

AlphaGo与人工智能



在之前的一篇[文章](#)中我指出，自动驾驶所需要的“视觉识别能力”和“常识判断能力”，对于机器来说是非常困难的问题。至今没有任何机器可以在视觉方面达到驴的水平，更不要说和人比。可是最近Google的[AlphaGo](#)战胜了围棋世界冠军，挺闹腾的，以至于对AI的误解又加深了。

本来玩个游戏而已，恁要吹成是“历史性的人机大战”，说得好像是机器挑战了人类的智能，伤了人类的自尊似的。这整个项目打着一个相当高大上的招牌，叫做“[Deep Mind](#)”。当然，其中的技术也有一些吓人的名字，什么“神经网络”啊，“深度学习”啊.....

听到这些，总有一知半解的人，根据科幻电影的情节开始展望，这样厉害的技术，应该可以用来做更加“智能”的事情，然后就开始对“人类的未来”作出一些猜想，比如自动车就要实现，人的工作很快都要被机器取代，甚至[Skynet](#)就要控制人类，云云。

我只想在这里给这些人提个醒：还是别做科幻梦了，回到现实吧。

棋类是相对容易的AI问题

一个常见的外行想法，是以为AlphaGo真的具有“人类智能”，所以Google利用同样的技术，应该可以实现自动车。这些人不但大大的高估了所谓“AI”的能力，而且他们不明白，不同的“AI问题”的难度，其实有着天壤之别。

围棋是简单的，世界是复杂的。机器视觉和自动车，难度比围棋要大许多倍，根本不在一个量级。要达到准确的视觉判断能力，机器必须拥有真正的认知能力和常识，这并不是AlphaGo所用的树搜索和神经网络，就可以解决的。由于需要以极高的速度处理“模拟信号”，这根本就不是人们常用的“数字计算机”可以解决的问题。也就是说，不是写代码就可以搞定的。

很早以前，人工智能专家们就发现一个很有趣的现象，是这样：

- 对于人来说很难，很烦的事情（复杂的计算，下棋，推理.....），对于计算机来说，其实算是相对容易的事情。
- 对于人来说很容易的事情（认人，走路，开车，打球.....），对于计算机来说，却非常困难。
- 计算机不能应付复杂的环境，只能在相对完美的环境下工作，需要精确的，离散的输入。
- 人对环境的适应能力很高，擅长于处理模糊的，连续的，不完美的数据。

从以上几点你可以看出，棋类活动正好符合了计算机的特点，因为它总是处于一种隔离的，完美的环境，具有离散的，精确的，有限的输入。棋盘上就那么几十，几百个点，不是随便放在哪里都可以的。一人走一步，轮流着走，不能乱来。整个棋盘的信息是完全可见的，没有隐藏和缺损的信息。棋局的“解空间”虽然很大，却非常规整，有规律可循。如果完全不靠经验和技巧的话，围棋的第一步可以有361种情况，第二步有360种情况，.....

这对机器是非常有利的情况，因为计算机可以有计划有步骤，兢兢业业的把各种可能出现的情况算出来，一直到许多步以后，然后从中选择最有优势的走法。所以下棋归根结底，就是一个“树搜索”问题，只不过因为规模太大，需要加入一些优化。围棋的解空间虽然大，却是一个已知数，它最多有 250^{150} 种情况。AlphaGo使用所谓“神经网络”，就是为了在搜索的时候进行优化，尽早的排除不大可能取胜的情况，免得浪费计算的时间。

这种精确而死板的活动，就跟计算一个比较大的乘法算式（比如 2463757×65389 ）的性质类似，只不过规模大很多。显然，人做这类事情很繁，很累，容易出错，计算机对此却任劳任怨，因为它本来就是机器。当年“深蓝”战胜国际象棋世界冠军的时候，我就已经预测到，计算机成为围棋世界冠军是迟早的事，所以没必要玩这些虐待自己脑子的游戏了。可惜的是，挺多人仍然把精通棋艺作为一种荣耀（因为“琴棋书画剑”嘛）。很多中国人认为，中国人下围棋总是输给韩国人，是一种耻辱。现在看来这是多么可笑的事情，这就像心算乘法不如韩国人快，就觉得是耻辱一样：)

认知是真正困难的AI问题

现在来对比一下人们生活中的琐事，就说倒水端茶吧。



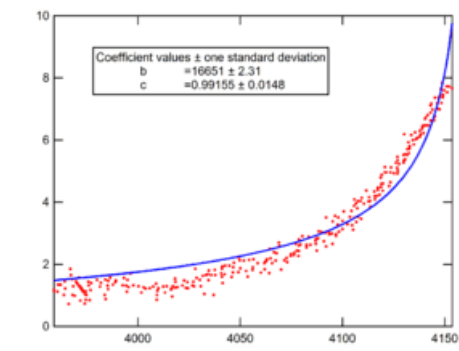
让一个机器来给你倒水，有多难呢？意想不到的难！看看这个场景，如果你的电脑配备有摄像头，那么它怎么知道茶壶在哪里呢？要知道，茶壶的材料，颜色，形状，和角度，可以有几乎无穷多的变化。甚至有些茶壶跟哈哈镜一样，会把旁边的物体的形状都扭曲反射出来。桌上的物品附近都有各种反光和阴影，不同材料的反光特性还不一样，这些都会大幅度的影响机器对物品的识别。

为了识别物体，机器需要常识，它的头脑里必须有概念，必须知道什么样的东西才能叫做“茶壶”和“茶杯”。不要小看这一步的难度，这意味着机器必须理解基本的“拓扑结构”，什么叫做“连续的平面”，什么叫做“洞”，什么是“凹”和“凸”，什么是“里”和“外”..... 另外，这机器必须能够分辨物体和阴影。它必须知道水是什么，水有什么样的运动特性，什么叫做“流动”。它必须知道“水往低处流”，然后它又必须知道什么叫“低”和“高”..... 它必须知道茶杯为什么可以盛水，茶壶的嘴在哪里，把手在哪里，怎样才能拿起茶壶。如果一眼没有看见茶壶的把手，那它在哪里？茶壶的哪一面是“上面”，要怎样才可以把水从茶壶的嘴里倒出来，而不是从盖子上面泼出来？什么是裂掉的茶杯，它为什么会漏水，什么是缺口的茶杯，它为什么仍然可以盛水而不漏？干净的茶杯是什么样子的，什么是脏的茶杯，什么是茶垢，为什么茶垢不算是脏东西？如何控制水的流速和落点，什么叫做“水溅出来了”，要怎么倒水才不会溅出来？

你也许没有想到，倒茶这么简单的事情，需要用到如此多的常识。所有这些变数加在一起，其实远远的大于围棋棋局的数量，人却可以不费力的完成。这能力，真是应该让人自己都吓一跳，然而人却对此不以为然，称之为“琐事”！因为其他人都可以做这样的事情，甚至猴子都可以，怎么能显得出我很了不起呢？人的自尊和虚荣，再一次的蒙蔽了他自己。他没有意识到，这其实是非常宝贵，让机器难以匹敌的能力。他说：“机器经过大量的学习，总有一天会做到的。看我们有神经网络呢，还有深度学习！”

机器学习是什么

有些人喜欢拿“机器学习”或者“深度学习”来吓唬人，以为出现了“学习”两个字，就可以化腐朽为神奇。而其实所谓机器学习，跟人类的学习，完全是两回事。机器的“学习能力”，并没有比石头高出很多，因为机器学习说白了，只不过是大量的数据，[统计拟合](#)出某些函数的参数。



比如，你采集到一些二维数据点。你猜测它们符合一个简单的函数 $y = ax^3 + bx^2 + cx + d$ ，但不知道a, b, c和d该是多少。于是你就利用所谓“机器学习”（也就是数学统计），推断出参数a, b, c和d的值，使得采集到的数据尽可能的靠近这函数的曲线。可是这函数是怎么来的呢？终究还是人想出来的。机器无论如何也跳不出 $y = ax^3 + bx^2 + cx + d$ 这个框子。如果数据不符合这个范式，还是只有靠人，才能找到更加符合数据特性的函数。

所谓神经网络，其实也是一个函数，它在本质上跟 $y = ax^3 + bx^2 + cx + d$ 并没有不同，只不过输入的参数多一些，逻辑复杂一些。“神经网络”跟神经，其实完全没有关系，却偏喜欢说是受到了神经元的启发而来的。神经网络是一个非常聪明的广告词，它不知道迷惑了多少人。因为有“神经”两个字在里面，很多人以为它会让机器具有智能，而其实这些就是统计学家们斯通见惯的事情：拟合一个函数。你可以拟合出很好的函数，然而这跟智能没什么关系。

AlphaGo并不是人工智能历史性的突破

这次AlphaGo战胜了围棋冠军，跟之前IBM的“[深蓝](#)”电脑战胜国际象棋世界冠军，意义其实差不多。能够写出程序，在这些事情上打败世界冠军，的确是一个进步，它肯定会对某些特定的应用带来改善。然而，这并不说明AI取得了革命性的进步，更不能表明电脑具有了真正的，通用的智能。恰恰相反，电脑能够在棋类游戏中战胜人类，正好说明下棋这种活动，其实并不需要很多的智能。从事棋类活动的能力，并不足以衡量人的智力。

著名的认知科学家[Douglas Hofstadter](#)（《[GEB](#)》的作者），早就指出AI领域的那些热门话题，比如电脑下棋，跟真正意义上的人类智能，几乎完全不搭边。绝大部分人其实不明白思考和智能到底是什么。大部分所谓AI专家，对人脑的工作原理所知甚少，甚至完全不关心。

AlphaGo所用的技术，也许能够用于其它同类的游戏，然而它并不能作为解决现实问题的通用方法。特别是，这种技术不可能对自动车的发展带来突破。自动车如果只比开车技术很差的人强一点，是不可接受的。它必须要近乎完美的工作，才有可能被人接受，然而这就要求它必须具有人类级别的视觉认知能力。比如，它必须能够察觉到前面车上绑了个家具，没绑稳，快要掉下来了，赶快换车道，超过它。可惜的是，自动车的“眼睛”里看到的，只是一个个的立方块，它几乎完全不理解身边到底发生着什么，它只是在跟随和避让一些线条和方块..... 我们多希望马路都是游戏一样简单，清晰，完美，没有意外的，可惜它不是那样的。每一个细节都可能关系到人的生死，这就是现实世界。



为AlphaGo热血沸腾的人们，别再沉迷于自动车和Skynet之类的幻想了。看清AI和“神经网络”的实质，用它们来做点有用的东西就可以，没必要对实现“人类智能”抱太大的希望。