《计算机器与智能》阿兰-图灵阿兰-图灵

【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。

1. 对新问题的评析

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。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。

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。

1. 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。

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。

1. 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。

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）存贮器 （ii）执行单元 （iii）控制器

**存储器用来存贮信息**，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。

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。例如其中一个是“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分被使用。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，将被称为无限容量计算机。

附加信息：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。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。

1. 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。

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。

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。

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。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？”

1. 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的动物或机器。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。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有关。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：即使是一台机器，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似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。

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似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。

（８）来自行为变通性的论证。我们不可能总结出一套规则来囊括一个人在所有可想象的环境中的行为。比方说，我们可以有这样一条规则：行人见到红灯止步，见到绿灯行走，但是， 由于某种错误，红绿灯同时亮了，那该怎么办？我们也许会这样决定，为安全起见最好止步。但是，这个决定还会有其他问题。要想总结出一套可行囊括一切人类行为的规则，哪怕是有关红绿灯的规则，看来都是不大可能的。对这些看法我全赞同。

从这一点可以得出，我们不能成为机器。我试图重新进行论证，但又恐怕做不好。似乎可以这么说：“如果每一个人都有一套行动规则来制约他的生活，那么，人同机器就会相差无几了。但实际上不存在这种规则，因此，人不能成为机器。”

有人将“行为规则” 和“行为规律”混为一谈，因此使这个问题有点模糊不清。所谓“行为规则”我指的是像“见到红灯止步”这样的规则。对这类规则你能服从，并能意识到。所谓“行为规律”我指的是自然律，若用在人体上的话，就像“如果你拧他，他就会叫喊”这样的规律。

如果我们将上面所引证中的“制约他的生活的行为规律”改为“他用以制约自己生活的行为规律”，那么，这个论证中的问题就不再是不可克服的了，因为，我们相信，不仅是，用行为规律来制约生活意味着人就是某种机器（尽管这种机器并不一定就是一台离散状态的机器），而且反过来说，如果是这样一台机器的话，那么，它就是由这种规律来制约的。然而，我们很难像否认完整的行为规则那样轻易地否认完整的行为规律。只有通过科学的观察，才能发现这种规律；无论在何种情况下，都不能说：“我们找得够久了，这种规律实际上不存在。”我们可以找到有力的证据，证明这种说法是不正确的。因为，假定存在这种规律，我们肯定能够找到。

对一台离散状态的机器来说，我们极有可能通过观察找到规律，预测其未来的行为，这要有一个合理的极限，比方说，在 1000 年之内。但事情并非如此，我曾在曼彻斯特计算机内输入了一个程序，这个程序用 1000 存储单元，这样，这个 16 位数的计算机就可以在两秒钟内作出回答。我绝对不相信任何人能仅从这些回答中充分了解这个程序，并预测对未试值的回答。

1. 学习机器

读者可能会错误的认为我并没有令人信服的正面论证。否则，我就不花费那么多精力来反驳对方的观点。其实并非如此，下面我将给出这样的证据。

第六部分开始时提出观点的真正令人满意的证明，只能等到本世纪末出现能够通过实验的机器出现。但是我们此时此刻能够拿出什么呢？如果实验将来会成功，现在应该采取哪些步骤呢？正如我所解释的那样，程序的编写是关键。工程设计上的进步也是需要的，但是看上去让它们满足要求是完全能够实现的。估计大脑的存储能力在 10^10 到 10^15 之间。我倾向于比较小的数字，因为我认为只有一个很小的部分被用来进行高级的思考。其余的大部分可能用来保存视觉影像。如果进行模仿者游戏需要的存储容量超过 10^9，是令人惊讶的。(请注意--英国大百科全书的容量为 2\*10^9。)即使立足于目前的技术，10^7 的存储容量也是非常实际的。也许根本就不需要增加机器的执行速度。一部分模拟神经细胞的现代机器的速度比神经细胞快1000 倍。这提供了一个”安全余量”，用来补偿各个方面产生的速度损失。剩下的主要问题就是找到给这些机器编程的方法，以使它们能够完成游戏。现在我一天大概能编 1000 字节的程序，所以如果大约 60 个工人在未来 50 年稳定的工作，并且没有东西扔进回收站，就能完成这个工作。看上去需要更快速的方法。

在模仿一个成人思维的时候，我们必须考虑它是怎样达到当前状态的。可以发现以下三点：

* 1. 思维的初始状态，也就是出生时的状态(b)它所接受的教育

(c)它所经历的，不能被称为教育的事情

除了尝试设计一个成人思维的，为什么不试试设计儿童一样的思维。如果它接受正确的教育，它就可能成长为一个成人的大脑。我们可以假设对机器进行教育的工作量和教育一个人类儿童基本相当。这样，问题被分为两个部分。设计一个儿童程序和对它进行教育。

惩罚和奖励通常是教学的一部分。在这些原则下，就可以建造或编写简单的儿童机器。如果某个行动后很立刻遭到惩罚，机器要能做到不再重复这样的行动；而接受奖励的时候，产生这个奖励的行动以后更有可能发生。我对一个这样的儿童机器进行了一些实验，而且成功的教会了它一些东西。但是教育方法有一些不正规，因此这样的实验还不能称为成功。惩罚和奖励最好能作为机器教育过程的一部分。概略的说，如果老师没有其他的方式和学生交流，教给学生的信息不会超过奖励和惩罚所给信息的总和。一个儿童学要学会说“Casabianca”，一定会经历一个痛苦的过程。如果只能通过“二十个问题”的方法才能发现一个词，每一个“No”都将是一个打击。因此应该寻找其他的“非感情”的交流渠道。如果能够找到，通过惩罚和奖励就能让机器学会以某种语言给出的命令。这些命令通过“非感情”的渠道传输。这种语言的使用将会大大降低需要进行惩罚和奖励的次数。

对于什么样的复杂程度更适合儿童机器可能有不同的看法。可能会有人主张尽可能的简单以保持通用性。另一种方式是嵌入一个完整的逻辑接口系统。在后一种情况下，大多数存储空间将被用来存储定义和命题。这些命题可能具有各种各样的形式。例如，确定的事实， 推测，数学上证明的定理，权威给出的判断，具有逻辑形式却没有确定值的表达式等等。一些命题可以被称为“命令”。机器应该能够在命令是确定的时候立即自动执行合适的动作。例如，如果老师对机器说“现在做你的家庭作业”。这将使“老师说'现在做你的家庭作业'” 成为确定的事实。另一个事实可能是“老师说的一切都是对的。”这两个结合到一起将使“现在做家庭作业”成为确定的事实。而根据机器的建造规则，这意味着立即开始写家庭作业。效果还是令人满意的。

学习机器的一个重要特点是它的老师通常不关心其内部发生了什么变化，尽管老师能够在一定程度上预测他学生的行为。*当被教育的机器是经过试验证实的好设计或编程的时候， 更应该如此。这一点和使用机器进行计算形成鲜明的对比。在后一种情况下，我们需要清楚的明白在任意时刻机器的计算状态。要达到这个目标需要付出艰苦的努力。这样，'机器只能按我们的要求做事'的观点就会显得很奇怪。我们输入机器的大部分程序将引起机器执行一些我们无法完全了解的事，或者一些我们认为完全是随机的事。智能行为应该和完全服从命令的行为方式有区别，而这种区别也不能太大，不应该产生随机的行为或无限循环。通过教学的方式使我们的机器能够进行模仿着游戏的一个重要结果是：“人类的不可靠”的特点很可能被相当自然的忽略，也就是不需要专门的辅导。通过学习得到的行为并不应该是百分之百的，确定的结果；否则，这些过程就不会被忘记。*

在一个学习机器中加入随机元素应该是明智的。一个随机的元素在我们寻找一些问题的答案时是相当有用的。例如我们想找到一个介于 50 和 100 的数，它等于各个数字的和的二次方。我们可以从 51，52 开始一直试下去直到找到满足条件的数。另一个方法是随机的选数直到找到满足条件的数。这个方法的缺点是一个数可能重复试两次，但是当有多解的时候， 这一点变得并不重要了。系统化的方法的一个缺点是可能存在很大一段数中并不存在解，但我们需要先验证它。现在的学习过程可以看成寻找一个行为规则满足老师的要求（或一些其他的标准)。因为可能存在大量的可能解，随机的方法可能比系统的方法更好。请注意，这和进化过程是相似的。但是系统化的方法并不是不可能的。如何能够跟踪不同的基因组合而不产生重复呢？

我们可能希望机器能和人在所有的纯智力领域竞争。但是首先从哪里开始呢？这也是一个困难的决定。许多人可能会说一个抽象的行为，例如下国际象棋可能是最好的选择。也可能需要给机器最好的传感器。然后教它听懂英语。这将和教一个正常的小孩一样。它应该被指出并命名等等。我并不知道正确的答案，但是我想这些方法都应该试试。

我们的目光所及，只是不远的前方，但是可以看到，那里有许多工作要做。

We can only see a short distance ahead，but we can see plenty there that needs to be done。