

关键词小程序 定向深度学习



请微信扫码

登录

新用户注册

首页	法学	经济学	政治学	国际关系	社会学	新闻学	教育学	文学	历史学	哲学	思想库	关键词	笔记	排行	导航
----	----	-----	-----	------	-----	-----	-----	----	-----	----	-----	-----	----	----	----

阿兰·图灵：计算机器与智能

选择字号：大 中 小 本文共阅读 5622 次 更新时间：2016-03-14 11:33:33

进入专题： [计算机器](#)

● 阿兰·图灵

小编按：图灵最伟大的论文当属1936《论可计算数及其在判定性问题上的应用》，该文提出图灵机概念，开创了计算理论，为今后的计算机科学奠定了基础。

对于人工智能，图灵1950《计算机器与智能》则更具意义。论文中提出的图灵测试至今还为AI界提供灵感。文中还有很多对反驳意见的辩护，在今天看来也还具有深刻意义。无疑，图灵用今日眼光看，是一位强人工智能拥护者。

【写在前面】本文是 A.M.Turing在 1950年创作，详细定义并解释了人工智能及其研究目的，发展方向，并驳斥了此前科学界及社会上普遍存在的反对观点，讲解通俗易懂，细致入微，有理有据，被称为人工智能科学的开山之作，直到现在仍有极重要的意义，几乎所有的人工智能教材都向读者强力推荐此文，读罢真的会让人切实感到，图灵不愧为计算机天才。

1.模仿游戏

我建议考虑这样一个问题：“机器能够思考么?”。要回答这个问题，我们需要先给出“机器”和“思考”的定义。我们可以用尽可能接近它们普通用法的方式定义这些词语。但是这种方式是危险的。如果使用这种方式，我们很可能会用盖勒普调查那样的统计方式来得出“机器能够思考么”这个问题的结论及其意义。显然，这是荒谬的。因此，我没有尝试给出一个定义，而是提出了另外一个问题。这个问题和原问题紧密相关，而且通过并不含糊的词语给出。

这个新的问题可以通过一个游戏来描述，不妨称之为“模仿游戏”。需要三个人来玩这个游戏。一个男人(A)，一个女人(B)和一个询问人(C)男女皆可。询问人呆在一个与另外两人隔离的屋子里。游戏的目标是询问人判断出外面的人哪个是男人，哪个是女人。询问人用标签 X，Y代表外面的两个人，游戏结束时，他要说出“X是 A，Y是 B”或者“X是 B，Y是 A”。询问人 C允许向 A 和 B提出下面这样的问题：

C： X，请告诉我你头发的长度。

现在假如 X实际是 A，那么 A必须回答。A在游戏中的目标是努力使 C做出错误的判断。他的回答可以是：

我的头发乌黑发亮，最长的一缕大概九英寸长。

为了排除声音帮助询问者得出结论，问题的答案可以写出来，打出来更好。理想的安排是使得两个屋子又远程打印通讯。也可以通过中间人传递答案。而 B在这个游戏中的任务是努力帮助询问者获得正确的答案。她的最优策略可能就是给出正确答案。她可以在她的答案中加入“我是女的，别听他的。”这样的词语。但是男人 A同样也能做出相似的评论。所以这并不能提供更多的帮助。

现在我们提出这样一个问题，“如果用一个机器担当 A的角色”，将会发生什么情况?同与两个人玩这个游戏相比，询问者判断错误的频率是否发生变化？

这个问题取代了原先的“机器能够思考么?”这个问题。

2. 对新问题的评析

[小程序](#) | [公众号](#) | [微博](#) | [手机版](#)



作者 

立即搜索

相同作者阅读

阿兰·图灵：计算机器与智能

相同主题阅读

阿兰·图灵：计算机器与智能

[>>更多相关文章](#)

热门专栏

秦晖	陈行之	郑永年	龙应台
丁学良	曹林	鄢烈山	傅国涌
陈嘉映	黄宗智	于建嵘	陈志武
徐贲	郭宇宽	马立诚	杨祖陶
向继东	沈志华	赵汀阳	戴建业
李昌平	张鸣	杨奎松	王海光
周濂	杨鹏	邓晓芒	王缉思
陈奉孝	郭世佑	马玲	王振东
狄马	袁伟时	史啸虎	熊培云
秋风	刘小枫	孟令伟	雷一宁
周枫	蒋兆勇	吴伟	沙叶新
葛剑雄	刘瑜	储昭根	吴稼祥
许之远	袁刚	吴励生	杨小凯
朱学勤	潘维	郑秉文	莫于川
羽之野	谢志浩	杨光	孙立平

你可能会问：“新问题的答案是什么？”。同样也可能问：“这个新问题真的有研究价值么？”

我们先解决第二个问题，因此不会进入一个无限的循环。

这个新问题的优势在于它把一个人的体力和智力完全区分开来。没有任何工程师或化学家宣称能够生产出和人的皮肤完全相同的物质。在未来的某天，这可能成为现实。但是尽管这样的发明是可能的，赋予一个“思维机器”人的皮肤对于让它更像人并不能提供更大的帮助。我们设置问题的方式考虑到了这个问题。因为我们让询问者无法看到，接触到或听到其它的游戏者。新的评判标准的其它优势在下面的样例问题和回答中显示了出来。

问：请写一首以 Forth Brige（译注：1890建成于英国，开现代桥梁建造的先河）为主题的诗。

A：我无能为力，我从来都不能写诗。

问：34957加70764等于多少？

A：（停了三十秒钟后给出答案）105621。

问：你玩象棋么？

A：玩。

问：我的王在 K1，没有别的棋子了，你只有王在 K6，车在 R1。该你走，你走哪步？

答：（十五秒钟的停顿后）车移动到 R8，将军。

这种问答形式几乎适用于我们想要包括的人类行为的一切形式。我们不希望因为一个机器不能在选美比赛中胜出就认为它无能；正如我们不能因为一个人不能在和飞机赛跑的比赛中胜出就认为他无能一样。我们的游戏设定让这些无能变得无关紧要。只要参与者认为合适，他们可以吹牛。可以把自己描述成魅力无限，强壮无比，英勇无畏。而询问人无法要求他们做实际的展示。

对这个游戏的批评可能是：机器在游戏的成败中占有了过多的比重。如果一个人试图假装成一个机器的话，他肯定是一个蹩脚的演员。他会由于算术方面的缓慢和不准确立即暴露。机器不会解决了一些应该是思考的问题，其解决方式却和人的方式相差甚远？这个反对意见确实很棘手。但是，至少我们可以说，尽管如此，机器可以被设计得可以很好的玩这个模仿者游戏。我们不需要对这个问题过分担心。

有一种情况是机器在玩模仿者游戏时的最佳策略是做别的事而不是模仿人的行为。这种情况是可能的。但是我认为这样的情况并没有付出多大的努力。不管怎样，本文并不试图研究这个游戏的理论。我们假定机器的最优策略是努力提供和人一样的答案。

3. 游戏中的机器

在 1 中提出的问题只有当我们确定“机器”这个词的意义的时候才是确定的。自然，我们希望一切工程科技都允许使用在我们的机器上。我们同样希望有这样一个可能：一个或一组工程师制造出一个可以工作，但是其工作方式却不能被它的建造者很好的描述，因为他们使用了一个基于试验的方法来设计它。最后，我们希望把通过生育诞生的人从“机器”的概念中排除出去。要让定义同时满足这三个要求是困难的。例如，你可能会要求这些工程师都是同个性别，但是这实际上也是无法满足要求。因为通过单个人的一个皮肤细胞产生一个完整的个体不是完全不可能的。这将是生物科技的具有划时代意义的突破。但是我们并不希望把它成为“建造了一个思维机器”。这就要求我们放弃允许一切的科技的打算。我们更愿意这样，因为目前“思维机器”的研究热点集中在一种特殊的，通常被称为“电子计算机”或“数字计算机”的机器上。因此，我们仅仅允许“数字计算机”参加我们的游戏。

第一眼看上去，这个限制过于严格。事实上并非如此。首先我们需要简要的了解一下这些计算机及其它们的一些性质。

你可能会说：万一数字计算机不能如我所愿，即不能在游戏中有所表现，把机器完全用数字计算机代替就只能令人失望，会像我们为“思考”定的评判标准一样。

目前已经有许多数字计算机在工作，人们可能要问：“为什么不拿一台直接作实验？很容易就能满足游戏的要求。同时用许多的询问者参加游戏，然后统计出判断正确的概率”。对这个问题的简要回答就是：我们并不是要问是不是所有的数字计算机都能在游戏中表现良好，也不是要问现

续费  
1.5  
会员  
上

在的计算机中有没有机器能够通过实验。而是要问：是否存在一台想象中的机器能够通过游戏。这仅仅是一个简要回答，后面将从另一个角度考虑这个问题。

#### 4. 数字计算机

数字计算机可以被解释成可以执行一切计算人员能够进行的操作。一个计算人员应该严格遵守规则；在一切细节上，都没有一丝偏离的权力。我们可以假设这些规则写在一本书上。每次被分配新的任务的时候，将会改变执行步骤。他有无数的纸进行计算，他可以用“算盘”进行乘法和加法运算，这并不重要。

如果使用上述解释定义数字计算机，我们可能陷入循环定义。为防止这种情况发生，我们列出期望结果得出的方式。一个数字计算机通常由一下三个部分组成。

(i) 存储器 (ii) 执行单元 (iii) 控制器

存储器用来存储信息，对应于计算员的纸。纸上既记录了他的计算过程，也记录了他所遵循的计算步骤。同样，如果计算员进行心算的话，一部分存储器将对应于他的记忆。

执行单元是一次计算中单个操作进行的场所。操作对象随着机器的不同而不同。通常一个相当长的操作可能是：“3540675445乘以 7076345687”，但是一些机器可能仅能够执行

“写下 0”一类的简单操作。

上面曾经提到计算员的“记录操作步骤的书”由机器中的一部分存储器代替。不妨把它们称为“指令列表”。控制器的职能就是保证这些指令按照正确的顺序得到正确的执行。控制器的设计使得这成为可能。

存储器上的信息通常被分为许多的小包存放。例如，在一个机器中，一个包由十个十进制数组成。每个存储信息的包都以某种系统的方式分配了数字。一个典型的指令可以是：“把存放在 6890的数加上存放在 4302的数相加并把结果存入后一个存储位置。”

显然，机器中的指令不会用英语标示。这个指令更有可能用 6809430217这样的形式来表示。这里，17指定在可以对这两个数进行的所有操作中的哪个被执行。这里的操作如上所示是“进行加法操作。。。”请注意，这里的指令共占用了 10个数字因此正好可以放在一个存储空间中，非常方便。控制器保证指令按照它们的存储顺序被执行，但是偶尔可能碰到下面的指令：

“现在执行存储在 5606的指令，并从那里继续执行。”

或者是：

“如果 4505位置是 0，那么执行存储在 6707的指令，否则继续”。

后面这种指令非常重要，因为它使得我们能够反复执行一段指令直到满足一些条件。要达到这个目的，在每次执行时不要改变指令，而是一遍一遍的执行相同的指令。可以用家庭生活情况做一个类比：如果妈妈想让汤姆每天上学时都到修鞋匠那里看看她的鞋是不是修好了。妈妈可以每天都告诉他一遍。另一种方式是，在一个汤姆每天上学都能看到的地方贴各个便条告诉他到鞋匠那里去看一下，当汤姆拿回他的鞋的时候，撕掉那个便条。

读者必须接受计算机可以而且事实上正是按照我们所提出的原则建造的。这些计算机几乎能够完全模仿一个计算员。

上面我们提到的计算员要遵守的写在书上的步骤实际上仅仅是一个设想。一个实际的计算员总是能够记得他们需要做什么。如果一个人想让机器模仿计算员执行复杂的操作，他必须告诉计算机要做什么，并把结果翻译成某种形式的指令表。这种构造指令表的行为通常被称为“编程”。“给一个机器编程使之执行操作 A”意味着把合适的指令表放入机器以使它能够执行 A。

一个数字计算机的一个有趣变化是让其成为“带有随机元素的数字计算机”它们有特定的指令进行掷色子或者别的等价电子过程。例如其中一个“掷色子并把结果存入 1000。”“有些时候这样的机器被描述为具有自由意志。（尽管我自己并不使用这个表述）通常并不能通过观察判断出一个机器有没有随机元素。因为一个相似的效果可以通过依据 Pi的小数部分进行选择来产生。

绝大多数的数字计算机具有有限的存储空间。让一个计算机获得无限的存储空间并不存在理论困难。当然，在任何时候都只有有限的不被使用。同样，只有有限的存储空间被建造。不过

我们可以想象越来越多的存储空间可以根据要求添加。这样的计算机具有特殊的理论价值，被称为无限容量计算机。

有关计算机设想的很早就产生了。1828至 1839年担任剑桥大学的 Lucasian数学教授的 Charles Babbage（查尔斯·巴贝奇）设想这样的机器，并称之为分析机。但是并没有完成它。尽管 Babbage有了所有的关键思想，他的机器在那个时代却没有吸引人的前景。它能够达到的运算速度肯定比一个计算员要快但是仅相当于曼彻斯特机的百分之一，而曼彻斯特机也是现代计算机中相当慢的一个。分析机的存储全部由用轮子和卡片组成的机械实现。

事实上，全部由机械实现的 Babbage分析机帮助我们破除了一个迷信。通常我们对现代计算机是电子的这个事实给予相当的重视。同样，神经系统也是基于电的。因为Babbage的机器是没有使用电，而所以的数字计算机在某种意义上都是等价的。我们就会明白使用是否使用电理论上并不重要性。当然，当需要高的运算速度的时候，需要使用电气结构。因此，我们就会发现在这些地方使用电是理所当然的。在神经系统中，化学过程至少和电过程同样重要。某些计算机的存储器主要基于声学原理。计算机和神经系统都使用电仅仅是表面的相似。如果我们希望寻找这样的相似，倒不如对函数进行数学分析。

### 5. 数字计算机的通用性

上一部分给出的数字计算机可以被归类为“离散状态机”。这类机器可以从一个确定状态向另一个状态突然跳变。为了不会有混淆这些状态的可能，它们之间要有足够的差别。严格的说，这样的机器是不存在的。一切事件实际上都是连续的。但是有许多种机器能够被看作离散状态机器。例如在照明系统中的开关，我们可以把开关看成只有开和关两个状态。它们之间肯定有转换过程，但是在绝大多数情况下可以忽略它们。作为离散状态机器的例子，我们可以考虑一个每秒滴答 120度的轮子，这个轮子可能因一个杠杆的阻挡而停下来，在轮子上有一个发光的灯。这个机器可以被抽象的描述为下面的形式。机器的内部状态 (通过轮子的位置来描述)可以是  $q_1$ ,  $q_2$ 和  $q_3$ 。输入信号是  $i_0$ 或  $i_1$ (杠杆的位置)。任何时候的内部状态可以根据上一次状态和输入信号由下表描述：

输出信号可以用下表描述，它是唯一能够被外部观测的内部状态指示器。

这个例子是一个理想的离散状态机。只要它们的状态是有限的，它们就可以用这样的表格描述。

可以看出，只要给出初始状态和输入信号，所有的未来状态都是可以预测的。这让我们想起了拉格朗日的观点。那就是，只要给出宇宙中任意时刻所有粒子的位置和速度，就能够预知未来的所有状态。我们考虑的预测和拉格朗日相比较更接近于实用性。因为“宇宙是一个整体”的系统，所以只要初始条件的一个很小的误差，就会引起系统在一段时间以后变得完全不同。某个时刻一个电子位置的亿万分之一厘米的偏移，将决定一个人会在雪崩中死去还是逃生。我们称为“离散状态机”的机械系统的一个核心性质就是，这样的现象并会发生。即使是考虑一个实际的物理系统，而不是理想机器，只要知道了某个时刻足够精确的状态信息，就可以足够精确的预测未来的一切状态。

正如我们所提到的，数字计算机属于离散状态机。但是这样的机器所能够达到的状态通常是相当大的。例如，现在在曼彻斯特工作的机器可以有  $2^{165000}$ 个状态，也就是大约  $10^{50000}$ 。而在上面的滴答轮的例子中，仅有三个状态。找到有如此多状态的原因并不困难。计算机具有一个存储器，对应于计算员的纸。这些存储器中应该能够写入任何能够写入计算员所用纸上的符号。为了简单起见，假设仅仅实用从 0到 9的数字作为符号。忽略手写体的差别。假如计算机具有 100张每张 50行，每行 30个数字的存储空间。那么状态的数目就会是  $10^{(100*50*30)}$ 即  $10^{150000}$ 。这大约是三个曼彻斯特机状态的总和。状态数的基为 2的对数通常被称为机器的“存储大小”。因此曼彻斯特机的存储大小是 165000而上面例子中轮子的存储大小是 1.6。如果两个机器加在一起，他们的存储大小应该是原来存储大小的和。因此我们可以说“曼彻斯特机具有 64个磁带存储器每个存储器的大小是 2560，还有 8个电子管，每个容量为 1280。各种各样存储器加在一起大约是 300个，总共 174380。

只要给出对应于离散状态机器的表格，就能够预测出机器将会做什么。这样的计算当然能够通过电子计算机进行。只要执行的足够快，电子计算机就能够模拟任何离散状态机的行为。这样，模范者游戏就变成机器作被问者(B)，善于模拟的计算机作 A，那么询问者就不能区分它们。当然，数字计算机必须有足够的存储空间，运行的足够快。而且它在模范不同的机器之前必须被重新编程。

数字计算机可以模拟任意离散状态机器的性质被表述为“他们是通用机器”。具有这样性质机器的存在带来的一个重要结果就是，不考虑速度，我们并不需要设计出不同的新机器来执行不同的问题。它们都可以用一个数字计算机来实现，只要根据没种情况进行适当的编程。可以看出，这表明所有的数字计算机在某种意义上是相互等价的。

现在重新考虑在第三部分末尾提出的问题。暂时把问题“机器能够思考么？”用“是否存在假想中的机器能够在模范游戏中表现良好”代替。如果我们愿意，我们还可以问“是否存在能够表现良好的离散状态机”，但是由于通用性，我们可以看出这两个问题都等价于“让我们把注意力集中在一个数字计算机 C 上。如果我们可以让它具有足够大的存储空间，足够快的计算速度，而且对它进行适当的编程。C 扮演模仿游戏中 A 的角色，人扮演 B 的角色，C 能不能在这个游戏中表现良好？”

## 6. 主要问题的对立观点

现在，我们认为基础已经打好，并准备就“机器能否思维？”这个问题继续进行辩论……我们不能全盘摒弃最初的那个问题形式，因为在考虑代换是否恰当这个问题时，会有不同的意见，因此，我们至少必须考虑这方面必须说些什么。

我想要是我先就这个问题阐述一个自己的看法，那么，读者就会觉得问题会更简单一点。

首先让我们来看看这个问题的更确切的形式。我认为在 50 年的时间里，计算机的信息存储量可达到大约  $10^9$ ，这样的话，计算机在模拟游戏中就会很顺利，一般的提问者在 5 分钟提问后，能准确鉴别的概率不会高于 70%。原来那个“机器能够思维吗？”的问题我认为没有什么意义，不值得讨论。然而，我认为本世纪末，由于词汇用法会有较大的变化，普遍的学术见解也会改变，那时候人们又能重新谈思维机器而不会自相矛盾。我还认为，如果掩盖这些看法，决不会带来任何益处。人们普遍认为，科学家进行科学研究工作，总是从可靠的事实到可靠的事实，从来不受任何未经验证的假设的影响。这种看法实际上是错的。假如能清楚地划分哪些是经过证实的事实，而哪些又是未经验证的假设，那也没有什么害处。假设往往是非常重要的，因为它们为研究暗示了有益的研究方向。

我现在谈谈与我的看法相对立的观点。

(1) 来自神学的反对意见。思维是人的不朽灵魂的一种功能。上帝赋予每一个男人和女人以一颗不朽的灵魂，但从未将它赋予任何其他动物或机器。因此，动物或者机器不能思维。

尽管我不能接受这种看法，但我试图用神学的语言来回答这个问题。如果将动物和人划为一个类别，我认为这个观点更有说服力；因为，在我看来，生物与非生物之间的差别远远要比人和其他动物之间的差别大得多。如果这种正统的观点出现在其他宗教社会里，它的主观武断性就更加明显了。伊斯兰教认为妇女没有灵魂，基督教对此有何感想？但是，现在暂不管这一点，让我们回到问题的焦点上来。在我看来，上面所引的论点对上帝威力的万能性有很大的限制。上帝对有些事情也是无能为力，比如，无疑他不能让 1 等于 2；但是，我们难道不应该相信，要是上帝觉得合适，他完全可以赋予一头大象以灵魂吗？我们可以希望，上帝可以通过自己的威力造成变种，由于变种，大象有了一个较发达的大脑，可以用来满足灵魂的需求。同一形式的论点也可以用来解释机器。只是看上去也许有点不一样，因为“轻

信”就不那么容易了。但这其实只能说明，我们认为上帝不太可能觉得这些环境适合于授予灵魂。关于环境问题我们将在本文的其余部分进行讨论。在企图制造这样的机器时，我们不应该无礼地篡夺上帝创造灵魂的权力，就像不应该剥夺我们生儿育女的权力那样；在两种情况下，我们其实都是上帝意志的工具，为他所创造的灵魂提供住所。

然而，这仅仅是猜想而已。不管能用神学的论据来证明什么，我仍不以为然。过去已经证明这样的论据是漏洞百出的。在伽利略的时代，就有人提出，“太阳一动不动地悬着……整天都不慌

不忙，不想落下”（《约书亚书》，10.13）以及“他为大地奠定基础，叫它永远不动摇”（《诗篇》，105.5）这些经文正好用来驳斥哥白尼的理论。从我们今天的知识来看，就觉得提出这种论据是徒劳的。过去没有这样的知识，情况便大不一样了。

（2）所谓“鸵鸟政策”式的异议。“机器思维后果太令人恐惧了。但愿机器永远不会有思维。”

这种观点不如上面的说法那样直言不讳。但它对我们许多人都有影响。我们都倾向于认为人类的某个微妙的方面比其他生物要优越。要是能证明人一定是高一等的，那再好不过了，因为那样的话，他高居一切之上的地位就不会有危险了。神学的论点那样流行，很明显是与这种情绪密切相关。这种看法在知识分子中会更普遍，因为他们比其他人更尊重思维能力，因此也就对人类思维能力的优越性更加深信不疑。

我认为这个论点不怎么重要，不值得一驳。安慰一下也许更合乎情理；这种安慰也许能在灵魂轮回说中找到。

（3）来自数学的异议。在数学逻辑里有一些结论，可以用来证明离散状态的机器的能力有一定限度。这些结论中最著名的是哥德尔定理，此定理声称，在任何一个足够有力的逻辑系统里，都能形成陈述，而所作陈述在本系统范围之内既不能被证明是对的，也不能被证明是错的，除非这个系统本身就不一致。丘奇、克利恩、罗瑟和图林等人也有别的在某些方面同哥德尔定理很相似的结论。图林的结论更容易考虑，因为它直接涉及机器，而其他人的结论相对来说是间接的：比方说，如要使用哥德尔定理，我们还需要某些附加手段，通过机器来描述逻辑系统，而要描述机器还需要再通过逻辑系统。这个结论涉及一种机器，它实质上是一台万能的数字计算机。即使是这样一台机器，它对有些事情也是无能为力的。如果计算机被设计成能在模拟游戏中回答问题的话，那么对有些问题它是无法给予正确答覆的，而对另一些问题，不管你给它多长时间，它也答不上来。当然，尽管这台机器回答不了许多问题，但另一台机器却能给予满意的解答。我们现在只假定，对这些问题只要回答“是”或者“不是”就行了，不会出现像“你认为毕加索怎么样？”这类的问题。我们知道机器必定无法回答的问题是下述这类问题：“这台机器有以下特点……这台机器会不会对任何问题作出‘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，我们就能知道，答案不是错了，就是压根儿没有答案。这就是数学的结论，此结论认定机器能力有限，而人类智能则没有这种局限性。

如果想对这个论点作一简要的答覆，我们就能指出，尽管它已经证明任何一台特定的机器都是能力有限的，但它并没有任何证据说，人类智能就没有这种局限性。但我认为这个论点不能就这么轻易了结。每当其中一台机器遇到一个合适的问题，并作出我们明知是错的回答时，我们无疑会产生一种优越感。这种优越感难道不会是错觉吗？这种感觉无疑是真实的，但我觉得这并没有多大意义。我们自己平时也经常对问题作出错误的回答，因此，就没有权利因机器犯了错误而沾沾自喜；当然，我们对付一台机器当然易如反掌，但我们无法同时对付所有的机器而且不出差错。一句话，有可能人比一台特定的机器聪明，但也有可能别的机器更聪明，如此等等。

我认为，那些持数学异议的人大多数愿意接受模拟游戏作为讨论的基点。而持前两种反对意见的人不大会对什么标准问题感兴趣。

（4）来自意识的论点。这个论点在斐逊教授 1949 年的李斯德演说中阐述得很明确，我摘引了其中的一段话：“若我们要承认机器与大脑是一样的，除非机器能够因为感受了思想与感情，而不是符号的偶然涂抹写出十四行诗或协奏曲来。也就是说，它不仅写了，而且也应知道自己确实这样做了。任何机器都感觉不到（不只是属于简易发明之类的人工信号）成功的喜悦，也不会因困难而沮丧，因受奉承而沾沾自喜，因犯错误而闷闷不乐，因见异性而神魂颠倒，也不会因欲望得不到满足而暴跳如雷或一蹶不振。”

这个论点看上去否定了我们测试的有效性。按照这种观点的最极端形式，你若要肯定一台机器是否能思维，唯一的途径就是成为那台机器并且去感受自己的思维活动。这样的话，他就能向众人描述他自己的感觉，当然，没有人会知道这些话是不是实话。同样，依照这个观点，要想

知道某人是否在思维，唯一的途径就是成为那个特定的人。这实际是唯我论的观点。这也许是符合逻辑的，但若真是这样，那思想交流就太困难了。A会说，“A在想，而B不在想”。而B会说，“B在想，而A不在想。”我们犯不着为这个观点争执不休，我们不如客客气气地认为大家都在想。

我肯定斐逊教授不愿意采纳这样一个极端的唯我论的观点。他倒有可能愿意把这个模拟游戏当作一个测试。模拟游戏（省略了游戏者B）在实际中经常采用“口试”形式，来鉴定某人的真的理解了，还仅仅是“鹦鹉学舌”，让我们看一看这种“口试”的情形：

提问者：你的十四行诗的第一行是这样的，“我欲比君为夏日，如何？”要是将“夏日”改成“春日”，是不是也可以，或许会更好？

证人：这样一改便不合韵了。

提问者：改为“冬日”怎么样？这样也会合韵。[1]

证人：是没问题。但是有谁愿意将自己比作冬日呢？

提问者：你认为匹克威克先生会使你联想起圣诞节吗？

证人：在一定程度上，会的。

提问者：但是圣诞节是在冬天，我认为匹克威克先生对这个比喻不会在意。

证人：我在想你也许在开玩笑。冬日的意思是指某一个典型的冬日，而不像圣诞节那样特殊的一天。

这里不再赘引。如果那台写十四行诗的机器在这场“口试”中能够这样对答，杰斐教授会作何感想呢？我不知道他是不是会认为，机器只是在“机械地通过发信号”而作出这些答复；但是，如果这些答复如上面所引那样令人满意，前后一致，我认为他不会再把机器当作“一个属于简易发明之类的人工信号”。所谓“属于简易发明的人工信号”指的是一架机器的一些设计功能，可以用来播放一个人念十四行诗的录音，只要拨动键钮，你随时都可以听到这段录音。

综上所述，我认为大部分支持来自意识的异议的人都可以经过劝说而放弃原来的主张，不致于陷入唯我论的困境。这些人因此也就有可能愿意接受我们的测试。

我并不想给大家留下这样一个印象，即我认为意识并没有什么神秘之处。比方说。要想确定意识在人体中的位置，就是一个谜。但是，如果我们没有解决这些谜，还是能够回答这个与本文有关的问题的。

（5）来自种种能力限制的论证。这些论证一般是这样一种说法：“我担保，你可以使机器干任何你刚才提到的事情，可你永远也不能使一台机器有X类的行为。”这类行为包括许多特征。我在这里援引一例：

要和蔼、机灵、美丽、友好……富于首创精神、富于幽默感、善于明辨是非、会犯错误……会坠入情网，喜欢草莓和奶油……能使别人钟情于它，通达世故……措词得当，长于反思……像人一样行为多姿多彩，富于创新……

说这些话一般都用不着证明。我认为这些话都是以科学归纳的原则为基础的。一个人在他一生中看到过成百上千台机器。他由所见所闻得出一些普遍的结论。它们形态丑陋，应用范围狭窄，只要范围略有变动，它们就束手无策。此外，它们的行为方式也非常单一，等等，等等。他很自然地认为，这就是机器大体上的必备特征。大部分机器的能力限制与机器储存量太小有关（我在设想，储存量这个概念可以通过某种方式进行扩充，它不仅仅包括离散状态的机器，也包括其他机器。因为目前讨论还不需要讲究数学的准确性，所以定义准确问题并不要紧）。数年前，由于数字计算机在社会上鲜为人知，要是你光说其特征而不提其构造，那么，你们就会以为你在信口开河。我想这也是因为人们使用了科学归纳原则的结果。当然，人们在使用这个原则时，大都是无意识的。一个小孩一朝被火烫，十年怕烛台，我认为他这就是在使用科学归纳（当然，我也可以用许多别的方式来解释这一现象）。人类的行为和习惯看上去不适合运用科学归纳。如果你想获得可信赖的结果，你就要对时空的大部分进行研究。要不然的话，我们会（就像许多说英语的儿童那样）以为世界上所有的人都讲英语，再去学法语真是傻透了。

然而，关于我们刚才提到的许多能力限制，还要特别说几句，说机器没有能力享受草莓和奶油，这种说法会使读者觉得有点轻率。我们有可能使机器喜欢这些美味，但任何强迫这样做的企

图都是愚蠢的。值得重视的是，这种能力限制对解释其他能力限制也有影响。比如，难以使人与机器之间形成那种像白人与白人之间，或是黑人与黑人之间的友好情感。

认为“机器不会出差错”这种想法有点令人费解。我们不禁要反问：“它们出了错就更糟了吗？”让我们站在同情的立场上，看看这究竟是什么意思。我想我们可以用模拟游戏来解释这种说法，有人声称，在游戏中提问者可以向被试问几道算术题来分辨是哪个机器，哪个是人，因为机器在回答算术题时总是丝毫不差。这种说法未免太轻率了。（带模拟游戏程序的）机器并没有准备给算术题以正确的答案。它会故意算错，以蒙骗提问者。机器在做算术题时，由于对要出现什么样的错误作出了不妥当的决定，因而显示了机械的故障。我们对这种观点作这样的理解，其实也不太富有同情心。但是，我们对这个问题限于篇幅不可能进一步讨论。在我看来，这个观点的根源在于混淆了两个不同性质的错误。这两个错误我们称之为“功能错误”和“结论错误”。功能错误是由某些机械或电器故障引起的，这些故障导致机器不能够按照指令工作。在进行哲学讨论时，我们很容易忽视发生这种错误的可能性；这样的话，我们实际上是在谈论“抽象的机器”。而这些抽象的机器与其说是实在的物体倒不如说是数学的虚构。从定义上讲，我们完全可以这么说：“机器从不出差错。”当某一意思与来自机器的输出信号联在一起时，就会产生结论的错误。比方说，机器能够自动打出数学方程或英语句子。当机器打出一个错误的命题时，我们就认为这台机器犯了结论错误。很明显，找不到丝毫理由说，机器从不犯这类错误。一台机器有可能别的什么也不能做，只会连续打出“0=1”。举这样一个例子也许太过分了，我们可以换一个例子：机器会想办法通过科学归纳来得出结论。这种办法有时无疑会导致错误的结果。

有人说，机器不能成为它自己思维的主体。如果我们能证明机器的某些思维是有某些主题的话，我们就能驳回这种说法。尽管如此，“机器活动的主题”确实有点意义，至少对于研究它的人来说是这样的。比如，如果一台机器试图解  $X^2-40X-11=0$  这个方程式，我们不禁会认为，这时，这种方式本身就是机器主题的一部分。从这种意义上说，机器无疑能够成为它自己的主题。这对编排它自己的程序，对预测因本身结构变化带来的后果都会有所帮助。机器能够通过观察自己行为的结果，修改自己的程序，以便更有效地达到某种目的。这并不是乌托邦式的空想，而是不久的将来可能办到的事。

有人批评说，机器的行为比较单一。这也就是说，机器不能够有很大的存储能力。直到最近，达到 1000 字节的存储量都很罕见。

我们现在这里考虑一些反对意见实际上是都来自意识的那个异议的改头换面。通常，如果我们坚持认为，一台机器有能力做完其中的一件事，并对机器所能采用的方法进行描述。那么，不会给别人多深印象。人们会认为机器所使用的方法（不管是什么方法，总是机械性的）实在太低级了。请参见前面所引斐逊演讲中括号内的话。

（6）洛夫莱斯夫人（译注：爱达，第一个程序员）的异议。

洛夫莱斯夫人的回忆录中曾对巴比奇的分析引擎作过详尽的记述。她这样写道：“分析引擎没有任何意图要想创作什么东西。它能做我们知道该怎样去指挥它做的任何事。”（重点为她所知）哈特里引用了这段话，并补充道：“这并不是说，就不可能制造能‘独立思考’的电子设备，（用生物学的说法）在这种设备里，我们能够引起可以用来作为‘学习’基础的条件反射。从最近的一些发展情况

看，这种设想从原则上说是不是有可能，已经引起人们的极大兴趣和关注。但是，当时制造的那些机器并不具备这些特点。”

在这点上我完全同意哈特里的看法，我们会注意到，他并没有断言当时的机器还不具备这个特点，他倒是指出了，洛夫莱斯夫人所能获得的证明还不足以使她相信这些机器已具备了这个特点。从某种意义上讲，这些机器已具备了这个特点，这是极有可能的，因为，我们可以设想某些离散机器有这个特点。分析机实际上是一台万能数字计算机。因而，如果它的存储能力和速度达到一定水准，我们就能通过适当的程序使它模仿我们讨论的机器。也许伯爵夫人或巴贝奇都没有想到这一点。无论怎么说，他们不必能提什么要求，便提什么要求。



洛夫莱斯夫人的异议还有另外一种说法，即机器“永远也不能创新”。这种说法可以用一句谚语“世上无新事”抵挡一阵。谁能保证，他的“独创性成就”就不是接受教育的结果，就不是因循著名的普遍原则的结果？这种异议还有另一个稍稍好一点的说法，即机器永远也不能“使我们惊奇”。这种说法有点直截了当，我能够针锋相对地加以反驳。机器经常令我吃惊。这主要是由于我对机器能做什么估算甚少，更确切地说是由于即使我做了估算，也匆匆忙忙，马马虎虎。我也许这样对自己说：“我认为此处的电压应与彼处相同；不管怎么样，就当是一样吧。”我自然经常出错，结果我大吃一惊，因为一俟实验完成，这些假设早被忘得一干二净了。我对自己的这些错误开诚布公，但在我证实了所惊讶的事时，人们也不会认为我在信口开河。

我的回答并不会使批评者就此缄口沉默，他也许会这样做，所谓大吃一惊都是因为我自己富于想象力的心理活动，与机器本身毫不相干。这样，我们重又回到来自意识的那个论证上去，而背离了吃惊不吃惊的话题。我们不得不认为这种论证方式是封闭式的，但是，也许值得一提的 是，要将某物认作会使人惊奇的，则需要许多“富于想象力的心理活动”，不管这件令人吃惊的事件是由一个人、一本书、一台机器还是任何别的东西引起的。

我认为，那种认为机器不会令人吃惊的观点是由这样一个为哲学家和数学家特别容易犯的错误引起的。它是这样一个假设，即心灵一接受了某个事实，由此事实所引起的一切后果都会同时涌入心灵。在许多情况下，这种假设十分有用，但是，人们会情不自禁地忘了这是个错误的假设。如果照这样做的话，其必然结果就是认为，仅仅从数据和普遍原则得出结论会毫无效力可言。

进入专题：[计算机器](#)

快来抢!

快来抢!

本文责编：[陈冬冬](#)  
发信站：爱思想（<http://www.aisixiang.com>），栏目：天益学术 > 哲学 > 科学哲学  
本文链接：<http://www.aisixiang.com/data/97818.html>

分享到新浪微博：

5

推荐

寄送给好友：

立即发送

在方框中输入电子邮件地址，多个邮件之间用半角逗号（,）分隔。

爱思想（aisixiang.com）网站为公益纯学术网站，旨在推动学术繁荣、塑造社会精神。

凡本网首发及经作者授权但非首发的所有作品，版权归作者本人所有。网络转载请注明作者、出处并保持完整，纸媒转载请经本网或作者本人书面授权。

凡本网注明“来源：XXX（非爱思想网）”的作品，均转载自其它媒体，转载目的在于分享信息、助推思想传播，并不代表本网赞同其观点和对其真实性负责。若作者或版权人不愿被使用，请来函指出，本网即予改正。

