

# 图灵机 \*

张江(email: [jakezj@163.com](mailto:jakezj@163.com))

自然中的一切过程都有可能在计算，碰撞的小球、流动的溪水、燃烧的火焰，大自然用自己的方式处理着大量的信息。著名的 Mathematica 软件发明人沃尔弗莱姆 (Wolfram) 甚至宣称，整个宇宙就是一台大的图灵计算机。究竟什么是计算？什么是图灵机？计算与人类智能是怎样的关系？

## (一) 图灵与图灵机

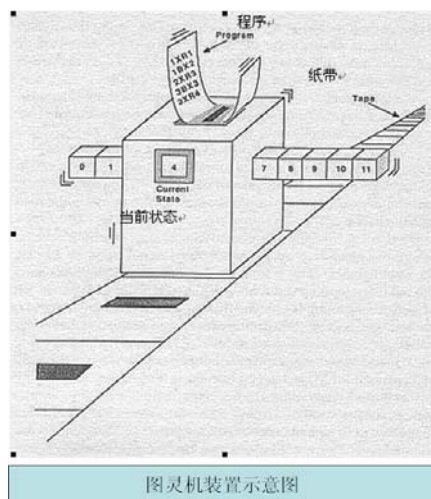
图灵机是计算机的理论模型，这个名字来源于它的发明人，阿兰·图灵 (Alan Turing)。



图灵 (1912~1954) 出生于英国伦敦，19 岁考入了剑桥皇家学院，22 岁就当选为皇家学会会员。1937 年，他发表了论文《论可计算数及其在判定问题中的应用》，提出了图灵机模型，后来，冯诺依曼就是根据这个模型设计出历史上第一台电子计算机的。1950 年，图灵又发表了划时代的文章：《机器能思考吗？》，成为了人工智能的开山之作。可惜的是，就在他的事业刚刚达到顶峰的时候图灵自杀了，享年仅有 42 岁。为了纪念这个伟大的学者，计算机界设立了最高荣誉奖：ACM 图灵奖。

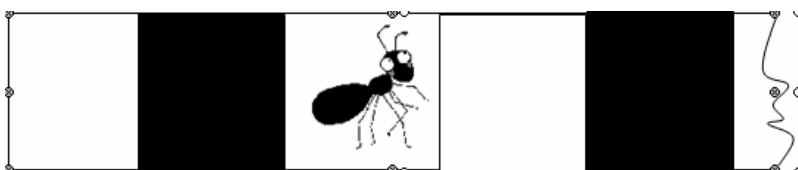
言归正传，我们开始讲图灵机的概念。你需要先认识一下它的轮廓，如右图：

这个装置由下面几个部分组成：一个被划分成方格的无限长的纸带，一个读写头。（中间那个大盒子），内部状态（盒子上的方块，比如 A,B,E,H），另外，还有一个程序对这个盒子进行控制。这个装置就是根据程序的命令以及它的内部状态进行磁带的读写、移动。也许这里的语言太抽象、死板，那么下面，我们用一个有趣的比喻让这个冷冰冰的家伙活起来。



### 1. 小虫的比喻

我们不妨考虑这样一个问题。假设一个小虫在地上爬，那么我们应该怎样从小虫信息处理的角度来建立它的模



\*\* 本篇文章介绍图灵机模型及其计算理论。\*号表示作者的推测。

型呢？

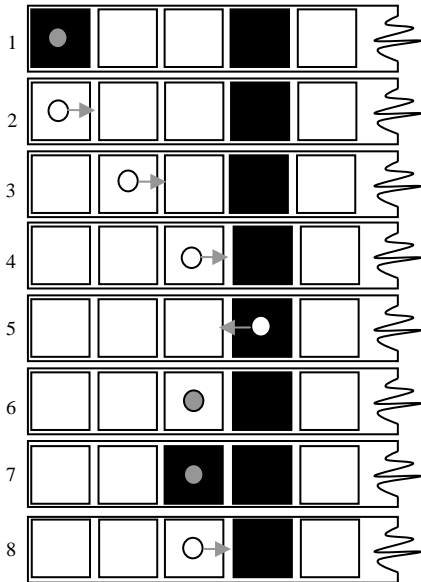
首先，我们需要对小虫所在的环境进行建模。我们不妨假设小虫所处的世界是一个无限长的纸带，这个纸带上被分成了若干小方格，而每个方格都只有黑白两种颜色。黑色表示该方格有食物，白色就表示没有。假设小虫仅具有一个感觉器官：眼睛，而且它的视力差得可怜，也就是说它仅仅能够感受到它所处的方格的颜色。因而这个方格所在的位置的黑色或者白色的信息就是小虫的输入信息。

其次，小虫有输出动作，它可以在方格上前移、后移，还可以涂写方格成黑色或者白色。最后，小虫还会有两种内部状态，即{饥饿，吃饱}。这样小虫的行动按照下面的程序进行：

程序：

输入	当前内部状态	输出	下时刻的内部状态
黑	饥饿	涂白	吃饱
黑	吃饱	后移	饥饿
白	饥饿	涂黑	饥饿
白	吃饱	前移	吃饱

即如果当前处于饥饿状态，则有食物就吃掉，没有食物就“自行解决问题”（即自己会吐出食物）；如果当前处于吃饱的状态，则如果没有食物就前移，如果有就后退，并且转入饥饿状态。那么当小虫子读入黑白白白……这样的纸带的时候，会怎样行动呢？让我们来看下面的图：



在图中，小虫用圆圈表示，它从最左边开始移动，灰色表示饥饿状态，白色表示吃饱状态。箭头表示移动的方向。从上到下，小虫一步一步地根据纸带的颜色和它自己的内部状态查找规则表中的对应项而采取行动。例如第 5 步读入方格是黑色，内部状态为吃饱，根据这两项输入信息查找规则表找到对应项是第二项，根据它，小虫应该后移，且内部状态变为饥饿。不难看到，到了第 8 步，情况跟第 4 步完全相同，输入都是白色纸带和饥饿状态，根据程序，小虫将重复 4~8 之间的动作，并一直持续下去……。

尽管从长期来看，小虫会落入机械的循环。然而当你输入给小虫白色信息的时候，它的反应可能完全不同（如第 4 步和第 6 步的行为），所以只要小虫子的内部状态和程序非常复杂，那么小虫的行为也会越来越超出你的想象！相信你已明白了这个小虫模型，那么你就掌握了图灵机的工作原理，因为从本质上讲，这个小虫模型就是一台图灵机！

## 2. 如何理解图灵机模型\*

这个模型是否太简单了？下面，我就要试图说服你，图灵机这个模型是伟大的！

首先，人都可以被抽象地看成一个图灵机。显然，输入状态集合就是你所处的环境中能够看到、听到、闻到、感觉到的所有一切，可能的输出集合就是你的每一言每一行，以及你能够表达出来的所有表情动作。内部状态集合则要复杂得多。因为我们可以把任意一个神经细胞的状态组合看作是一个内部状态，那么所有可能的神经细胞的状态组合将是天文数字！

似乎你会说，还有很多思维本质的东西没有概括进去，人有记忆，图灵机有么？其实，只要图灵机具有了内部状态，它就相应的具有了记忆。因为记忆的本质就是在头脑的内部存储了某种信息，这无疑可以表达成一个内部状态。比如上面讲到的具有饥饿和吃饱两种状态的小虫就会记住它的经历：如果吃到食物就用吃饱状态来“记住”食物的存在。

似乎图灵机模型中不包括学习，因为学习就意味着对程序的改变，而图灵机是不能在运行过程中改变它的程序的。然而，我们不难假设，你实际上并不能打开一个人的脑袋来看，所以它的实际程序规则你是不知道的。很有可能一个图灵机的规则非常多，在大多数情况只在少数几个规则里面转，但一旦不同的外部信息激活了它的某些内部状态，它就可能激活另外一些之前没有使用过的规则，这样它的行为就发生了本质上的变化。对于外界观察者来说，它似乎会学习了！而实际上，这个图灵机的程序一点都没变。

## (二) 图灵计算

### 1. 什么是计算

什么是计算？广义上讲，一个函数变换，把输入信息  $x$  变成输出信息  $f(x)$  就是一个计算！如果我们把一个小球扔到地上，小球又弹起来了，你完全可以把小球的运动都抽象成一些诸如位置、速度、形状等等信息，而地面把小球弹起来就无非是对小球的这些信息进行了某种变换，因而地面就完成了了一次计算！

按照这种理解，你会发现，现实世界到处都是计算了！因为我们完全可以把所有的自然界存在的过程都抽象成这样的输入输出系统，所有的大自然存在的变量都看作是信息，因而计算无处不在！正是采取了这样的观点，国外才有可能提出什么 DNA 计算机、量子计算机这些新鲜玩艺儿！因为人们把 DNA 的化学反应、量子世界的波函数变换都看作是计算了。

回到图灵机！为什么说图灵机是一个计算的装置呢？很简单，图灵机也是一个会对输入信息进行变换给出输出信息的系统。比如前面的小虫，开始时纸带上的黑白方格就构成了输入给小虫的信息，而小虫对纸带的涂抹后构成的黑白序列就是输出信息。但是，也许你会疑惑那个简单的小虫怎么可能完成复杂的计算呢？不要忘了，复杂性来源于组合！虽然每一次小虫的输出动作很简单，然而当把所有这些输出动作组合在一起，就有可能非常复杂！

### 2. 征服无限的方法！

可以说计算是一种用有限来应对无限的方法！例如：如果你拥有了  $2x$  这两个字符就相当于记住了无限个数对：(1,2), (3,6) ... (102,204)，例如输入给你一个 102，你会根据  $2x$  动态生成一个 204。

下面再考虑下面的三组数对，排列成一个表：

1,2	3,6	4,8	100,200
3,9	2,6	8,24	100,300
1,4	2,8	3,12	100,400

首先，每一行都可以用规则来描述，如第一行是  $2x$ ，第二行是  $3x$ ；进一步当把所有行都串起来后，我们能够发现这些规则本身也是有联系的，因此能够用规则的规则来描述。如果行数是  $n$  的话，那么这一行的规则就是  $(n+1)x$ 。我们显然能够根据输入的  $n$  和  $x$  计算出数值。不难看出， $(n+1)x$  是每一行的规则，也就是规则的规则。我们还可以把多个这样的表排列成三维的结构，这就需要规则的规则的规则加以概括，这个过程还可以继续下去……！这种规则的无限性意味着，我们似乎不可能找到一种终极规则，使得它可以产生所有的可能输入—输出关系对。

对于图灵机来说，虽然给定了规则之后，它可以完成计算，但是它却不能自发得到这些规则，所有的规则都是人编程告诉它的。而我们人似乎总能在一些个别的事件中通过归纳学习自动得出规则，即自动生成一些计算的程序。那么，机器可以归纳吗？

### 3. 归纳\*

金庸在他《倚天屠龙记》中曾讲述过这样一段故事，武林泰斗张三丰要将新创的太极拳传授给张无忌。张三丰打了两趟太极拳，可是张无忌越在那里揣摩太极拳的奥秘，忘记的招数却越多。旁人糊涂了，忘的这么快，怎么可能学会武功呢？然而，当张无忌说已经忘掉了所有的招式。张三丰却笑着说：“不错，你终于学会了‘太极拳’”。

张无忌学太极拳的过程就是从特殊的输入输出提升到了一般的计算过程。他已经能够从具体的一招一式之中抽象出了更高层次的武学规律，通常我们称这种能力为归纳学习法。用到图灵机模型中，这就相当于，图灵机可以通过输入一些特殊的输入输出数对，如 (1,2), (3,6), (5,10), (18,36)，而自动得到程序  $2x$ 。目前，科学家们已经研制出一些具有归纳能力的图灵程序，但是这些程序往往适用于一些特殊的场合。那么，是否存在某种通用的能够归纳出任何一种程序的图灵机呢？

如果这样的通用归纳程序  $P$  存在。那么，给  $P$  输入一些特殊的数对， $P$  就会输出能够生成这些数对的程序。也就是说输入具体的“招术”，输出的是这些“招术”的一般规律。如果说  $P$  已经设计出了，那么  $P$  就必然可以归纳出所有的规律，当然也包括  $P$  这个程序它自己。换句话说，程序  $P$  必须能够产生它自己， $P$  自己能把自己给归纳出来！这是否可能目前尚无定论，但可以肯定的是，这陷入了某种怪圈之中！

### (三) 超越图灵? \*

我们看到，图灵机都是按照固定的规则进行行动的机械装置，我们能否超越图灵计算呢？我认为，如果想彻底超越图灵计算的限制，我们必须放弃程序的实在性。也就是说程序在每时每刻都要变化成不是它自己了。那么这样的一个不断变化得不是它自己的怪东西存在么？

这个问题似乎是那个古老的赫拉克里特问题了：一个人不能两次踏入同一条河流。河流在每时每刻都不再是它自己了。河流是一大群流动的水滴构成的整体，在每时每刻这些水滴都在不停的运动、流逝，因而当你两次踏入这条河的时候，所有的水滴可能都不一样了，那么我们怎么能说这些水滴构成的整体还是同一条河呢？其实，我们人何尝不是这样一个不断流动、不断变化的个体呢？我们知道由于新陈代谢，每隔4年，人体的所有细胞就已经全部更新了，然而4年前的你还和4年后的你的确是一个个体。这意味着，所谓的生命、智能的本质不是一个一成不变的东西，而是一个永恒流动着的高级的空间和时间上的构型（Pattern）。

如果这样的理解是对的，那么我们就有了一条全新的创造人工智能的思路。即我们从最根本的超越图灵计算的机器入手。无论是专家系统还是神经网络，它们总有一个一成不变的规则，因此都是固定死的程序。但是我们又如何创造出一个总在变化的机器呢？近年来人们研究的涌现智能似乎已经非常接近我们的目标了。科学家用一群小机器人创造出一个动态的机器人群体，试图用类似简单蚂蚁构造复杂蚂蚁王国的思路来创造群体的智能。这个时候，无意识的小机器人组合到一起就有可能完成某种并不存在于个体之中的群体行为，这被科学家们称之为涌现。这样的涌现智能群体是否可能最终超越图灵计算呢？

也许，智能和生命的本质就存在于万事万物永恒不变的流动中。

参考文献：

Harry R. Lewis (1998). Elements of the Theory of Computation, Prentice-Hall, Inc

Michael Sipser, 张立昂译 (2000), 计算理论导引, 机械工业出版社

Wolfram(2002): A New Kind of Science, Wolfram Media.

侯世达(1995). 哥德尔、艾舍尔、巴赫——集异璧之大成. 郭维德等译. 北京：商务印书馆.

彭罗斯 (1993): 皇帝新脑, 湖南科学技术出版社