

本篇文章给大家谈谈人工智能 选择题, 以及人工智能选择题库对应的知识点, 文章可能有点长, 但是希望大家可以阅读完, 增长自己的知识, 最重要的是希望对各位有所帮助, 可以解决了您的问题, 不要忘了收藏本站喔。

本文目录

1. [人工智能的深度学习是什么意思？好学么？](#)
2. [人工智能会给人类带来哪些危害？](#)
3. [人工智能导论知识点总结](#)
4. [假如人工智能有自我意识后，杀了人，怎么判断是故意还是过失？](#)

人工智能的深度学习是什么意思？好学么？

我们来一起梳理一下人工智能与深度学习的关系。

人工智能

首先, 大家所谈论的人工智能可以分为两个层面: “强人工智能” 和 “弱人工智能”。其中:

弱人工智能

希望借鉴人类的智能行为, 研制出更好的工具以减轻人类智力劳动, 类似于 “高级仿生学”。

强人工智能

希望研制出达到甚至超越人类智慧水平的人造物, 具有心智和意识、能根据自己的意图开展行动, 可谓 “人造智能”。

AI技术现在所取得的进展和成功, 是缘于 “弱人工智能” 而不是 “强人工智能” 的研究。要想让AI借鉴人类的智能行为, 关键的一个环节是让AI模拟人类的学习行为。

所以, 这里面有个非常关键的技术, 叫做机器学习。

机器学习

机器学习是专门研究计算机怎样模拟或实现人类的学习行为, 以获取新的知识或技

能，重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域，它主要使用归纳、综合而不是演绎。

目前的机器学习可以分为三大类：

(1) 有监督的学习

数据具备特征 (features) 和预测目标 (labels) ，又分为：

a.二元分类

简单粗暴地理解，即让AI做是非题

b.多元分类

可以理解为，让AI做选择题

c.回归分析

可以理解为，让AI做计算题

(2) 无监督的学习

从现有数据并不知道预测的答案，无预测目标 (labels) 。

(3) 强化学习

通过定义的动作、状态和奖励不断训练，使其学会某种能力。

机器学习有一个很有意思的技术，叫做人工神经网络。

人工神经网络 (ANN : ArtificialNeuralNetwork) 是一种模拟人脑神经网络以期能够实现类人工智能的机器学习技术。它可实现函数逼近、数据聚类、模式分类、优化计算等功能。因此，神经网络广泛应用于人工智能、自动控制、机器人、统计学等领域的信息处理中。

通过这两张图的对比，我们可以看到，机器学习中的神经网络很好地借鉴了人类神经网络的特点，是一种非常有意思的仿真。

深度学习

而深度学习是一种特殊的机器学习，是机器学习研究中的一个新领域，其动机在于建立、模拟人脑进行分析学习的神经网络，它模仿人脑的机制来解释数据，例如图像，声音和文本。

深度学习能直接对大量数据进行表征学习，来替代手工获取特征。深度学习与传统机器学习最主要的区别在于：随着数据规模的增加其性能也不断增长。引发深度学习热潮的一个标志性事件是：2016年3月，AlphaGo（谷歌旗下DeepMind研发）击败了李世石九段。

相应的，深度学习有一个非常重要的技术，叫做卷积神经网络。

卷积神经网络（Convolutional Neural Network, CNN）是一种典型的深度神经网络，它避免了对图像的复杂前期预处理，可以直接输入原始图像，因而得到了更为广泛的应用。

总结

最后，我们用一张图来梳理一下人工智能、机器学习、深度学习的关系：

人工智能会给人类带来哪些危害？

先说结论：我个人的浅见是，AI化生存是人类演化的必然，不以人类的意志为转移。马斯克：人类做的所有事情中最可怕的，就是试图唤醒计算机，研发强人工智能是迄今为止人类可能面临的最大的生存危机，应对这个危机的唯一办法就是：成为AI。

人工智能会给人类带来灾难吗？其实这句话实质的含义是人工智能的觉醒会不会威胁到人类的生存？

下面我将带你逐一剖析人工智能觉醒话题的方方面面。

觉醒问题的科学实质

人工智能，简称AI，这个名词现在几乎已经到了随处可见的地步。

2016年，一只会下围棋的“阿法狗”把人类对人工智能的警惕推向了又一个高潮。如果你也是一个科幻爱好者，可能会跟我一样追美剧《西部世界》。这部片子向

我们细致描绘了人工智能是如何被制造出来，并且最终觉醒试图取代人类（截止到本文写作时，《西部世界》第二季正在上映）。而美国畅销书之王丹·布朗今年推出的最新力作《本源》，也是以人工智能为题材的小说。

2017年国内非虚构类图书畅销冠军就是尤瓦尔·赫拉利的《未来简史》。在这本书中，作者向我们历数了现在人类的大部分职业都将被人工智能取代，包括一向被认为是高知识含量的医生、律师、记者等等。

大多数人能够做的工作，未来人工智能都能做，这一点几乎没有什么争议。然而，人工智能是否能够觉醒，也就是具备自主意识，像人一样自主思考，这就有非常大的争议了。

本文将带你深入了解科学界对此的研究历史和各方观点。

实际上，从科学的角度来说，觉醒的路径有两条：

一条是由人类编制出来的电脑程序觉醒。另一条路径是把人类的思维和记忆全部上传到计算机中，让人类在电脑中觉醒。

到底哪一条路径更有可能发生呢？让我们带着这个问题开始今天的探索之旅。

首先，我们先来讨论一下到底什么是意识？

这是一个非常难以定义的词，有点儿像“精神”或者“我”这样的哲学味道很浓的词。对意识的探究自古就有。

古希腊时代，亚里士多德撰写了《灵魂论及其他》，阐述了他对生与死、睡与醒、以及其他心脑问题的深入思考。而另一位古希腊学者希波克拉底认为，人类应该知道，我们的喜怒哀乐都来自大脑，且只来自于大脑。这可能是最早的心物一元论。

笛卡尔在17世纪提出了心物二元论，他的名言是“我思故我在”，认为人由完全不同的两种实体组成，一个是心灵，另一个是身体。心灵是非物质实体，永生不灭，不占空间，具有意识。而身体是物质实体，可以被摧毁，占据空间。心灵寄居在人的身体，或者说寄居在人的大脑内，大脑也属于身体的一部分，心灵接受身体传来的信息，并向身体发送指令。

科学的进步使我们对意识的研究不再仅限于哲学领域。

到了19世纪末，西班牙神经解剖学家罗曼尼·卡哈尔运用并改进了银染色法，对神

经系统进行了大量研究，创立了神经系统的神经元理论。他认为神经系统并非是连续一片的，而是由一个个的神经元组成。卡哈尔的研究，为现代神经科学奠定了基础。

现代科学认为，意识是一种“神经反应”，是一种自我感受、自我存在感与对外界感受的综合体现。换句话说，现代科学已经否定了笛卡尔的心物二元论。科学家认为意识必须依赖于神经元这种实体，它是人自主或者不自主地产生的一系列神经反应的体现。通俗地说，意识就是一种大脑活动。这是一个重要的前提，正是因为意识是基于物质的，我们才有了继续研究觉醒的可能性。

好了，有了以上这些知识后，我们就可以这样认为，所谓的觉醒问题，它等价于是否可以用我们现在的计算机系统来完全模拟神经元的活动。

如果能模拟，那么计算机系统就必然有觉醒的理论可能，反过来，如果神经元的活动是现在的计算机结构根本不可能模拟的，那么，觉醒就是一个伪命题，至少可以说，人类的意识与电脑的意识是不同的。

于是，觉醒问题就暂时转换成了计算机系统能否模拟神经元活动的问题。

你看，我刚才的这一番论述就体现了科学思维中非常务实的一面，如果我们只有哲学思维，那么，光光是研究到底什么是“我”，什么又是“精神”，就可能永远争论不休了。科学思维则把这个问题转换成了可以通过逻辑和实证来研究的具体问题。

我们现在拿到了两个研究对象，一个是计算机系统，另一个则是大脑的神经元系统。现在我们要研究的是这两个系统之间到底存在哪些异同点，是否是等价的，或者说，计算机系统未来是否有可能可以和大脑神经元系统等价。

图灵机

你要知道，这个问题相当的复杂，不过已经有许多科学家给出了非常精彩和深入的思考，为了让你能充分领略这些科学家们的非凡见解，我要先给你讲解计算机系统的本质是什么？

现代所有的计算机，从办公桌上的电脑，到手机、Pad以及各种智能电子设备，其实从本质上来说，都有一个共通点，它们在本质上都是一台图灵机。

怎么理解这个概念呢，我打一个比方。我小时候喜欢帮妈妈拆旧毛衣，每次我拆毛衣的时候，总是会有一种很神奇的感觉，因为不论有着多么复杂结构和图案的毛衣

，拆的过程中就发现，其实都是一根或者几根长长的毛线缠绕出来的。所以，在普通人的眼中，计算机无比复杂，可是在有些科学家的眼中，计算机就像是拆成毛线的毛衣，无非就是一台图灵机。

但这个比方只能帮你建立一个总体的概念，为了深入理解我后面要讲的东西，我还需要继续给你解释图灵机到底是怎么回事？

图灵机是英国传奇科学家图灵提出来的一种抽象的计算机，我估计你对这个名称大概是不陌生的，它与我们这次的主题关系重大，所以，我必须让你对图灵机的工作原理了解得比别人更多一点。下面的内容可能会比较难，但只要有点耐心，一定难不倒你。

图灵机只有两个部件，一根无限长的纸带，纸带上画好了一个个的格子，每一个格子有三种可能，写着0或者1，或者什么也不写，我称作“空”。然后，有一个读写头，这个读写头可以在纸带上前后滑动，可以读取纸带上的内容，也可以在纸带上写字或者把已经写好的字擦除。你可以把读写头想象成自己拿着一个橡皮擦和一支铅笔。

（图灵机的原理）

这个读写头可以给自己标记一个当前状态值，例如，我用“起”表示读写头的当前状态是起步，用“终”表示读写头当前的状态是终了，用“停”表示读写头的当前状态是停机。读写头还有一个本事，就是可以根据一组条件规则来执行一组动作。

好了，关于图灵机就介绍完毕了，就是这么一个假想出来的简单机器，但是它却无比强大，它就像是那根可以编织任何华丽毛衣的毛线。图灵机可以完成任何有限次数的数学和逻辑运算。

我来举一个具体的实例帮助你理解图灵机的工作原理。我们只要给图灵机赋予六条基本的规则，就能让它完成给任意数字加1的工作。这六条基本的规则是这样的：

第一条：如果读写头的状态是“起”，并且读到的数据是空，就写入数字1，并把自己的状态设为“终”，读写头右移一格。（从下一条开始，我就把上述规则中那些描述性的“废话”都删除了）

第二条：起01终右

第三条：起10起左

第四条：终空空停左

第五条：终00终右

第六条：终11终右

以上就是让图灵机完成加法运算的六条规则（程序）。当然，根据初始的数字不同，并不是一每一条规则都会被执行到。但不论遇到什么情况，读写头一定会从“起”的状态执行到“停”的状态，当读写头停下来了，就表示纸带上显示的数字是初始数字加1后的数字。

下面这张图是 $111+1=1000$ 的计算详细步骤。（有些读者可能会不懂，为啥 $111+1$ 会等于1000呢？因为这是二进制计算，不是逢十进一，而是逢二进一）

111+1=1000的计算详细步骤（二进制）

既然可以加1，那么加2无非就是重复执行2次，能执行加法就等于能执行所有的加减乘除，甚至是更高级的幂运算或者开方运算了，读过高中的读者都应该明白这其中的道理，一切运算都可以还原为加法运算。

好了，如果听到这里，你完全听懂了，那么恭喜你，你的见识已经成功地提升了一个境界，你就好像是升级成功的尼奥，可以看透黑客帝国母体的本质了。一切现代的计算机在本质上都是这样一台图灵机。

不过，你马上就会生出更大的疑问：一个会做四则运算的计算机怎么就能变成今天可以看电影、打游戏、发微信的手机的呢？图灵机和人的意识到底有什么关系？

计算机算法的局限

其实，计算机上的所有行为都可以转换为数学和逻辑运算。如果把计算机的行为抽象出来看，无非就是输入、计算、输出这样三个过程。例如你在屏幕上点一下“磨皮”的按钮，这个叫输入，经过一定的运算后，返回你一张更漂亮的自拍照，这个叫输出。

解释“磨皮”背后的算法会太复杂，我们简化一下问题：将图片从彩色变成黑白。它背后的算法就是当年研究彩电的人鼓捣出来的，计算过程是这样：大家都知道任何彩色图片都是由红绿蓝三原色组成的，也就是照片上的每一个像素都是一个彩色的点，而这个点都是红绿蓝三种颜色的混和，现在我们只需要利用一个经验公式，把红色的颜色值乘以0.2989，绿色值乘以0.587，蓝色值乘以0.114，再把结果求和

，得到的数值就是图片的灰度数值，把照片上的每一个像素点计算一遍，转换成黑白模式，就得到了一张从彩色变成黑白的照片了，它的背后就是数学。

上一节介绍的那六条完成加1运算的图灵机规则也可以叫做程序，计算机程序就是用来描述某种运算过程的机器语言，所有程序还原到最后都是图灵机中两个条件加上三个动作的组合。

好了，有了这些前置知识，我们就可以继续来探讨人工智能是否能觉醒的问题了。

我们现在有两个研究对象，一个是计算机系统，一个是大脑神经元系统，解决问题的关键在于比较这两者之间的异同点。就我所知，在这个问题上，最早做出深入思考的科学家是英国著名的物理学家彭罗斯，他专门就此问题写了一本厚厚的专著，叫《皇帝新脑》。彭罗斯认为：只要我们的计算机系统依然是图灵机，就不可能觉醒。话说得斩钉截铁，没有任何含糊。

那彭罗斯何以能得出这么斩钉截铁的结论呢？他当然是有论据的。彭罗斯的论据很多也很复杂，我这里选取其中最重要也是最核心的二个论据解释给你听。

第一个论据叫做哥德尔定理。这是大数学家、也是爱因斯坦的好朋友哥德尔证明的。哥德尔定理有两条：

任何相容的形式系统，只要蕴涵皮亚诺算术公理，就可以在其中构造在体系中不能被证明的真命题，因此通过推演不能得到所有真命题。任何相容的形式系统，只要蕴涵皮亚诺算术公理，它就不能用于证明它本身的相容性。

我估计绝大多数人看不懂，因为术语太多。但是这没关系，有时候知道自己并不是真正的懂反而是好事，带着问题去学习效果最好，要真正搞懂哥德尔定理，那需要很多前置知识，再下一番功夫。

我这篇文章是为了科普，所以不得不放弃一些严谨性，上面这两条哥德尔定理可以得出一个推论：在数学中，必定存在既不能证实也不能证伪的命题，假如我们只能用数学本身来证明的话。我务必提醒你注意后半句“假如我们只能用数学本身来证明的话”。这也是很多人对哥德尔定理的一个误解。哥德尔定理只是说数学本身无法证明所有的数学问题，而并不是说所有的数学问题根本无法证明。假如我们能找到一种比数学更高一个层次的形式逻辑，那么就有可能证明所有的数学问题了。

再换句话说，就是存在无数个数学结论，但这些结论根本不可能用数学本身给算出来。图灵机只能用数学来计算数学问题，而且仅仅只是那些数学中可以被有限次计算步骤计算出来的问题，这些问题只占所有数学问题的一小部分。

但是，我们的人脑好像不是这样的。

人似乎有一种很厉害的本领叫做“洞察”或者说“直觉”，我们可以凭空冒出许多最终被证明是正确的想法，这些问题完全不像是计算出来的。例如，在数学中有一个经典的不可知问题，就是集合论中的连续统假设，你不需要去搞明白这个假设到底说的是什么，你只要知道，这个数学问题就是用数学本身无法证实也无法证伪的问题。假如我们的大脑也是像图灵机一样工作，那么按理说，我们只要一想这样的问题，就应该死机了，大脑陷入了无限死循环。但事实上没有，而且人类还可以凭空发现这个问题是不可计算的。

在彭罗斯看来，这就说明了，我们人类的意识是高于数学系统的，只要我们的计算机还是图灵机，还是完全基于数学的框架，那就不可能达到人类意识所表现出来的种种令人惊叹的行为。

第二个论据叫做算法无法演化。彭罗斯的原话是这样说的：假如我们的大脑只是在执行一系列非常复杂的算法，那么我就应该询问这种非常复杂有效的算法从何而来。生物学家告诉我们的答案是“自然选择”。也就是说，生物在演化过程中，那些更有效、更有利于生存的算法会保留下来，遗传给下一代，如此逐步升级。我完全相信自然选择是生物演化的原理，但是我却看不出自然选择本身如何能演化算法。对比我们所有的电脑程序，你会发现，程序的有效性和概念本身最终要归功于至少一个人类的意识。

彭罗斯的意思是说，人的大脑可以凭空创造出算法，但是图灵机不能自己创造出算法，只有被人类的意识赋予了那些按照条件执行动作的规则，图灵机才能运行。

当然，彭罗斯还提出了其他一些论据，包括著名的图灵停机问题，我这里不再展开谈。总之，彭罗斯的结论就是，人类目前的计算机工作原理是无法创造意识的，电脑和人脑不在一个层次上。

在这里，我必须提醒大家，彭罗斯《皇帝新脑》这本书的出版年份是1989年，那个时候的计算机和今天的相比显得非常原始，那时候人们对大脑的认识也与今天不可同日而语。但是，即便是在那个年代，即便以彭罗斯当年在科学界的声望，他的这本书一出，也几乎是立即就遭到了强烈的反对声。

例如，著名的认知科学和人工智能专家马文·明斯基（Marvin Lee Minsky）就是反对声最大的几个科学家之一。明斯基是麻省理工学院人工智能实验室的创始人之一，1969年就获得过图灵奖。有些文章中把他称为人工智能之父，当然，有这个称号的科学家其实很多。明斯基2016年去世。

明斯基在看完彭罗斯的书后，在一次学术会议上，非常激动地做了一次演讲，题目是“有意识的机器”，其中一段原话是：在某些思想领域，更多的人对我们的无知本性采取了不同的立场。他们努力工作，为的不是找到答案，而是表明根本没有答案。这就是彭罗斯在《皇帝新脑》中所做的，他在书中一章接一章地反复叨念着，人类的思维并不基于任何已知的科学原理。

另外一位著名的人工智能专家，也就是在1971年就获得过图灵奖的约翰·麦卡锡（John McCarthy）在1998年的一篇论文中指出：彭罗斯忽略了一点，他宣称“形成判断是有自我意识的标志，编程人员是没法把这一点变成计算机程序的”，而事实上，大多数的AI文献都讨论过在机器记忆中有关它做出的事实和判断的表现。用AI的术语表达就是，AI的认知论部分和启发式部分一样突出。

总之，一直到现在，对彭罗斯观点的反对声也是不绝于耳。总的说来，所有的反对声可以总结为这样一个观点：图灵机的局限性彭罗斯已经说得很清楚了，但是他却没法严格证明人脑已经突破了图灵机的局限性，我们今天所知的一切看上去令人惊叹的大脑行为，完全有可能依然是在可计算的数学框架内的，并没有逾越这个大框框。

是不是真的是这样呢？下一节，我将带你去了解一下目前人类已经掌握的人工智能算法的实质到底是怎样的。

人工智能与机器学习

可能在你的认知中，人工智能的技术原理是根本无法理解的高超技术，只有那些疯狂的科学家才懂。但是，我却想告诉你，今天所谓的人工智能技术可能离你心目中的“智能”、“智慧”相去甚远。准确地说，人类今天所掌握的人工智能技术只是一种“机器学习”和“概率预测”的技术，不但离不开“人工”，也并不“智能”

。

下面我将用一个实例来为你揭开你在各个大佬口中听到的那些“人工智能”的实质

。

现在，我给你一份泰坦尼克号所有船员的名单。这份名单包含了以下这些信息：

乘客编号、姓名、舱位、性别、年龄、是否与亲戚同行、船票号码、票价、房间编号、登船码头，以及是否存活。

我将这份名单随机分成两半，这两半数据只有唯一的一个差别：我隐藏了其中一组数据中“是否存活”这项。现在，我交给计算机一个任务，就是通过学习其中一半

的数据，然后来预测另一半数据中每个人的生死。

这个程序可以写得极其简单，也可以写得很复杂。

最简单的算法是：统计一下乘客的死亡率，发现已知的一半人的死亡率是62%；有了这个数据，那么另一半人我把他们全部预测为“死亡”，也有62%的正确率了。那怎么继续提高预测准确率呢？我们继续统计已知数据中男女的死亡比例会发现，女性的存活率是74%，男性的存活率是18%。好了，有了这个数据，我们马上可以大幅度提高预测的准确性了。

算法可以继续复杂和优化，乘客的每一个属性都有可能影响他的存活率，而很多属性又会交叉影响。但无论怎么复杂，都是一种数学统计模型：通过已知的这一半数据，不断地优化每一项参数在存活率中所占的权重，最终得到的是一个数学公式——把乘客的每一个属性的数值（例如票价、年龄等）代入公式。你可能会问，像性别这种只分男女，没有数值，怎么办呢？这不难办，在数学建模中，我们可以给男女人为规定一个数值，比如男=1、女=2，或者男=0、女=1，这就看你的算法怎么设计。最终，每一个乘客根据这个公式，都会计算出一个表示生或者死的数值。这就完成了从机器学习到预测的全过程。

上面这个例子并不是我杜撰的，而是布鲁萨德的新书《人工不智能》中一个AI算法的实例。我只是将这个例子提炼精髓，做了进一步的简化。在这个实例中，计算机算法对于泰坦尼克号上的乘客死亡率的预测准确性可以达到97%。

从这个例子中，我们可以得出这样两个结论：

人工智能依赖已知的数据工作：你喂给它的数据越多，它就可以预测得越准确；反之，如果吃不到数据，它就无法工作。人工智能本质上只是一种数学统计模型的具体应用，本质上还是一个计算器：只是计算公式超复杂，运算速度超快而已，计算机并没有“思考”。

所以，人类今天所掌握的人工智能技术还只是一种“机器学习”和“概率预测”的技术，不但离不开“人工”，也并不“智能”。现在我们人类所开发出的所有人工智能程序，不论是“阿法狗”，还是谷歌的打电话AI，它们在实质上依然符合上面两条。理解了这些，希望你对人工智能不再感到神秘和膜拜。

显然，像这样的人工智能还远远谈不上觉醒。

因此，搞人工智能的圈子把这种类型的人工智能称为弱人工智能，而像《西部世界》中那样有自我意识的人工智能则被称作强人工智能。

但是也有一些哲学家和科学家认为，弱人工智能和强人工智能之间并没有一条泾渭分明的线，换句话说，无意识和有意识之间也不是生和死这样的明确差别。我们每个人自以为的自由意志或许只不过是一种幻觉，而这种幻觉只不过是条件和规则足够复杂后涌现出来的罢了。

他们认为，弱人工智能本质上是“机器学习”和“概率预测”的技术，但你怎么知道我们人类所谓的思考其本质上就不是这样工作的呢？人类做出的任何判断，也必须依托于过去的经验。差别仅仅在于条件和规则的复杂程度，当弱人工智能处理的数据足够多，执行动作的可能性也足够多时，自我意识就涌现出来了，也就逐步成为了强人工智能。

关于人工智能与自我意识之间关系的哲学思考，有三个著名的思想实验。知道了这三个思想实验，你差不多也就大致了解了哲学家是怎么看待人工智能的觉醒问题。

三个思想实验

图灵测试

在1950年之前，对于到底什么才算是“智能”，不论是哲学家还是科学家，都很难给出一个大家都比较认可的定义，直到图灵在1950年发表了那篇意义非凡的论文。图灵在这篇论文中大胆地预言具有真正智能的机器一定会出现。为此，他给智能下了一个定义，这就是著名的图灵测试：假如一个人与另外两个“人”（A和B）进行不见面的对话，A是机器，B是一个真人。但是，无论这个提问者怎么努力发问，他都无法区分出A和B到底哪个才是机器，那么就可以说这台机器通过了图灵测试。

图灵提出，假如一台机器通过了图灵测试，我们就可以认为它具备了智能，也就是说，它觉醒了。图灵测试之所以出名，因为这是在人工智能探讨方面第一个非常严肃而且具备可检验性的定义。从中也可以看出当一个受过严格科学训练的科学家进行哲学思考时，他能够把科学思维带入到哲学思辨中。

图灵测试提出后，表示反对的声音当然一直就有，只是，没有人能提出一个比图灵测试更具备可操作性的替代定义。

在所有对图灵测试的反驳中，最出名的是1980年美国哲学家约翰·罗杰斯·塞尔在一篇论文《心灵、大脑和程序》中提出的一个思想实验。

中文房间

假设有一个对汉语一窍不通，只说英语的人关在一间只有一个开口的封闭房间中。

房间里有一本用英文写成的手册，指示该如何处理收到的中文问题，及如何以中文回复。房外的人不断向房间内递进用中文写成的问题。房内的人便按照手册的说明，查找到合适的回答，将相应的中文字符组合成对问题的解答，并将答案递出房间。

塞尔认为，哪怕房里的人能骗过房外的人，让对方以为自己懂中文，其实房里的人压根就不懂中文。如果把这个思想实验类比图灵测试，那么房里的人可以比作是试图通过图灵测试的计算机，而那本手册可以类比为计算机程序。虽然房里的人可以正确地回应房外人的问题，但是，正如房里的人压根不理解中文一样，计算机也不可能通过程序来获得理解力。既然计算机没有理解能力，那“智能”便更无从谈起了。

但是，关于这个思想实验的争议也很大。有些哲学家从这个思想实验中得出的结论反而是所谓人类的理解力也不过就是一种“幻觉”，本质上和查手册并无区别。有些哲学家反驳塞尔，假如你说的这台机器能通过查手册以假乱真，那么，我完全可以认为它具备了对中文的理解力。所谓的理解，就是人、纸、笔、手册等所有这些工具联动起来的一种涌现而已。

另一位美国哲学家丹尼尔·丹尼特就提出了一个与中文房间针锋相对的思想实验，试图论证机器人和人不存在本质区别。

亿年机器人

丹尼特是美国塔夫茨大学的哲学教授，他在2013年出版了一本书《直觉泵和其他思维工具》，就是在这本书中，亿年机器人的思想实验被首次提出。这个实验是这样描述的：

假如我们爬进一个休眠仓，希望自己能在1亿年之后成功地苏醒。为了达成如此艰巨的目标，我们必须制造出一个能感知环境、回避风险、寻找资源的机器人，我们只留给他一个指令“让我活着”，然后我们就休眠了。在这1亿年中，我们再也无法对这个机器人进行任何干预。机器人为了完成终极目标，必然会把这个大目标分解成无数个小目标，在执行这些目标的过程中，机器人就会开始演化，就会表现得越来越像人。

丹尼特讲到这里，话锋一转，他提醒每一个读者：在这个思想实验中，我们每个人不就是机器人嘛，而那个思想实验中的人就是基因。基因创造了生物，然后就放手了，什么都不管，唯一的指令就是把我传递下去。人就是基因为了让自己永生而演化出来的产物，当基因把控制权让渡给了人类的那一刻起，人类就拥有了自由意志。

那么，既然人本来就是机器，机器为什么就不可以拥有意识呢？

好了，我把对人工智能是否会觉醒的三个最重要的哲学思考给你介绍完了，不知道你会更倾向于哪一个思想实验？稍微换了一下脑子后，我还是要继续带你回到科学的轨道上来，科研是非常务实的，但每一步都是积累确定的知识。

在科学家看来，觉醒问题可以转换成计算机系统能否模拟神经元活动的问题。他们有两个研究对象，一个是计算机系统，另一个则是大脑的神经元系统。从下一节开始，我要给你介绍人类对大脑的研究历史。

脑电图

英国生物学家理查德·卡顿1875年发现，把电极放置在与视觉相联系的脑区表面，然后对实验中的动物发出闪光的刺激，能够检测到电压的变化。他也观察到了自发的脑电变化，但他对此没有特别的兴趣，也就没有进一步探究下去。

几乎与此同时，波兰科学家阿道夫·贝克进行了相同的实验，并得到了相同的结论。贝克对脑表面自发的脑电活动做了这样的总结：“在第一次实验和所有的验证实验中我都注意到，如果将两个电极放在大脑半球的两个固定位置时，两个电极之间的电位差并不稳定，会有增减变化。这种变化既与呼吸节律无关，也与脉搏不同步，而且这种变化的方式也不取决于动物的运动。我们在被毒素麻痹了的狗身上也发现了这种变化。因此我相信，这些电位变化是脑中枢自发活动的结果。”

在卡顿和贝克从动物实验得到各自发现的五十年后，人们终于确切地证实了：人脑也有自发的脑电活动。

当时在精神病院工作的德国精神病学家汉斯·伯杰发现，在人的颅骨表面能够跟踪记录到脑电活动。伯杰是在研究了心灵心理现象之后，转而研究脑电活动的，他最开始的研究内容包括思想感应和心灵致动。他根据自己的发现下了断言，在思想感应活动期间，仪器能够跟踪记录到人的脑电活动。一年后，也就是1924年，伯杰第一次成功记录到了病人的脑表面自发的脑电活动，并将其命名为“脑电图”（也就是我们熟知的EEG）。他的结论是：当受试者的感觉器官受到刺激，或者精神活动活跃时，受试者的脑电压会降低，并且波动加剧。伯杰对自己的实验守口如瓶，并禁止其他人进入他的实验室，也从不与同事讨论他的发现。

在第一次发现并记录到人类脑电图的五年后，也就是1929年，他才第一次公开发表了关于脑电活动的研究成果。从1929到1934年，他每年发表一篇补充文章。1938年，他公开蔑视纳粹党的种种针对犹太人的暴行，纳粹党人解除了他精神病院主管的职务，他被迫退休。三年后，他因抑郁症被送入了精神病院，同时患有严重的

皮肤病，他不堪打击，选择了上吊自杀。伯杰的发现完全改变了脑科学研究，他本人也几次被提名诺贝尔奖候选人，但德国纳粹从没有承认过他的重大贡献。

目前，脑电图仍然是探测大脑意识最重要的工具，并且，这种方法在未来很长一段时间内可能仍然会是主流。除了脑电图，还有其他记录大脑活动的方法，最常见的有两种，一种是用磁共振扫描仪检测大脑内部的血液动力学反应，另一种是用脑磁图（MEG）检查大脑周围的磁场模式。最近还新诞生了一种名叫“近红外光谱仪”的技术。但这三种方法，在理论和实际操作上都存在各种问题，因此还不能作为常规方法被应用在临床上。

脑电图技术能够检测大脑新皮层上由脑电活动产生的微小的电压变化，变化值在10-100微伏之间。新皮层位于大脑外表面，主要负责知觉、动作、记忆和思维等认知活动。脑电信号是一种通过“体积传导”机制产生的集群电信号，主要由皮层椎体神经元产生，皮层椎体神经元因形状呈四面体而得名。但EEG无法对大脑皮层下的结构进行直接的测量，医生只能通过EEG对皮层细胞的作用进行推断。

EEG的一大优势就在于它不需要受试者进行侵入性手术，比如把核心设备穿透头骨连接到大脑内。EEG只需要将电极直接放置在头皮上，就可以采集信号。随着EEG技术的发展，它能采集的脑电信号的密度也越来越高。现在你甚至最多可以在头皮上放置256根电极。所以，用EEG记录整个大脑的脑电活动的分布情况已经很容易就能实现了。

脑电图虽然不需要侵入性手术，但放置电极的过程极其繁琐。先要擦洗头部的皮肤，然后将湿润的电极膏涂在几十甚至几百个电极上。在采集的时候受试者还要尽量保持不动，如果电极移动，收集到的数据就很容易产生误差。所有这些问题都限制了脑电图技术的发展。

不过现在，干电极横空出世了，它更敏感，并且使脑电图仪从临床专用设备变成了一款消费品类设备。用户可以连续数小时使用干电极设备，获得关于自己的生物反馈信息。极客们可以用它分析自己的思维活动，失眠症患者可以用它记录睡眠状态。

说了那么多脑电，那脑电和意识之间的关系又是什么呢，是不是有脑电就有意识？

其实从20世纪40年代后期开始，科学家就下了一个判断，个体有自主意识的最重要的标志是：能够检测到“活跃的”脑电信号。那什么又是活跃的脑电信号呢？简单说来，它的特征是，电压很低且上下快速波动。整个颅骨上的脑电波都是不同步的。一般说来，脑电波越向低频偏移，意识就越不可能出现。但是，也有很多相反的例子能推翻这一规律。所以，现在的结论是，“活跃的脑电信号”不能作为特定

个体是否存在意识的基本标准。

于是，科学家和临床医生需要寻找更可靠的测量方法。

意识的分界线

为什么需要一个判断人是否有意识的客观标准呢？可能你觉得这仅仅是一种研究大脑的科学探索活动。其实，这个客观标准有着非常重要的现实意义。

有两类患者，他们的意识状态非常需要医生进行判断。

第一类患者包括创伤性脑损伤、脑炎、脑膜炎、中风或药物、酒精中毒后具有严重意识障碍的人。

患者从这些来势汹汹的疾病中幸存下来后，身体状况逐渐稳定，但会出现残疾。他们有可能卧床不起、无法说话，或者虽然能张口但说不清自己的想法和意图。这些病人经过适当的护理，避免褥疮和感染，还可以存活很多年。

这些处于植物人状态的患者，用术语来说就是无反应性觉醒综合征的患者，总是循环性地进入和脱离睡眠状态。医护人员在床边对患者说，如果你听到我的话，捏下我的手或转动下你的眼睛，但患者并没有反应。植物人可以把食物吞下去，也可以打哈欠、睁眼、转动眼球或头部，但他们并不是有意为之。

患者剩下的只有似乎是无意识的行为，也就是通过脑干反射来控制生命的基本过程，这些条件反射包括呼吸、从睡眠到苏醒的过渡、心跳、眼动和瞳孔反射。特丽·夏沃是一个很有名的患者。她来自佛罗里达州，心脏骤停后，变成了植物人，这种状态持续了15年，直到2005年，她被医学手段结束了生命，也就是俗称的安乐死。要求法院结束她生命的正是她的丈夫，而她的父母坚决要求继续维持女儿的生命迹象，官司一直打到了美国联邦最高法院，最后她的丈夫胜诉。植物人是现代社会特有的现象。由于有了紧急救护系统和先进的医疗护理，植物人得以在大病后存活下来。在美国，有超过一万名植物人，他们住在临终关怀中心、疗养院或自己的家中。

尽管从行为上看，植物人似乎对外界没有意识，但缺乏证据并不意味着这些病人肯定没有意识，这种科学的怀疑态度，对病人是有益的。毕竟也有植物人醒过来的奇迹般的事例存在。对植物人患者，医学上存在一个诊断的“灰色地带”，也就是在他们是有一部分意识还是完全丧失意识上，模棱两可。有研究表明，20%的植物人是有意识的，他们其实是被误诊了。对那些照顾患者多年的亲友来说，知道自己关心的人是有意识还是无意识，心理上肯定会有完全不同的感受。

在临床诊断中，MCS代表最小意识状态，这类患者不能说话，但却可以发出信号。他们发出的信号往往只是零碎、微小、没有规律的，比如在适当的情绪状态下微笑或哭泣，偶尔发出声音或打手势，或者用眼睛跟踪明显的物体等。我们猜测，这样的患者存在着微弱的意识，可以意识到一些事情，起码在某些特定的时刻是这样。

第二类患者完全不同，他们属于像我们一样大脑运转正常的普通人，只是因为要做手术，进行了麻醉。很多常见的疾病，比如去除恶性肿瘤、固定膝关节等等，都需要使用麻醉。麻醉可以消除疼痛，但也抑制了其他的意识，麻醉会阻止患者移动身体，也会稳定呼吸，控制自主神经系统。这样的麻醉，一次的效果可以持续数小时。

一般来说，患者不会在手术过程中醒来，所以他们也不会被血淋淋的手术过程吓到，留下创伤性记忆，困扰终生。倒霉的是，麻醉觉醒可能会在少数的手术中发生，比例是每1000例手术会发生一次。如果麻醉师在手术过程中给麻醉患者插管，这种情况就很有可能发生。现有的脑电图技术能够监测手术中的麻醉深度。但在种类繁多的麻醉药物中，没有一种能对各个年龄段的患者都有效。

这就是临床医生认为的最需要判定是否存在意识的两大类患者。

为了清楚地了解他们的意识水平，我们需要一种可靠的探测个体意识的工具。所以，就临床实践来说，关键问题不是问到底什么是意识，与其争论定义，不如发明一个能够客观检测大脑活动强度的工具，有了大脑活动的指标，至少可以脱离只定性不定量的哲学思考。

本世纪初，为了探测意识的整合度，威斯康星大学麦迪逊分校的精神病学家、神经科学家朱利奥·托诺尼和另一位神经科学家马西米尼的团队合作，他们基于脑电图和一种名为“整合信息论”的意识理论，通过一种计算机算法，对意识水平进行了比较粗略的估计。虽然算法算不上精确，但他们用这种方法，准确区分出了六名健康志愿者的状态，比如他们是清醒地静静地闭着眼睛，还是处于深度睡眠的无意识状态。这初步证明了该方法的有效性。

处于深度睡眠的人，大脑就像一个运转功能出问题的时钟。在刚刚所说的实验中，志愿者脑电波的初始波幅大于他醒来时的波幅，但脑电波的持续时间要短得多，并且不会让相互连接的脑区彼此回应。深度睡眠时，虽然神经元仍然保持活跃，这一点可以从局部脑区的强烈反应中得到验证，但大脑的整合功能已经不再起作用了。清醒大脑中的脑电活动模式，在深度睡眠状态下的大脑中很难见到。扩展该方法，可以区分一些更复杂的大脑状态。在过去几年里，托诺尼、马西米尼和另外十七位神经科学家，已经在许多受试者中测试了这个方法，并于2016年发表了一篇具有里

程碑意义的研究论文，详细总结了这套方法。

研究人员使用经颅磁刺激（TMS）技术，在患者的头皮上放置了一个封闭线圈，线圈会发送一个脉冲磁场，刺激大脑的相应部位，在皮层神经元中触发短暂的电流。从不同脑区对脉冲磁场的反应可以看出，如果一个人是清醒的，他的大脑是神经连接完好的，那么他大部分的大脑皮层的活动都非常复杂，既不是完全规律的，也不是完全随机的，只能用“复杂”这个词来形容。

研究人员用一种数学方法评估了大脑的复杂性，想弄清楚不同的脑区在不同的时间范围上，神经反应究竟存在着多大的差异。这种评估方法来自计算机科学，它的算法也是著名的zip压缩算法的基础。在计算机领域，zip算法主要用于压缩图片和电影，以减少存储它们所需的磁盘空间。这也是为什么科学界将检测意识的过程称为刺激与压缩（zapandzip）的原因。最后，每个人的脑电反应转换成了一个确切的数字，这个数字就是“扰动复杂度指数”，PerturbationalComplexityIndex，PCI。如果大脑没有对磁刺激作出反应，比如大脑皮层受到抑制时就会出现这种情况，又或者脑电仅有微弱的波动，那么PCI的数值就接近于0。而大脑反应复杂度最高时，PCI的数值就接近于1。PCI是一个在0到1之间浮动的数值，数值越高，大脑对磁脉冲的反应就越复杂和丰富多样。

PCI研究的方向非常明确，接下去就是找到PCI的临界值，也就是在0到1的数值间定个位，小于这个临界值，患者就是无意识的，大于这个临界值，患者就是有意识的。这个临界值至关重要。它是能够产生意识的大脑活动复杂度的最小值。那些灰色地带的患者，用常规方法无法知道他们是否存在意识，那么就需要用PCI去判断一下他们的大脑意识水平。

2016年，研究开始了。研究人员在比利时和意大利的一些专科诊所，对受试者使用了这项技术。研究发现，当把PCI的临界值设定为0.31时，可以完全准确地推断出那些健康受试者是否有意识。接着就要进入灰色地带了，先去判断一下MCS，也就是最小意识状态的病人，到底是否存在意识。最小意识状态的那组病人，除了具有基本的反射功能，还存在其他意识行为的迹象。在38例最小意识状态患者中，研究人员正确辨识出了其中36例存在意识，只把2个患者诊断为无意识状态。接着研究人员就对植物人进行实验，在43例无法沟通的植物人患者中，有34个患者被判定为无意识。而另外9名植物人患者的PCI值大于0.31，也就是说，他们大脑反应的扰动复杂性，与许多对照组个体在有意识的状态下一样高。这些患者的皮层反应高度复杂，有可能存在意识，只是无法与外界及亲人沟通。

除了PCI指标，还有一项检测植物人是否具备意识的方法。2010年，英国剑桥大学脑科学博士亚德里安·欧文带领的研究小组通过功能性核磁共振成像技术（fMRI），对23名植物人的脑部进行了扫描，以观察他们对于外界刺激的大脑反应。

这23名植物人都能够像正常人那样交替进入睡眠和醒来的状态，甚至眼睛也可以睁开，但对外界事物毫无反应，言语、意识和思维能力几乎接近于零。根据医学界对于植物人的定义，他们都已经处于了“永久性植物人状态”。这些植物人中，有4人被扫描出，脑部对于外界刺激会产生反应，他们可以想象运动场景，也可以想象空间场景。当他们思考不同问题时，脑部血流会在核磁成像中呈现出不同大脑区域的亮点。

而在这4名植物人当中，最令研究人员震惊的是一个已经深度昏迷5年的29岁男子。他竟然可以对“选择性问题”做出回答。比如，当问到“你的父亲是托马斯吗”，他可以通过思考后作出“不是”的反应，而当问他“那你的父亲是亚历山大吗”，他可以立即做出“是”的反应。研究人员总共问了6个类似的问题，其中包括“你是否有姐妹”等。在6道是非选择题中，他竟然答对了5题！

这是世界首例科学家与永久性植物人进行的“交谈”，当然，这个交谈是打引号的。欧文博士说，这不仅仅是一个简单的检查或者实验，最重要的是，它首次提供了一种方法，可以使病人向外界传达他们的想法。我想，未来，越来越多的技术会帮助植物人通过自我意识来主动选择他们活着还是离去的命运。这些被禁锢在躯壳中的灵魂也有人权。

不过，科学家们还有很长的路要走。比如：如何改善刺激与压缩技术，如何进一步验证PCI的临界值，以保证能100%检测到MCS和植物人患者存在的意识。还有一个问题，是否能找到其他生理学或行为学上的方法，帮助我们证明某些植物人患者真的存在意识。又或者，PCI能否预测植物人患者有多少康复的几率。这些都是PCI研究的漫漫长路。但在寻找答案的同时，我们可以庆幸一下，在探索意识这个古老的问题上，我们终于取得了里程碑式的突破。随着技术的发展，我们还可以去探究婴儿是从什么时候开始有意识的，猫狗又是否有意识，这些都是非常有趣的课题。

大脑神经元

接下来，我们要进入本文最核心的部分了，就是比较大脑神经元与计算机系统在物理结构上的异同。

今天，脑科学家认为，我们每个人所谓的“我”其实就是一层大脑皮质而已。它的外形像一颗核桃，表面布满了褶皱。如果我们把这层大脑皮质取下来摊平的话，大小大约是48厘米见方，就像是放在餐盘底下的那块餐巾布大小，厚度大约是2毫米，比1元的硬币略微厚一点点。

我们的感知、思考、理解、表达、判断以及七情六欲都只不过是这块餐巾产生的电信号。在这块餐巾中，分布着大约200亿个神经元。你可以把神经元想象成是一只

章鱼，只不过这只章鱼的每一根触手都像是一颗大树一样又细分出无数的小触须。

巧合的是，2018年人类能够制造的最复杂的单块芯片所包含的晶体管数量也是200亿个左右。从工作方式的角度来说，晶体管与神经元都相当于一种电位开关，因为它们都是通过两种状态来传递信息，有或者没有动作电位。但是，是不是这样看来，人类制造的集成电路的复杂程度可以媲美大脑皮质了呢？

这样想就太天真了，因为决定复杂程度的不仅仅是单个元器件的数量，还有一个更重要的因素，就是这些元器件之间如何连接。

在集成电路中，主要的器件是二极管、三极管和MOS管，出现最多的就是MOS管，它有四个端，你可以想象成长出四条触手的章鱼，这只章鱼的每条触手又与另外一只或者几只章鱼连接。大致来说，一块包含200亿个晶体管的芯片中，还包含着约1000亿根连接线路。下面这张图是集成电路的一个局部设计版图：

（上图，每一个方块都表示一个“器件”，而线条则表示金属连接线）

我们再来看一下神经元的连接，每一个神经元会和上千个甚至上万个其他神经元连接。包含200亿个神经元的大脑皮质层总共包含了约20万亿个神经连接。所以，仅从连接复杂性上来说，大脑皮质层的复杂程度依然比今天最复杂的一块芯片要高出200倍。

神经元的连接

神经元局部放大图

但这还没完，集成电路一旦制造完毕，所有的元器件和连接就固定下来终生不变了。但是，大脑神经元的连接不是固定不变的，而是可变的。比如我们学习一项新技能，今天学会了，明天就可能生疏，通过不断地重复练习，就能掌握一个技能，而且能够长久不忘。

这个过程其实就是神经元改变自己的形状、位置和与其他神经元之间的连接。虽然我们还不清楚细节，但是我们知道某一些神经元以及由它们组成的某一种固定通路就使得我们能长久掌握一项技能。

听到这里，你可能觉得我们好像已经蛮了解我们的大脑了，其实，脑科学界都承认，我们现在对大脑复杂性的认识依然只是一点皮毛。这就好像刚刚来到了一大片森林的入口，我们只不过知道了这片森林的面积和入口处一些最典型的大树情况，这片神秘的森林中到底还藏着多少令人惊讶的东西，我们真的不知道。

说到这里，我最后想告诉你的结论就是，我们现在能制造的计算机系统的复杂性与大脑相比，至少还有3、4个数量级的差异。在我们没有能力制造出足够复杂的硬件系统之前，就不要奢谈人工智能的觉醒。

美国艾伦脑科学研究所的所长、首席科学家克里斯托夫·科赫认为：意识是达到一定复杂程度的物理系统所具备的一种属性。并且，一个物理系统越复杂，它所拥有的意识水平也就越高。如果未来，人类能打造出比人脑还要复杂的计算机系统，人工智能的意识水平就可能超越人脑。

科赫曾经与DNA双螺旋结构的发现者之一、诺贝尔奖得主克里克一起合作过。他和克里克都认为应该抛弃关于意识的本质到底是什么的争论，这种争论是一种哲学式的思辨，不可能争出结果。科学研究应该更务实一点，寻找一个判定是否具备意识的客观标准，也就是一个可以被测量的物理量，用这个物理量来判定一个系统是否具备意识。他主张，用典型的科学的方法，来研究意识的神经机制，比如，哪些特定的分子、神经元和神经回路与意识有关。可以说，克里克专注于研究意识在大脑中留下的“足迹”，使这个问题回到了科学的正轨上。而我们前面在介绍意识的分界线时，包括科赫在内的测量派已经取得了阶段性的里程碑式的进展，可以说他们找到了第一代用来判定意识是否存在的可测量值。

现在让我们再回到本文一开篇就提出的问题：人工智能觉醒的两条路径，即程序觉醒和意识上传，到底哪条路径更有可能走通呢？

我的回答是：殊途同归！

当我们的计算机系统复杂到足以模拟大脑神经元结构时，两种觉醒就具备了硬件基础。但以我所掌握的知识来看，要实现用计算机来模拟大脑神经元的复杂度，或许只有等到量子计算机的成熟才有可能。但是，无论如何，我不认为有哪条物理法则禁止人工智能的觉醒。

量子计算机

我们现在的计算机称为“电子计算机”，它的最基本的工作单元是二极管、三极管、MOS管。但不论是什么“管”，它的本质就是一个带开关的管道：关上管道，电子通不过，表示0；打开管道，电子通过，则表示1。有了0和1，就可以用二进制来表示一切信息。

但是，这里面有一个最关键的核心点，电子计算机用来记录信息的基本单元其实不是电子，而是开关。每一个开关，我们称之为1位，也就是我们常说的1个比特（bit）。在某一个给定的时刻，信息单元有多少位，就代表能表达多少比特的信息量，

我们现在的操作系统一般是64位的，也就是说用64个开关排成一串作为一个最基本的信息单元。

而量子计算机则从根本上改变了信息的存储方式，它用到了量子的一种奇异性质。像电子、光子这样的基本粒子，我们统称为量子，所有的量子都有一种很奇异的性质，这种性质被科学家们称为“量子叠加态”。

比如说，电子有一种性质叫“自旋”，每当我们去测量一个电子的自旋态时，我们总能随机得到两种结果（A或者B），但非常奇特的是，物理学家们发现，当我们不去测量电子时，电子的自旋态既不是A也不是B，而是同时处在A和B两种状态的叠加态中。不仅是电子的自旋，量子的很多性质都有这种奇异性，例如光子的偏振态

。

量子计算机用来记录信息的基本单元不再是“开关”，而是用电子的自旋态或者光子的偏振态等来记录信息。比如规定，自旋A表示1，自旋B表示0，那么，一个电子就可以同时表示1或者0，因为电子的自旋态在没有观察之前，是处在叠加态中的

。

所以，在任何一个给定的时刻，电子的数量和信息量的关系是2的N次方，这里的N就表示量子的数量，也称之为量子比特（qubit）。

我们来对比一下普通比特与量子比特在信息量上的差异，前者有多少位就有多少比特的信息量，而后者是2的N次方。2的10次方就是1024，指数增长是非常惊人的，如果涨到2的20次方，那么信息量就会变为1048576比特，如果到了30次方，就突破10亿了，如果指数增加到64，我的电脑已经无法计算出2的64次方到底是多大的数字了，估计比银河系中原子的总数还多。

由量子计算机的基本原理可知，评价量子计算机性能的一个最基本指标就是看它能同时精确操纵多少个量子，也就是拥有多少位量子比特。

2013年5月，谷歌和NASA在加利福尼亚的量子人工智能实验室发布D-WaveTwo，这台量子计算机的量子比特数是9位，同时操纵9个量子，换句话说，他的信息长度是2的9次方，也就是512位。2016年8月，美国马里兰大学发明世界上第一台由5量子比特组成的可编程量子计算机。中国科技大学潘建伟团队在2017年5月3日发布了一台光量子计算机，量子比特数是10位，因此，信息长度就是1024位。量子比特数每增加一位，性能都是一次飞跃，难度也是极高的。这说明，我国在量子计算机上的科研水平已经走在了世界的前列。当然，量子比特的位数并不是全部指标，我们在可编程性上还没能超越国际同行。

量子计算机不仅仅是在信息长度上有着电子计算机无法比拟的巨大优势，在计算速度上也有着巨大优势。电子计算机的计算速度取决于开关的频率，频率越高，则算得越快，但不管频率有多高，电子计算机只能按次序老老实实地一次又一次地做加法运算。所以，我们经常看到，评价超级计算机的运算性能都是用每秒钟运算多少次来评价，这个“次”就是一次最基本的加法运算。

但是，量子计算机却可以充分利用奇异的量子叠加态，在同一时间，并行处理加法运算，而不用排队。而且，运算速度会随着量子比特数的增加呈指数级的增加。我在电视新闻中看到潘建伟在一次讲座中就打了个比方，他说如果量子比特数能达到30位，那么我们现在最快的电子计算机用15年才能完成的运算量，量子计算机1秒钟就够了，就是这么夸张的差距。

不过，以目前人类的理论储备，哪怕我们实现了30位可编程的量子计算机，也无法完全取代传统电子计算机。这是因为，量子计算只对特定的、可以用大规模并行计算解决的需求有优势，对普通的文字、图像处理没有任何优势。换句话说，上网看个电影、发个邮件什么的，量子计算机根本发挥不出优势，至少在目前现有的理论框架中，还没有优势。但量子计算理论还是一门非常年轻的学科，随时都有可能在理论上出现突破性的进展，这个就要靠数学家们的努力了。

但是，量子计算机用来模拟超级复杂的对象却有着电子计算机无法比拟的巨大优势，像大脑皮质层的200亿个神经元，如果可以用差不多同等数量级的量子来模拟，那这些量子全部加起来也不会超过一个针尖大小。

量子计算机是完全不同于传统电子计算机工作原理的计算平台，它的编程模式也完全不同于传统的编程模式。虽然我们现在只是刚刚起步，刚刚做出了原型机，但是，这就像70年前的第一台电子计算机、30吨重的埃尼亚克诞生一样。当时的人们绝对想不到今天我们已经能把计算机装进口袋中，运算速度却是埃尼亚克的百万倍。

我们现在也绝对想不到30年或者50年后量子计算机能发展成什么样，量子操纵技术会发展，同时量子计算的原理也会继续发展，这两者的发展相辅相成，未来充满了种种可能性。

未来人会怎样

前面讲的这些还是停留在我们可以把握的近未来。那么如果把时间的尺度加长，在远未来，人类与人工智能的关系又会怎样呢？我想跟你谈谈我的思考。

我先请大家思考一个问题，怎么区分人类智能和人工智能。

你现在可能会觉得很清晰，人工智能不就是计算机中的程序嘛，而人类必须是活的生物体。但是再过100年，可能根据这个定义就很难区分出它们来了。因为我们很快就可以在头脑中植入芯片，利用最新的生物和计算技术来帮助听不到声音的人听到声音；我们还可以为瞎子植入电子义眼，经过计算机芯片处理后，刺激神经产生图像。在这种技术的进步过程中，都需要用到计算机程序，也就是所谓的人工智能。这种技术再往下发展的话，我们就可以直接用脑机接口获取知识。我们可以把这个过程分成四个等级，等级越高，需要的脑机接口的性能也就越高。

这里假设：我需要获取的知识是一本书，比如《三体》。第一级，我在脑中用文字搜索了它，得到了文字的搜索结果；第二级，我访问云端，一秒就回忆起了这本书的全部信息，虽然我压根没看过；第三级，整个云端都是我大脑的一部分，我想起这本书的时候就知道这本书的全部信息；第四级，我需要的时候不但能把书倒背如流，还能对佳句有切身感悟。那么在这种情况下，人工智能和人类的大脑思考这两者之间的区别就已经变得模糊了。

或许再过200年，我们人类就有可能让自己在计算机网络中永生，也就是把自己全部的记忆知识全部上载到现在的计算机云中。到了那个时候，你可能就是一个活在计算机网络中的智能程序了。但是他拥有你全部的记忆和历史，还有你自己的思维方式和观点。如果这一天真的到来，那么到时一个外部的活人，面对这样一个计算机程序中的亲人，他会把他当成一个人工智能呢还是当成一个活人呢？

或许再过300年，我们就可以为这些上载到计算机程序中的程序人外加一个机器人的身躯，当然也可能是生物体的身躯，这就看你自己的喜好了。那么这个时候，这个人到底是一个机器人，还是一个活人，你还能说得清楚吗？

实际上我们已经在朝着这个方向进化，这一切已经发生了。我们每一个人都已经是半机械人，只是你没有意识到而已。现在让我来点破吧。你觉得你的眼睛、嘴巴是不是就是你？你可能一下子没反应过来我这个问题问的是什么。所谓的人其实就是一张展开只有一块桌布大小的大脑皮质层，你所有的记忆、性格、意识都在这里存着，你就是大脑皮质，而人的嘴和耳朵只不过是皮质层的工具。我们用嘴和耳朵这样的工具与其他大脑皮质发生信息交流。而人体的一切器官也只不过仅仅是为了让大脑皮质保持活性的工具而已。那么，我们现在人手一个的手机也是用来让我们和其他大脑皮质交换信息的工具，至于这个工具是拿在手里还是嵌在人的头骨上，有什么本质区别呢？把一个电子心脏塞进你的胸腔里面和别在腰上有什么本质区别呢？很快，我们就要穿戴上各种智能设备了，万物互联的时代也很快会来到，请记住我今天的话，我们已经是半机械人。

我认为自然选择和进化无时无刻不在发生，现在正在发生，将来也一定会继续发生。只要是更适合在这个宇宙生存，那么进化就必然朝着这个方向发展。而在我看来

，“变形金刚”这样的金属生命，显然更适合在宇宙中生存。因为他们不需要水，不需要空气，可以在真空中活着。所以从这个意义上来说，如果我们把时间拉长到几万年来看的话，人类朝着“变形金刚”这种方向的发展是不可逆转的。

很可能在未来的某一天，当人类第一次与真正的外星文明接触的时候，从飞船上走下来的就是这样一个一个的“变形金刚”。如果用时髦一点、更科学的话来说，可以把他们看成是硅基生命体，而我们是碳基生命。这些硅基生命给我们讲述了一个在宇宙中普遍发生的古老故事。那就是在几万年以前他们也曾经是跟我们一样的碳基生命，但是经过了百万年的进化之后，他们终于摆脱了母星，自由地翱翔在太空中。现在请你想象一下，你就是第一个听到这个故事的人。这时候你再回想一下今天社会上关于人工智能会不会毁灭人类的讨论，你大概已经有答案了。

结语

2017年马斯克收购了Neuralink公司，这位科技界的传奇人物开始正式进军脑机接口这个行业。我在一篇对马斯克的采访中，看到了几乎和我在上一节中表达出来的一模一样的看法。

但在马斯克眼中，人类做的所有事情中最可怕的，就是试图唤醒计算机，这种努力开弓没有回头箭，一旦开始，就不可能停止了，人类会一直试图唤醒计算机，直到它真的醒来的那一天，这是正在发生的事情。

现在有两种观点：

一种认为强人工智能会帮我们解决所有问题，

另一种认为我们就是一群把自己不懂的炸弹当玩具玩的小孩。

对于马斯克和其他许多人来说，研发强人工智能是迄今为止人类可能面临的最大的生存危机。应对这个危机的唯一办法就是：成为AI。

而我个人的浅见是：AI化生存是人类演化的必然，不以人类的意志为转移。

人工智能导论知识点总结

以下是人工智能导论中的主要知识点总结：

1.人工智能的定义和历史

-人工智能的定义和目标

-人工智能的历史和发展

2.智能代理和问题解决

-智能代理的定义和分类

-问题解决方法和搜索算法

-问题求解的启发式搜索算法

3.知识表示和推理

-知识表示和表示方法

-逻辑推理和规则推理

-知识表示和推理的应用

4.机器学习

-机器学习的定义和分类

-监督学习、无监督学习和强化学习

-机器学习的应用

5.自然语言处理

-自然语言处理的定义和任务

-词法分析、语法分析和语义分析

-自然语言处理的应用

6.计算智能和神经网络

-计算智能和神经网络的定义和基本原理

-神经网络的类型和训练方法

-计算智能和神经网络的应用

7.人工智能的伦理和社会问题

-人工智能的伦理和社会问题

-人工智能的风险和挑战

-人工智能的未来和发展

以上是人工智能导论中的主要知识点总结。这些知识点是人工智能的基本概念和技术，对于学习人工智能的人来说是非常重要的。

假如人工智能有自我意识后，杀了人，怎么判断是故意还是过失？

由于法律的拟定是需要具体案例来推动的，现阶段从法律条款上并无相应约束。不过按照题主的意思，这是一个脑洞题，那么就进行一番推演。

按照目前法律，脑补可能结果是不追求刑事责任，民事责任厂商负责

类比来说明这个问题，现在把人工智能比作一台榨汁机，只需要按动一个按钮，将食材推压到机器中，就可以获得相应的果汁，就好像下图所示：

现在出现了一点儿小意外，使用者的手被卷了进去，导致手指被榨汁机内的刀头所伤，那么现在对于判断案例可能就会集中在：

1.使用者是否由于违规操作导致，应该自行承担，例如有人开着车，路况正常，车辆正常，自己开到了河里。

2.制造者由于设备的安全隐患导致，对市场上销售了有问题的设备，厂商来赔偿，例如有人开着新买的车，操作没有问题，但是突然车的刹车出了问题导致开到了河里。

上述的所有可能都是建立在榨汁机是属于生产者或使用者高度可控的程度。

那么，增加其“自主意识”属性。现在这台榨汁机突然有能耐了，不完全受生产厂家和使用者控制了。这个时候，是属于其“故意”还是“过失”？

在这个选择题前，需要明确一个关键，即，对于人工智能产品，其所具有的“意识”，法律上是否认可其能够独立承担责任。

现阶段肯定是无法认可，依旧会按照上述两条找责任方是属于生产商还是使用者。但是往长远想，认为一个人工智能产品具有“高度意识”。

还是在现有法律里猜答案，对于未成年人犯罪，先行法律则认为其监护人有承担赔偿责任的责任。也就是对于未成年人，即未满14岁，无刑事责任能力的人，不负刑事责任。所以，不满14周岁的人是不构成犯罪的。

因此，我猜之后的法律更加不会认可，即刑事责任不追责，民事责任依旧会找他的监护人，即“生产厂家”来进行赔偿。也就是，只需要其生产厂家对被害人进行赔偿。

人工智能的真实案例是怎么判的

先看一个旧闻。2018年3月，苹果工程师WalterHuang（黄伟伦）开着自己的特斯拉ModelX上班时不幸发生了致死车祸，他的车一头撞上了高速公路下道口的隔离墩。当时，特斯拉Autopilot处于开启状态，但却未能识别路上的障碍。更可怕的是，在车辆接近隔离墩时，它还控制车辆加了速。

事故发生后，美国国家运输安全委员会（NTSB）开始着手对其进行深入调查，最近它们才终于公布了调查报告。

报告内容大家有兴趣可以自己来看，我这边先说一下具体的结果：

1.WalterHuang要为自己的意外去世买单，因为他心里很清楚Autopilot在这个下道口力有不逮（此前曾两次遇到过类似问题）。此外，开车打游戏也是他的不对。现在的Autopilot只是一套ADAS系统，司机不能指望它“全知全能”。

2.负责公路管理的Caltrans在工作上也有需要完善的地方，比如尽快修复受损设施并更新道路标线。

3.此次事故特斯拉不用负什么责任。

目前的人工智能，还不是题主说的具有“自我意识”，但是可以看到，目前对待科技的态度，还在找到事故背后人的因素。

总之

不要过度信赖科技，看到不伤手的刀，也不要用自己的手去试，毕竟最后依旧是人来买单，刀还是刀。

人工智能 选择题的介绍就聊到这里吧，感谢你花时间阅读本站内容，更多关于人工智能选择题库、人工智能 选择题的信息别忘了在本站进行查找哦。