人香农和维纳对图灵机的贡献。使用熵度量信息构建了通信学和热力学之间的桥梁，这也是过去几年一些畅销书的主题。55例如，麦克斯韦妖（Maxwell's Demon），这是詹姆斯·克拉克·麦克斯韦（1831—1879）发明的一个假想精灵。用一个隔板将容器分成两格，这个精灵可以操作隔板上的一扇小门，使运动快的分子流入容器的一边，而让运动慢的分子呆在另一边，这样熵就减少了。这种想法被证明是不可行的，因为这个精灵自身也从系统中带走了部分熵。

55 Tom Siegned，*The Bit and the Pendulum: From Quantum Computing to M Theory — The New Physics of Information*（John Wiley & Sons，2000）。Hans Christian von Baeyer，*Information: The New Language of Science*（Harvard University Press，2003）。Charles Seife，*Decoding the Universe: How the New Science of Information is Explaining Everything in the Cosmos from Our Brains to Black Holes*（Viking，2006）。

美国物理学家约翰·阿奇博尔德·惠勒（1911—2008）把宇宙的存在形式和人类的感知联系在了一起。我们在观察的基础上问是或否的问题，并接收信息予以回答。对这个过程，用惠勒著名的三个词讲就是“万物源于比特”（it from bit）：

“万物源于比特”象征这样一个观点：物理世界的万物从根本上，最根本上，都有非物质的来源和解释。也就是说，我们所称的现实都得于对是否问题的分析和对仪器引起的响应的记录。简而言之，所有具有物理实体的东西都源于信息论的范畴，而这就是参与的宇宙。56

56 约翰·阿奇博尔德·惠勒，“Information, Physics, Quantum: The Search for Links”（1989） in Anthony J. G. Hey，ed.，*Feynman and Computation: Exploring the Limits of Computers*（Perseus Books，1999），311。

虽然提出宇宙是由信息构建的，但是惠勒拒绝接受宇宙是任何形式的机器的概念，因为它“还得明确或隐含地假设，存在超级计算机，存在预定的计划方案，存在执行某工作的设备，存在奇迹的事件，而这就会让宇宙陷入无穷的种类和无穷的数量中”。57

57 同上，314。但是惠勒引用了他写自1988年的另一篇论文。

另一个相当不同的观点来自大卫·多伊奇（1953—　），他是量子计算的先驱之一。多伊奇是“多宇宙理论”最为坚定的支持人。多宇宙这个概念最早来源于美国物理学家休·埃弗莱特（1930—1982）。我们所认为的波粒二象性矛盾就是发生着不同的量子事件的多宇宙间互相干涉的结果。我们所知的这个宇宙只是多宇宙中的一个可能实例。

1997年，多伊奇出版了《宇宙的构造》。在书中，他从四个互有交织的部分来解释宇宙的本质：

* 维也纳出生的哲学家卡尔·波普尔（1902—1994）所刻画的认识论；
* 休·埃弗莱特在量子物理框架下的多宇宙论；
* 英国自然学家查尔斯·达尔文（1909—1982）和生物进化学家理查德·道金斯（1941—　）描述的进化论；
* 图灵开创的计算理论。

当讨论虚拟现实生成器时，多伊奇使用了他称为“图灵原则”（Turing principle）的概念。一开始，图灵原则似乎是关于计算机制的：“**存在一种理论上的通用计算机，它可以模拟任何可能的现实物理实体的行为**。”多伊奇确认这种计算机可以模拟一切物理过程。很快，多伊奇就指出了这种计算机的计算能力就等同于创造一个虚拟现实的宇宙。图灵原则逐渐演化为一个更强的版本：“创造一个虚拟现实生成器，它的所有指令包含了现实可能具有的一切环境，这是可能的。”58显然，这也就隐含了虚拟一个我们生活的宇宙的可行性。

58 大卫·多伊奇，*The Fabric of Reality*（Penguin Books，1997），132-135。

MIT机械工程学教授塞思·劳埃德（1960—　）更愿意将自称“量子力学”的量子物理形容为“奇异”而不是多宇宙的，但他也用计算和信息论的观点描述这个宇宙：“宇宙大爆炸也是比特大爆炸。”劳埃德拒绝宇宙可以用图灵机构建模型的观点，“宇宙本质上是量子机制的，传统的数字计算机不能模拟量子机制的系统”。59这也是他认为量子计算机更适合这种任务的原因之一：

宇宙是一个物理系统，它可以等效地由量子计算机模拟，量子计算机在尺度上和真实的宇宙一样大。因为宇宙支持量子计算，可以等效地被量子计算机模拟，所以宇宙与一部通用的量子计算机的能力并无二致……从技术角度而言，对于宇宙是否就是一部量子计算机这个问题，我们现在可以给出一个确定的答案了：是的，宇宙就是一台量子计算机。60

59 塞思·劳埃德，*Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*（Alfred A. Knopf，2006），46，53。

60 同上，54-55。

量子计算机的一个特性是传统图灵机所没有的：由量子过程产生真正的随机数的能力。

在英国物理学家、数学家、著名数学软件Mathematica的创始人斯蒂芬·沃尔夫勒姆（1959—　）的作品中，元胞自动机作为一种宇宙物理定律的模型，再一次出现在公众面前，而他于2002年出版的一本极厚的、雄心勃勃而且非常畅销的书《新科学》（*A New Kind of Science*）又将此推向高潮。沃尔夫勒姆观察到元胞自动机是如何基于简单的规则而繁衍出非常复杂的模式的，受此启发，他将元胞自动机和图灵机的普遍性联系起来，以此说明它们都可以对物理过程建模。沃尔夫勒姆并没有在他的系统中引入量子机制，但是他表示他不需要引入量子机制是因为“我强烈地预感到，我讨论到的各类程序……最终都会展示出（即使不是全部，也是大部分）量子理论的主要特征”。61

61 斯蒂芬·沃尔夫勒姆，*A New Kind of Science*（Wolfram Media，2002），538。

在《新科学》中，沃尔夫勒姆在很多现实表现中找到了计算的普遍性，他定义了一种计算等价性的原则：

引入了一种新的自然准则，其中没有任何其他系统产生的计算能比元胞自动机和图灵机所做的计算更复杂……那么，我们大脑中所进行的抽象计算又是什么呢？它们更复杂吗？答案应该是否定的，至少当我们想知道确切的结果，而非泛泛空想时是如此。如果一个计算要被显式执行，那么它一定最终会被实现为一个物理过程，因此它必然会受到其他类似物理过程所受到的限制。62

62 同上，720，721。

一旦我们相信宇宙中所有的形态都是可以计算的（不管是通过传统的数字计算机还是量子计算机），那么万事万物就要遵守这样的准则。例如，生命就是可计算宇宙的一部分，生命里最神秘的形式——人类的思维也是一样。

很多世纪以来，哲学家、生物学家、精神学家，甚至是平民百姓都一直在追寻思维的本质。我们经常认为，我们身体的大部分机能都是各种器官中一系列物理和化学过程的机械结果，我们还没有将人类思维归为此类。我们感受到的思维是如此地特别。大脑显然和思维有一定程度的关系，但是，我们同样认为大脑并不等同于思维的全部。

在西方文化里，这种思想通常称为“意识/肉体二元论”，并且通常和勒内·笛卡儿（1596—1650），尤其是他的《形而上学的沉思》（1641）联系在一起。笛卡儿相信，我们身体的大部分器官（和所有我们称之为低等动物一样）都像机器一样，但是思维不一样。

20世纪40年代，二元论遭受重大打击。对于神经学家和计算机科学家迈克尔·阿尔贝勃而言，麦卡洛克和匹茨已经在1943年的论文中解决了神经元计算的问题。大脑天然有着适合进行计算的结构，因此麦卡洛克和匹茨“展示了所有可设想的有穷计算都可以被神经网络计算出来。他们否定了二元论”。63

63 Arbib，“Warren McCulloch's Search for the Logic of the Nervous System”，213。

几年后，哲学家吉尔伯特·赖尔（1900—1976）在他的著作《心的概念》（*The Concept of Mind*，1949）中建立了一个很强的事实，狠狠抨击了二元论，这个事实里并没有引用麦卡洛克和匹茨的论文。今天，二元论早已黯然失色。大多数研究思维的科学家们（包括哲学家、神经学家）都默认了思维仅仅是人体物理过程，尤其是神经系统和大脑物理运作的一种表现。

在二元论渐渐被否定的同时，我们对计算和算法的认识也在不停地增长，这并不那么让人吃惊。构想中的图灵机作为人类计算机的模型，可以执行被精确定义的算法任务，所以从自动计算这一门学科诞生开始，机器和大脑之间的联系就受到关注。同样不那么让人吃惊的是最早研究人工智能概念的人中就有阿兰·图灵本人。他在其1950年最著名的论文《计算机与智能》中，发明了今天称为“图灵测试”的测验。

一旦二元论被抛弃，思维就必然被看作是大脑物理活动（协同身体的其他部分）的一种自然表现，而不是什么超自然的东西。虽然我们在情感上有一丝排斥，但是结论是昭然的：首先，思维在能力和局限上等同于图灵机；其次，理论上完全可能制造人工的思维。

就像美国哲学家丹尼尔·丹尼特（1942—　）说的：“阿兰·图灵做出了基础性的开创，让我们得以将康德曾经提出的问题：怎么可能存在思维，转换成一个工程性的问题：怎么才能制造出思维。”64

64 丹尼尔·丹尼特，interview in Susan Blackmore，*Conversations on Consciousness: What the Best Minds Think About the Brain, Free Will, and What it Means to be Human*（Oxford University Press，2006），81。

图灵测试最让我们困扰的，同样也是大脑是计算机这种见解带来的困扰，是以第一人称一直在我们脑袋里喋喋不休的“意识”（consciousness）。意识让我们感到主观上的自主性和信仰的自由。

不过意识是难以捉摸、阴晴不定的。我们大多数人都会宣称自己在每天清醒的时候都会在心中自言自语，这让我们感到意识的存在；但当自言自语消失的时候，意识就是透明的。大多数人与人交流的时候，都假定对方有着和我们一样类似的意识，但我们并不确定，也不知道如何才能让对方意识到我们自己的意识的存在。

判断我们的大脑如何产生自我意识，是澳大利亚哲学家大卫·查默斯（1966—　）所谓的意识的“难问题”，而判断大脑如何与感知器官进行信 号的输入输出，是相比而言较为简单的问题。

图灵测试（让人类测试者觉得对方像人类一样聪明）隐含着一种行为主义的观点，即不必了解个体内部是如何运作的就能将其归入或排除出“智能”的类别。我们谈论的是一种“黑盒”测试。这也是我们如何与其他人交流的方式，因为我们不能证明其他人也是有意识的。即便我们不能分辨人与机器，也非常希望能够分辨机器和我们自己。

我们所知的计算机只不过是一部遵守一套规则的机器。它们不像人类知道自己做的是什么。在这个方面，美国哲学家约翰·塞尔（1932—　）做了一个著名的思想实验，也称为“汉语房间”。向一个不懂汉语的人提问，他有一本能让他给出合理答案的书，那么这个人可以通过汉语图灵测试，虽然他完全不懂问题或答案的意思。65

65 约翰·塞尔，“Minds, Brains, and Programs，”from *The Behavioral and Brain Sciences*，Vol. 3（Cambridge University Press，1980。Republished in Douglas R. Hofstandter and Daniel Dennett，eds.，*The Mind's I: Fantasies and Reflections on Self and Soul*（Basic，Books，1981），353-373。

最大的问题是电脑只懂得语法，而人类还懂得语义。在塞尔看来，这说明了数字计算机（不管它会变得有多复杂）永远不能像人类一样理解它们正在做什么。

英国数学物理学家罗杰·彭罗斯（1931—　）同样确信，思维不仅仅是一个计算器官的产物。在他1989年的《皇帝新脑》和1994年的《思维的影子》中，彭罗斯断言意识超出了计算的范畴。他猜测在大脑运行的是一种量子过程，这种量子过程不是算法式的，超出了图灵机的计算能力。

彭罗斯认为哥德尔的不完备性定理揭示了某种规律。我们人类能够理解哥德尔推导出的那些正确却不能证明的命题，但是任何计算都无法证明它，因为它并不是从公理衍生出的。这不是一个新的发现，早在1958年的《哥德尔证明》（纽约大学出版社）中，欧内斯特·内格尔和詹姆士·纽曼在哥德尔的定理中找到了类似的对机器智能的反驳，同样有类似发现的还有哲学家约翰·卢卡斯（1929—　）在其1961年著名的论文《心灵、机器与哥德尔》66中。这些论据都显示了，虽然机器能够容易地计算公理系统内的数学，但不能运行元数学，因为这需要对公理系统之外有所理解。

66 约翰·卢卡斯，“Minds, Machines and Gödel”，*Philosophy*，Vol. 36，No.137（Apr.-Jul. 1961），112-127。

丹尼尔·丹尼特也许是将哲学家的深思熟虑和科学家的实证主义结合得最好的人之一，他在《意识的解释》（1991）等一些精彩的书中，对思维有着不同的描述。丹尼特吸收了可计算性的概念，并将其融合到对进化和现代神经前沿研究成果的理解当中。他眼中的大脑和思维并不是像图灵机一样的媒介：大脑是神经系统的一部分，也是身体的一部分，不能隔离地讨论大脑。大脑中有一点兴奋的想法，心跳会加快，以便让更多的氧气进入大脑。很多药物可以影响大脑。大脑从眼鼻耳等其他器官中持续接受大量的刺激，不断地通过身体与这个现实世界交流。

大脑不是一个线性处理系统，它是大规模并行分散系统，没有像“笛卡儿剧场”（丹尼特诙谐地将笛卡儿的“意识的中心”称为“笛卡儿剧场”）那样有中心作用的区域。丹尼特的大脑模型是由思维的“很多草图”构成的，包含了感官输入、视觉数据、语言等不完整的支离破碎的部分。如果大脑是一台计算机，那么这也不是一台可以由正常工程师设计出来的计算机！因为它里面一定是混乱的。

进一步讲，我们所认为的意识其实是在这种并行结构之上的一系列活动。丹尼特提出：

假定人类的意识：(1) 所具有的创新力是不能被硬编码在机器中的；(2) 是早期经过人类文明训练的产物；(3) 能否成功建立取决于大脑的可塑性中无数的细微设置，也就是说，意识最重要的特征对于神经解剖学来说很可能是无法剖析出来的，尽管它们起了非凡的作用。67

67 丹尼尔·丹尼特，*Consciousness Explained*（Back Bay Books，1991），219。

至少在某种意义上，意识会和自己“说话”，而这就需要文明的产物——语言的支持。

显然设计一台能模仿人类思维的计算机是没有意义的，这需要输入很多的数据，而且如果没有多年的训练和积累的经验，效果也不会好。不过，理论上是否可以制造一台机器可以通过无限制图灵测试呢？（丹尼特认为无限制图灵测试是很难但很公平的测试）这样的机器有意识吗？丹尼特认为这两个问题的答案都是肯定的。

不管你更倾向于大脑是以什么样的机理进行工作的，一个令人胆寒的隐含结论是，机械运作的结果决定了我们的决定，而不是其他东西。那么我们所认为的自由意识（free will）又是怎么一回事呢？

自由意识在机械化运转的宇宙中消失了，这一看法早就暗含在了决定宇宙每个粒子运动的严格的确定性法则中。皮埃尔-西蒙·拉普拉斯（1749—1827）在他的《概率论》（*Essai Philosophique sur les Probabilites*，1814）中写道：

如果一个智能体理解某一时刻所有激发自然运动的力和组成宇宙的万事万物的各自状态，假如它能够在很宽广的空间分析这些数据，那么对于大到宇宙中的最大星体，小到最轻的原子，它对宇宙万事万物的运动的计算都会包含在一个公式中。对于它而言，没有什么事情是不确定的。未来就像过去一样，呈现在它眼前。68

68 皮埃尔-西蒙·拉普拉斯，*A Philosophical Essay on Probabilities*，译者Frederick Wilson Truscott 与 Frederick Lincoln Emory（John Wiley & Sons，1902，Dover，1995）。

这个观点通常称为拉普拉斯妖（Laplace's Demon）。我们很难避免这一推理：在大爆炸之后，宇宙中每个原子（包括组成大脑中细胞的那些）的运动就按照一种已经确定的模式固定下来。

当然，拉普拉斯妖并不真的存在。为了跟踪宇宙中每个粒子的运动，必须用一台比宇宙自身还大的计算机存储数据。海森堡测不准原理告诉我们，基础粒子的位置和时间不能同时确定。在数学上，把研究这些原子碰撞结果的问题归类为“多体问题”（many-body problem），而即使是3体问题的计算就足够让人头疼的了。

如果宇宙确实是一台图灵机，即使我们知道当前的“完全格局”以及这个机器具有的所有格局，还是不能够预测它未来的走向，除非真正地“跑”一遍“程序”。

不确定性是自由意识的基础。塞思·劳埃德指出，

停机问题不仅适用于传统的数字计算机，也适合于能进行数字逻辑运算的系统。因为粒子碰撞本质上进行的是数字逻辑的计算，所以它们的未来是不可计算的……我们面临抉择的时候所感知的主观随意性就类似停机问题：一旦我们脑中有一些想法，我们并不知道它会引领我们走向何方。即使它确实引领我们去了某个地方，在到达之前，我们也不知道是在哪里。69

69 劳埃德，*Programming the Universe*，98，36。

大卫·多伊奇仔细考虑了大脑是“经典”的非量子计算机，而不是量子计算机的可能性：

都说大脑可能是一台量子计算机，而且直觉、意识和我们解决问题的能力都基于量子计算。这可能是正确的，但是我没有看到任何证据或任何让人信服的论据，证明这是正确的。我的看法是，大脑如果被认为是计算机，那么它是一台“经典”的计算机。70

70 丹尼特，*The Fabric of Reality*，238。

然后他承认，“图灵对于计算的解释，即使是从原理上，也似乎仅仅给从物理角度进一步探索诸如意识和自由意识等精神属性留下了很小的空间”。记住，在量子物理的多宇宙理论里，世界是不断分裂的，在一个世界里你可以选择做这件事，而在另一个世界里你也可以选择做另一件事。如果这都不是自由意识，那么什么才是呢？多伊奇总结道：“图灵对于计算的概念似乎与人类的价值观不太相关，在多宇宙的框架下理解这些，对于我们认识人类主观意识等精神属性并没有阻碍。”71

71 同上，336，339。

斯蒂芬·沃尔夫勒姆在研究元胞自动机表现出的复杂结构时，他试图寻找预测结果的方法，或者至少能够找到可以减小繁衍代数而保持结果不变的捷径。但是他不能，“完全无法预测系统将会如何表现，除非像系统自身进化的过程那样一步一步地计算……对于很多系统，根本无法进行系统性的预测，也没有普遍意义上的进化捷径……”不可能进行有效预测这一事实给了系统以行使自由意识的自由，沃尔夫勒姆甚至还给出了一个图表，展示了一个“行为表现出类似自由意识的元胞自动机”。72

72 沃尔夫勒姆，*A New Kind of Science*，739，741，750。

这也算是一种慰藉了。即使宇宙和大脑像元胞自动机和图灵机一样，都是以一套简单的规则为基础，并繁衍出复杂的结构，我们依然无法基于这些规则预测未来。在我们运行到属于未来的那一行代码前，它并不存在。

就像在《回到未来》三部曲73的结尾中，布朗博士对麦克弗莱和帕克说的那样：“这说明你们的未来还未书写，每个人的未来都还未书写。你们的未来取决于你们如何打造它。所以你们二人要好好把握。”

73 编剧鲍勃·盖尔，根据其与罗伯特·泽米吉斯的故事和人物改编。