

忙总谈数学：外行瞎扯人工智能

January 22, 2022

Abstract

趁最近讨论计算机围棋，凑个热闹。

我没有搞过人工智能，只是在80年代初参加某军方课题时，因为我参加的模块必须大量使用属于人工智能专业的知识分类，知识提取，知识辨识和专家系统等等内容，所以和人工智能模块人员有很多接触和合作，道听途说知道一点。

我只说说我自己的认识，与我们国家具体做过什么和做到什么水平无关。而且我不是这个专业的，所以可能是乱说。由于我90年代初离开中科院，所以我知道的人工智能的东西只限于80年代，以后的就不知道了。

豆瓣小组“管理实践与学习”

<https://www.douban.com/group/542139/?ref=sidebar>

外行瞎扯人工智能

来自: wxmang 2016-03-17 10:43:52

<https://www.douban.com/group/topic/84570941/>

趁最近讨论计算机围棋，凑个热闹。

先声明我没有搞过人工智能，只是在80年代初参加某军方课题时，因为我参加的模块必须大量使用属于人工智能专业的知识分类，知识提取，知识辨识和专家系统等等内容，所以和人工智能模块人员有很多接触和合作，道听途说知道一点。

不过由于这是一个水很深的领域（30多年过去了，课题尚未解密），而且我离开太久，所以不要指望我说什么干货。

我只说说我自己的认识，与我们国家具体做过什么和做到什么水平无关，这样大家都幸福。而且我不是这个专业的，所以可能是乱说。由于我90年代初离开中科院，所以我知道的人工智能的东西只限于80年代，以后的就不知道了。

现在使用的人工智能概念是在1956年提出的。基本定义是：

人工智能是计算机科学的一个分支，人工智能是对人的意识、思维的信息过程的模拟，人工智能不是人的智能，主要为生产能以人类智能相似的方式做出反应的智能计算机服务。

目前人对自身智能的理解非常有限，对构成人的智能的必要元素也了解有限，所以人工智能的研究目前只集中在与知识有关的部分，例如知识分类，知识创造，知识学习，知识获取方面，一般说来包括计算机人、语言识别、图像识别、自然语言处理、深度学习和专家系统等。

人类的情感，审美，信仰等等高级阶段的智能活动尚不能介入。

所以美国斯坦福大学人工智能研究中心尼尔逊对人工智能的定义是：

人工智能是关于知识的学科——怎样表示知识以及怎样获得知识并使用知识的科学。

美国麻省理工学院的温斯顿定义是：人工智能就是研究如何使计算机去做过去只有人才能做的智能工作。

形象解释一下上述定义。

我们都知道利用计算机解决问题时，必须有程序，程序其实就是步骤，所以只要一个问题能够分解成步骤，就能用计算机解决。这个就是算法可以解决的，不需要什么人工智能。对一些已经有步骤，但是过于繁琐，需要优化步骤，缩短解决问题时间和难度的，也不是人工智能，也只是数学优化。

可是大量问题往往是一开始是没有步骤的，步骤是穷举试错出来的，是经验归纳的结果，如果计算机能够通过穷举试错，归纳演绎出解决问题的步骤，也即自动给出算法，而且这种归纳演绎能力因学习而得到不断的提高，这就是人工智能了。

目前人工智能主要是利用计算机来模拟人的基于知识的思维过程和智能行为，如学习、推理、思考、规划等，重点还是模拟人类的逻辑思维，尚不能研究灵感，审美，信仰等等形象思维、灵感思维。

人工智能其实起源于维纳发现的反馈控制（最简单的反馈控制是自动调温器，将收集到的房间温度与希望的温度比较，并做出反应将加热器开大或关小，从而控制环境温度），维纳从理论上证明所有的人类逻辑思维活动都是反馈机制的结果，而反馈机制就可以用计算机模拟。维纳实际上是早期人工智能理论的奠基人。

1955年西蒙（我们曾经在管理帖子里面介绍过这个唯一获得诺贝尔经济学奖的管理学家，其实他也是人工智能的奠基人之一）和他的团队写出逻辑专家

(LOGIC THEORIST) 程序，这个程序是人类第一个人工智能程序，它将每个问题都表示成一个树形模型，然后选择最可能得到正确结论的那一枝来求解问题（这次下围棋的阿尔法狗仍然是这个结构，采用的是蒙特卡洛树型搜索算法）。

逻辑专家是人工智能的里程碑。

后来是所有人工智能进展，其实都是在逻辑专家基础上，不过都是如何减少搜索和建立可以自我学习的系统。

1957年逻辑专家团队写出通用解题机(GPS) 程序，这个程序在树型搜索算法上增加了维纳的反馈原理。

1958年MCCARTHY创造出人工智能专业语言：LISP语言。LISP的意思是表处理(LIST PROCESSING)，很快人工智能标准语言。

1963年开始，美国军方开始资助人工智能，大量人工智能项目开始保密。（其实现在世界大国中，在精确制导武器和洲际打击武器中，在防空反潜武器中，在电子战中，在指挥自动化中，大量使用了人工智能成果）。

70年代以后，出现了大量的人工智能程序，例如SHRDLU，主要是在有限数量的几何形体空间中的逻辑问题自动编程（有点像吴文俊先生的计算机证明）；STUDENT主要是计算机解决代数问题；SIR可以理解简单的英语句子。

70年代另一个进展是专家系统。专家系统可以预测在一定条件下某种解的概率（参考以前介绍过的贝叶斯思路）。到80年代，专家系统被用于股市预测，疾病诊断，确定矿藏位置，反潜，防空，情报辨识等等。

70年代还发展了计算机视觉，例如如何通过一副图像的阴影，形状，颜色，边界和纹理等基本信息辨别图像，推断出图像可能是什么（现在是卫星遥感情报判读的主要方法）。

1972年人工智能更常用的语言PROLOGE语言诞生。

1980年代开始，人工智能开始实用化，国防项目和军火企业大量使用，而且制造业开始使用，例如DEC用XCON专家系统为VAX大型机编程，杜邦，通用汽车公司和波音公司开始在产品研发阶段使用专家系统，而且为了查找和改正现有专家系统中的错误，又有另外一些专家系统被设计出来。

对普通民用产品而言，80年代后常见的是计算机视觉，例如利用生产线上的相机和计算机进行质量控制（主要通过识别产品标识和精确识别形状的不同来控制质量）。

80年代开始，美军开始战地任务计算机人项目（现在已经基本成功）。

同时80年代模糊逻辑（不确定条件下的决策）和神经网络（用计算机模拟人类大脑）技术被发展起来，而且计算机视觉，指纹识别，人脸识别，视网膜识别，虹膜识别，掌纹识别，专家系统，自动规划，智能搜索，定理证明，计算机博弈，自动程序设计，智能控制，计算机人，语言和图像理解，遗传编程等等开始实用化。

下面先介绍一下图灵测试。

1950年，图灵在《计算计算机与智能》这篇论文中讲到：计算机有可能具有人的智能，而“学习计算机”可能就是实现计算机智能（当时尚无人工智能这个概念）的一种方法。

如今的人工智能其实都是图灵所论述的图灵机。

这篇论文主要观点是：

1、提出了判断计算机是否能思考、是否具有智能的方法，即“图灵测试”。

2、论证所有的数字计算机都是通用的、等效的。

3、反驳了9种认为计算机不能思维、没有智能的观点。

4、提出了计算机智能的实现方法，即如何设计一台“学习计算机”。

但是“图灵测试”只是一种判断计算机是否有智能的方法，无法证明计算机一定可以有智能（或者一定不能有智能）。实际上2014年俄罗斯数学家设计的一款软件（Eugene Goostman），基本推翻了图灵的这个假设，也即图灵测试无法判断计算机是否有智能。

所以图灵没有证明计算机一定可以有智能（或者一定不能有智能）。

图灵也没有完成“学习计算机”的设计，也没有让这台计算机通过“图灵测试”。

上面这些内容不是这篇帖子的重点，因为太专业，我只想介绍点业余的内容，那就是图灵在论文中其实有一个悖论，这个悖论可以诱导我们思考什么是人工智能。

图灵测试的要点是：如果一台计算机能够与人类展开对话而不能被辨别出其计算机身份，那么称这台计算机具有智能。

例如一个人（代号C）使用测试对象皆理解的语言去询问两个他不能看见的对象任意一串问题。对象为：一个是正常思维的人（代号B）、一个是计算机（代号A）。如果经过若干询问以后，C不能得出实质的区别来分辨A与B的不同，则此计算机A通过图灵测试。

但是这个测试其实是一个悖论。

在图灵测试中，图灵有一个非常强的假设或先决条件：测试者（询问者）必须是人，或者说只有人才能以测试者的身份，来判断一个被测试者是计算机还是人。

这个先决条件本质是什么呢？显然就是：人的智能要高于计算机，因为人至少会一样计算机不会的事情：测试一台计算机的智能。

所以图灵在测试之前实际上已经假定：人的智能永远高于计算机，如果测试结果是：人的智能与计算机相同或人的智能低于计算机，那么结论与前提就会矛盾，就出现悖论，但是如果结论是：人的智能高于计算机，那就只前提的重复，而毫无价值，因为这是我们测试的假设，这种测试就没有意义。

所以，用图灵测试就会得到矛盾，也即图灵测试无法判断人的智能高于计算机或者低于计算机，计算机不可能像人一样思考或者能够想人一样思考，计算机不可能具有人的智能或者有人的智能。

由于图灵机是现代计算机的基础结构，所有计算机在图灵机意义上等价，也就说目前的计算机技术无法制造和判断人工智能是否存在。

图灵在论文中反对神学的观点（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一颗不朽的灵魂，但从未将它赋予任何其他的动物或计算机。因此，动物或者计算机不能思维）：在试图制造这样的计算机时，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。

但是图灵测试实际上也无礼地篡夺了计算机作为测试者的权力。也即

图灵赋予人唯一的“测试者”功能，与上帝赋予人唯一的“不朽的灵魂”其实本质一样。上帝赋予人唯一的“不朽的灵魂”是不合理的，那么图灵赋予人唯一的“测试者”功能就同样是不合理的。

所以图灵测试（人是唯一的测试者）假设不成立。

但是如果测试者是计算机，那么当计算机来判断你是一个人还是一台计算机时，或者计算机认为你是一台计算机，这会是什么结果，尤其是在一些涉及经济利益（例如法律文书签字）和人身安全（例如选择手术治疗或者去生产线维修）情况时，画面美得不敢简直不敢想象。

计算机认为你是计算机，你就能把自己当成计算机，按照计算机的方式去维护生存？

所以图灵机（现代计算机）建立在假设计算机是没有能力判断计算机或者人有没有智能的，所以必须依赖人来判断自己有没有智能，否则计算机就没有智能。

所以计算机的智能只能来自于测试者也就是人的判断。也就是现代所有基于图灵机结构的计算机，必须经过人的测试才能具有智能，这个命题等价于：

所有智能计算机都离不开人，因为它们必须由人来判断它们的智能。

再推广一下：计算机必须经常性地由人来判断智能。一旦脱离人的判断范围，计算机就将丧失智能。人说它有它就有，说它没有它就没有。

所以人类掌握了最后判断计算机有没有智能的能力，所以只要是基于图灵机结构的计算机，不管性能多么超级，都只是人的工具，不存在超越人类的可能。

所以人于计算机的本质区别是：人具有判断智能的能力，一个人不需要由另外一个人或计算机判断自己有无智能，他自己就能做。计算机则必须依靠人来判断。

其实目前计算机的运算速度、记忆容量和逻辑单元的数目都超过了人脑，而且还有很多人类无法掌握使用的算法，但是他们仍然不是人类：因为他们目前无法创造自己判断自己智能的能力，不能创造出新的方式，新的概念，并利用公理对概念进行推理，得到人类还没有的知识。

其实这种看法在1980年西尔勒提就用“中文房间”实验论证了。

西尔勒的实验是：图灵测试的问题用中文而不是用英文来讲，这显然是非本质的改变，因为信息量一样。

西尔勒自己不懂中文，所以这些问题只是一些符号，但是只要正确地执行了那些构成算法的一系列运算——已给他用英文写的这一算法的指令，他就能和一位真正理解这一问题的中国人做得一样好。所以机器可以在不理解问题前提下成功执行算法。

所以证明：仅仅执行程序的计算机本身并不具有智慧，就算是他真的通过了图灵测试，人类还是不能断定其真有理解能力。

所以图灵测试来定义人工智能是不够充分的。

这也就说目前不存在我们常识理解的人工智能。

下面介绍一下目前人工智能的研究路径。

目前人工智能主要是采用对人的思维的信息过程的模拟来研究。因为我们已经证明了运算能力和数据存储和快速提取能力并不能让计算机变得有智能。

对于人的思维模拟目前有两条道路，一是结构模拟，也即仿照人脑的结构机制，制造出“类人脑”的计算机；二是功能模拟，对思维过程进行模拟。

结构模拟也即逆向工程人脑。不过目前还没戏，因为目前还不知道为什么人脑能够如此高效、快速的运行，并且能从中获得灵感来进行创新。

结构模拟的主要方法是人工神经网络，大概可以想象为由晶体管作为“神经”组成的网络，晶体管和其它晶体管互相连接，有自己的输入、输出系统，通过做任务来自我学习，比如下围棋。目前基本算法是树形搜索算法（例如阿尔法狗的蒙特卡洛树型搜索算法），这种算法的最开始的神经处理和猜测都是随机的，但是当它得到正确的回馈后，相关晶体管之间的连接就会被加强；如果它得到错误的回馈，连接就会变弱。经过一段时间的测试和回馈后，这个网络自身就会组成一个智能的神经路径，而处理这项任务的能力也得到了优化。人脑的学习是类似的过程。

更加极端的模拟方式是“整脑模拟”，把人脑切成很薄的片，用软件来准确的组建一个3D模型，然后把这个模型装在强力的计算机上。不过目前人类刚刚能够模拟1毫米长的扁虫的大脑，这个大脑含有302个神经元。人类的大脑有1000亿个神经元，所以还遥遥无期。

人脑的功能模拟，主要是模仿生物演化过程。这种方法目前主要是基因算法。

基本思想是：建立一个反复计算的表现/评价过程程序（例如生物生存通过能否繁殖后代为评价）。一组计算机程序并行执行各种任务，最成功的将会繁殖，把各自的程序融合，产生新的计算机程序，而不成功的将会被剔除。也即计算机自己修改自己的代码，改进自己的结构和算法。经过多次的反复后。这个自然选择的过程将产生越来越强大的计算机程序。而这个方法的难点是建立一个自动化的评价和繁殖过程，使得整个流程能够自己运行。由于融合经常产生性能突变，就可能创造出超出人类控制的程序，也即产生智能。这也就是计算机实现递归的自我改进，越来越聪明。

不过目前还进展缓慢。

最后再说说计算机学习，从数学角度让，计算机学习就是多重非线性变换对数据进行多层抽象的算法。深度学习是计算机学习中表征学习方法的一类。

现在已有多种深度学习框架，如神经网络和卷积神经网络等等，已被应用于计算机视觉、语音识别、自然语言处理等领域。

目前深度学习其实只是统计学算法，通过数据分析和统计，得到相关性结果，解决语音识别、图像识别、自然语言处理中的问题。

但它并不能得出因果关系，例如它可以把词汇与图片关联起来，因为这就是数据库匹配，如果加上模糊搜索需求，也就是一个数据库匹配的统计问题。但它不能真正理解词汇的含义。所以目前深度学习缺乏表达因果关系的手段，缺乏进行逻辑推理的方法，而且远没有具备集成抽象知识，例如物品属性、代表和典型用途的信息。

人工智能，还在婴儿摇篮阶段。

纯属非专业瞎扯，不要当真。

问答部分（注意：排列不一定是按照时间顺序）

踏月无声：这次AlphaGo与深蓝完全不一样，用的是非常通用的技术——深度学习，可以广泛用于其它问题。事实上自从2010年CNN技术（卷积神经网络，深度学习技术一种）在微软用于语音识别，将错误率一下子降低了30%，那以后美国各大技术公司开始了深度学习方面的“军备竞赛”，并在语音

识别、图像识别、机器翻译、广告推荐系统等领域取得了突破性进展。深度学习目前在理论上处于一种比较尴尬的状态，由于缺乏理论基础经常被理论界鄙视。比如说，为什么AlphaGo的CNN是13层，而FaceBook的围棋程序DarkForest却是12层？没人知道为什么。这些数字都是在实践中调出来的。人工智能史就是一部“打脸史”，多少大牛的预言都落了空。不知道缺乏理论支持却在实践中效果良好的深度学习这次能走多远？

踏月无声：对我来说这次最震惊的不是AlphaGo赢了，而是AlphaGo的算法这么简单（理论上很简单，工程上还是有相当复杂度的），居然就把围棋这么复杂的问题解决得七七八八了。不过就是蒙特卡洛搜索（MTCS）算法加上两个深度学习网络嘛！策略网络（Policy Network）负责选点，价值网络（Value Network）负责形势判断，前者对应棋手说的所谓“棋感”，后者对应所谓“大局观”。赛前都以为人类的棋感和大局观对AI有优势，计算力则不如。结果是AI在棋感和大局观上完全碾压了人类，反而是在复杂局面下的计算出了误差（问题出在MTCS的搜索深度，据说只有25），在第四盘给李世石捡了漏。

刹那芳华：外行跟贴一下。。。不过我曾经十分接近人工智能。。。以前我做了几届计算机系的研究生班主任，其中一个专业就是“模式识别与人工智能”，其他还有“计算机软件”、“计算机应用”等等。同学们都很好，有趣的是评奖学金的时候，原来是按照全系统一论文权重打分。模式班的同学总是论数目文要比其他专业的多，被SCI/EI检索的也多，就造成奖学金分数排行榜前面大多数是模式班的。模式的同学分遥遥领先其他专业。。之后是软件和应用，有些是涉及偏硬或结构的专业，论文难发，还不容易被检索，一个专业连1个名额都排不到。。。同学们有来查询和发牢骚的。这个的确不太好，后来规则更改为总名额的四分之三按照专业人数分配给各专业（专业内部排序），四分之一为机动，大家全系pk。这样才让大家基本满意。。。不过话说回来，人家模式文章多也是努力多的结果啊。。。现在不做班主任有六年多了，过节还是会收到同学们的祝福短信，聊聊得知职业发展都很好。也是开心事。。。我也差点儿去考模式识别与人工智能的博士。后来有畏难顾虑没有考，现在离学术很远很远了，离计算机也很远很远了。。。

知行合一：拜读了！忙总对区块链技术的看法？

wxmang：不了解这个专业，不过银行现在采用的分布式数据库可能就是你说的区块链。这个技术应该只是一个噱头，因为不同企业间不存在任何数据交换和共享的可能，例如用户数据，定价数据，成本数据，研发数据等等是企业主要数据，但是这些数据涉及企业的商业秘密和生命要点，不可能开放给任何系统，连自己的副总裁往往都看不到全部数据。

一休哥们儿：这东西要是平移到无人机上，无人机扫描战场形势，通过价值网络和策略网络选出最优打击点，完全可以实现自主打击。如果无人机群之间再加上联网，实现分布式计算，就有点类似于电影《蚁人》里的场景了。

wxmang：我以前的室友就是搞组织无人机打群架的，现在都因功晋升少将了。

乌金沙：30多年来，AI这个领域进步确实有限。忙总下判断很谨慎。这一波商业炒作，连带学术炒作，的确会给未来带来新变数。专家系统会从中央扩散至地方。福祸难料。

tswang: 这一拨人工智能热, 恰恰是摆脱了专家系统的局限. 以前所谓专家系统, 不要靠人整理规则, 一条条都是人工设定的. 现在这一拨主要靠数据量, 专家的作用是在前期准备数据, 后期评估系统效果, 中间阶段靠算法就可以搞定. 类似alphago, 写程序做系统的团队里面没有真正的职业棋手. 这条路子大大提高了解决问题的能力, 是非常大的进步.

wxmang: 你理解的专家系统其实只是里面的专家库或专家知识库, 专家系统很复杂, 不但要有实时监测数据, 有海量数据快速存储提取, 还要有判断系统 (或者叫决策系统), 一般是是一个树型搜索算法构成. 例如我们在三峡库区建立的地质灾害专家系统, 就是依靠实时监测数据和专家知识经验, 再加上搜索算法, 判断地质灾害可能发生的概率, 并实现24小时预警, 启动应急预案. 专家知识库在专家系统中占的比例不大, 建立的工作量不到10%.

wxmang: 你对专家系统理解望文生义了, 专家系统的基础是知识辨识、知识提取和知识生成, 是由算法来自动实现的.

塔里木河: 人工智能的特点之一就是: 不论是否曾经研究过基本的算法, 是不是研究过图灵机, 每个人都可以就这个问题发表自己的看法. 这是一个大家喜闻乐见的话题.

wxmang: 就像在网络上讨论政治军事话题一样可笑.