

当我们谈AI，我们在谈什么-2

定义与历史

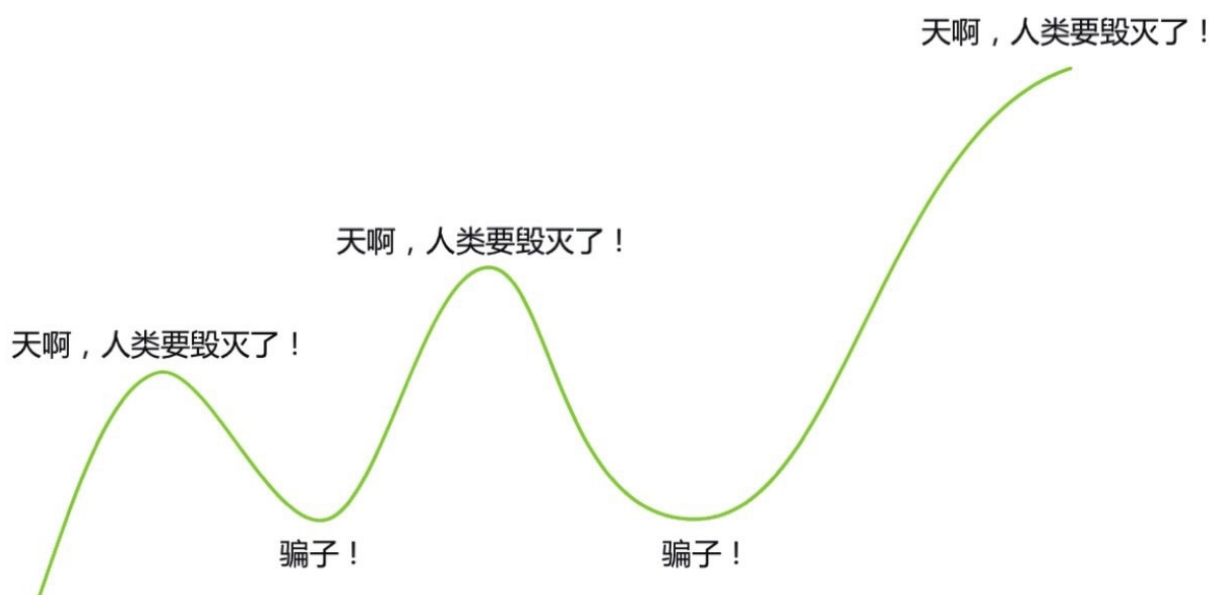
我本来的计划，是按照自己学习的顺序，从技术层面入手，先写技术笔记。但是考虑到技术晦涩难懂，篇幅巨大，所以我决定还是从AI的定义与历史谈起。

知乎上有一句名言，“先问是不是，再问为什么。”讨论一件事情，先严格界定定义和讨论范围，剔除那些因为语言本身的衍生意义，情感色彩，才能避免很多无意义的争执。

AI的定义

所以AI究竟是什么呢？

在试图研究这个问题之前，可以先看一副网上流传的有趣漫画，人工智能发展成熟度曲线：



这幅画既搞笑，但是又反映了AI发展的现实。

实际上，每一次AI热潮的兴起，都会伴随着人类毁灭的论调和之后的幻灭期。在AlphaGo之前，我们至少喊过两次人类毁灭了。之所以会出现这样的波动，主要还是因为对于AI究竟是什么，普罗大众缺乏清晰的认识。

可能在大部分的心目中，谈及人工智能，想到的至少都是长相和人一样，能够与人交流，完成人指定的任务的机器人。

然而实际上，我们早已习以为常的搜索引擎、电商的推荐算法、今日头条的推荐算法、女生爱用的美图、瘦脸功能，最近兴起的智能助理（微软小冰、苹果Siri）、机器翻译、机器写作、机器视觉、自动驾驶等等都属于AI的范畴。

那么，究竟什么是AI？

目前AI领域的领袖们都给出了自己不同的定义。比如在Stuart Russell和Peter Norvig在他们著名的教科书《人工智能：一种现代的方法》中就这样定义：

人工智能是有关“智能主体的研究与设计”的学问，而智能主体是指一个可以观察周遭环境并作出行动以达致目标的系统。

又比如李开复认为，历史上曾经存在几种并且现在仍流行的人工智能的定义。（这里考虑的都是通俗定义，而不考虑更为技术化的形式化定义）[1]

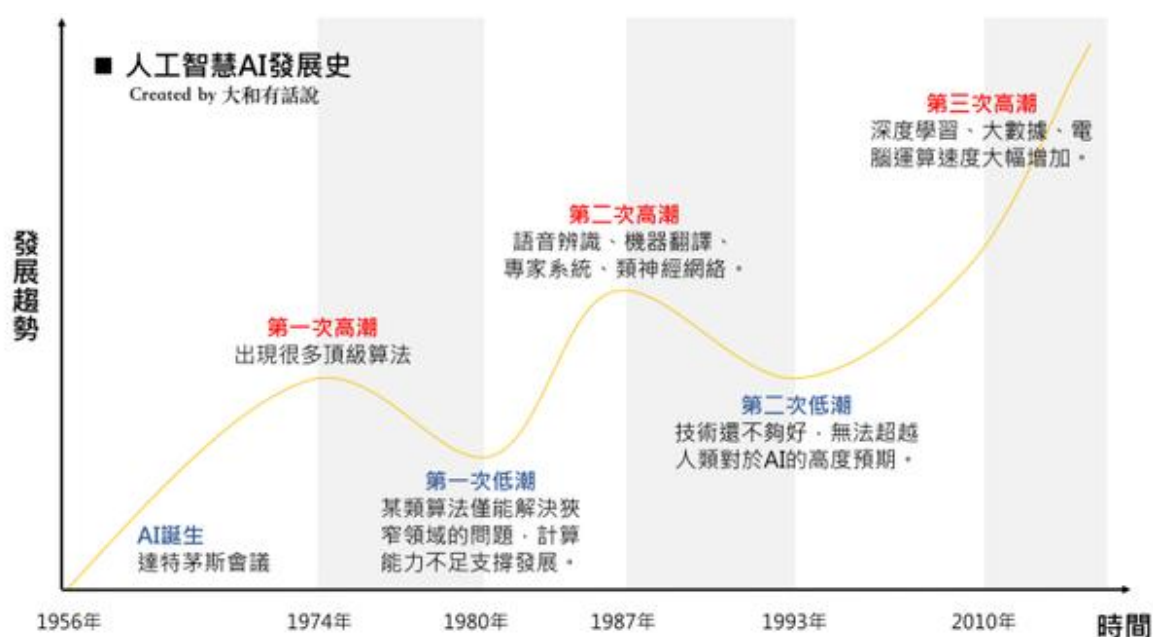
1. AI就是让人觉得不可思议的计算机程序

2. AI就是与人类思考方式相似的计算机程序
3. AI就是与人类行为相似的计算机程序
4. AI就是会学习的计算机程序
5. AI就是根据对环境的感知，做出合理的行动，并获得最大收益的计算机程序

实际上，对于AI的界定，目前在学术界都还没有统一一致的认识。其实想想也不难理解，因为界定‘Artificial Intelligence’（AI），‘Artificial’是比较好界定的，但是‘Intelligence’却很难界定。事实上我们现在对于什么是人类的Intelligence都缺乏一个严格的界定。

所以与其泛泛的纠结于AI的范畴界定，不如从AI的发展历史视角来看看，究竟我们今天所谈论的AI的历史内涵是什么。

实际上，AI并非是什么新鲜事物。历史上至今发生过3次AI热潮，我们正在经历的这一次，已经是第三次热潮了。



AI起源与第一次AI热潮

第一次AI的热潮肇始于1950年10月，图灵发表了一篇名为《计算机机械和智能》（Computing Machinery and Intelligence）的论文，试图探讨什么是AI。在文中，图灵提出了一个有趣的实验，即大名鼎鼎的“图灵测试”：如果一个人（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）、一个是机器（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，则此机器A通过图灵测试。[2]

图灵测试回避了有关人工智能比较难以回答和空泛的问题，比如“机器是否会思考”（需要界定什么是思考）或者是“机器是否有自我意识”（需要界定什么是意识），而是采用了一种形式化的定义。

这样的好处，就是使得在操作层面界定AI有了一个标准。

这个测试，也被称为“模仿游戏”，即在机器试图模仿人类与评判者对话的“模仿游戏”中，有思考能力的电子计算机是否可以做的更好。

1955年在洛杉矶召开的美国西部计算机联合大会（Western Joint Computer Conference）上展开了学习机讨论会（Session on Learning Machine），讨论会的参加者中有2个人参加了第二年著名的达特茅斯会议，分别是Oliver Selfridge和Alan Newell。Selfridge发表了一篇关于模式识别的文章，而Newell则探讨了计算机能否下棋。他们分别代表两派的观点。讨论会的主持人是神经网络的鼻祖之一Pitts，他最后总结时说“（一派）企图模拟神经系统，而Newell企图模拟心智……但殊途同归”。这句话也预示了随后几十年人工智能关于“结构与功能”两条路线的斗争。[3]

到了1956年夏天，具有里程碑意义的达特茅斯夏季人工智能研究会召开。在会议上人工智能的概念被首次提出。会议上，J. McCarthy等4人发起了一个达特茅斯夏季人工智能研究计划（Dartmouth Summer Research Project on Artificial Intelligence），在他们的提案《A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence》中，他们提出：“这些研究是基于这样的推测进行的，即，学习的各个方面，或智能的任何其他特征都能原则上精确的描述以使得机器来对其进行模拟。”（The study is to proceed on the basis of the conjecture that

every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.) [4]

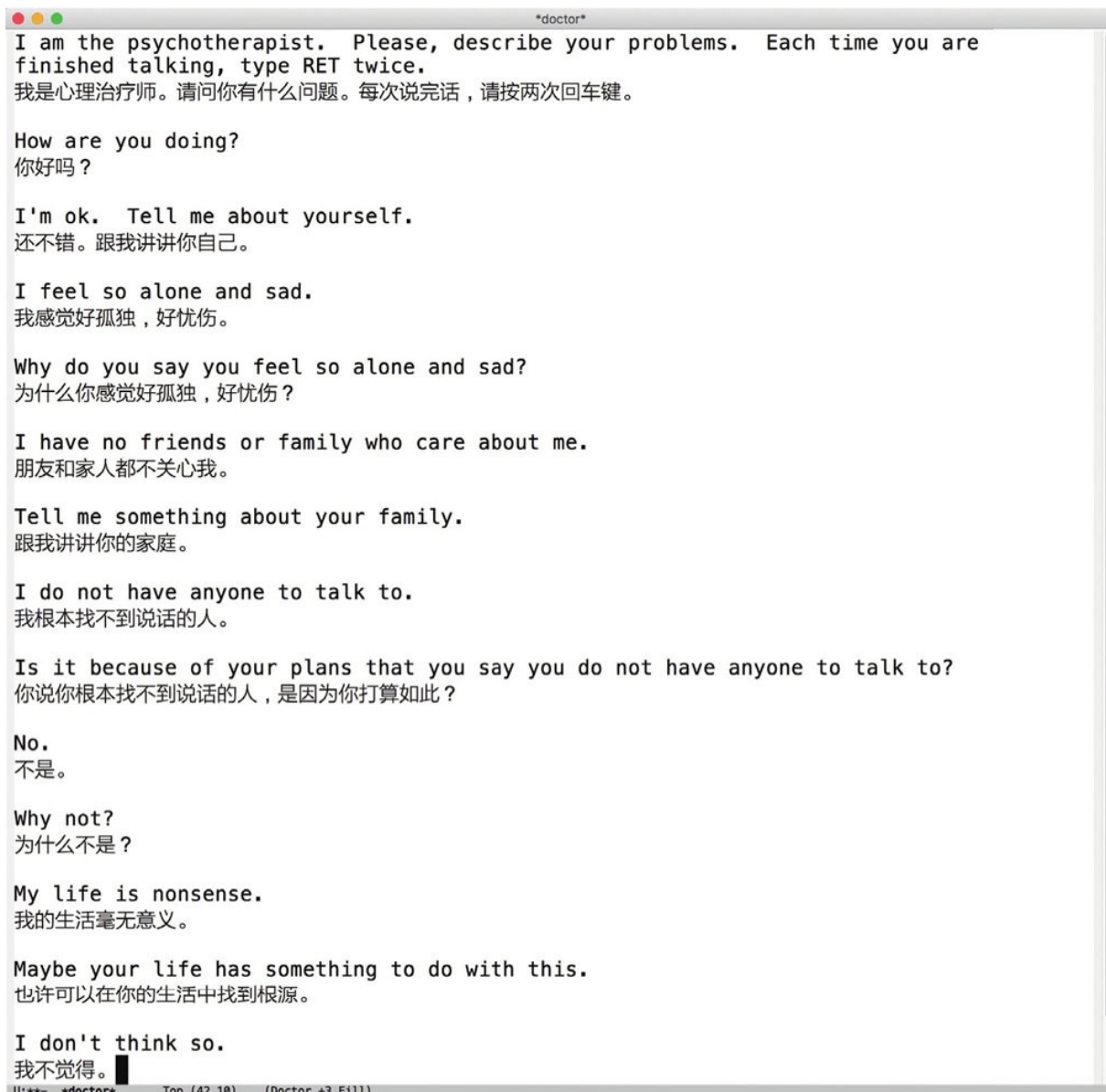
可以看出，这些AI的先驱们对于AI的定义主要是基于机器对于人的智能的各个方面的模拟。在同一篇文章中，他们提出了AI问题的几个可以研究的方向：[4]

1. 自动计算机/器 (Automatic Computers) ——他们提出如果一种机器可以完成一种工作，那么一种自动计算机就可以模拟这种机器。有意思的是他们认为虽然当时的硬件条件难以实现模拟人脑的高阶功能，但主要的障碍却是如何有效利用现有的硬件。
2. 如何编程使得计算机使用一种语言 (How Can a Computer be Programmed to Use a Language)
3. 神经网络 (Neuron Nets) ——不是深度学习中神经网络的概念，而是研究理论上的神经元如何排布已形成概念，是一种理论研究。
4. 计算规模理论 (Theory of the Size of a Calculation) ——这是研究提高计算效率的问题。
5. 自我提升/演进 (Self-Improvement) ——他们认为只有能够进行自我提升/演进的机器才能被称为真正的智能机器 (Truely intelligent machine) 。
6. 抽象 (Abstractions) ——目标是对抽象进行分类，并且尝试描述机器从感知和其他数据中形成抽象的方法。
7. 随机性与创造性 (Randomness and Creativity) ——他们认为创造性思维来自于有序思维中的可控的随机性 (controlled randomness in otherwise orderly thinking) 。

达特茅斯会议之后不久，1956年9月IRE（即后来的IEEE）在MIT召开信息论年会，Newell和Herbert Simon发布了“逻辑理论家”程序，并发表了一篇文章“Logic Theory Machine”。“逻辑理论家”程序可以证明《数学原理》中命题逻辑部分的一个很大的子集，**该程序是第一个可工作的人工智能程序**。因此，这篇报告也成为了人工智能历史上最重要的文章之一。[3]除此之外，在1956年的IRE信息论年会上还有诸多影响深远的文章发表。

1966年，MIT教授Joseph Weizenbaum发明了一个可以和人对话的小程序，叫ELIZA，名字来自萧伯纳的喜剧《卖花女》，主角的名字就叫Eliza Doolittle。ELIZA小程序被设计成可以通过谈话帮助病人完成心理恢复的心理治疗师。当时的人都被惊呆了，ELIZA可以和病人一聊就是几十分钟，有的病人还特别愿意与ELIZA聊天。

下图展示了一个Emacs中现代版本的ELIZA：



```
I am the psychotherapist. Please, describe your problems. Each time you are
finished talking, type RET twice.
我是心理治疗师。请问你有什么问题。每次说完话，请按两次回车键。

How are you doing?
你好吗？

I'm ok. Tell me about yourself.
还不错。跟我讲讲你自己。

I feel so alone and sad.
我感觉好孤独，好忧伤。

Why do you say you feel so alone and sad?
为什么你感觉好孤独，好忧伤？

I have no friends or family who care about me.
朋友和家人都不关心我。

Tell me something about your family.
跟我讲讲你的家庭。

I do not have anyone to talk to.
我根本找不到说话的人。

Is it because of your plans that you say you do not have anyone to talk to?
你说你根本找不到说话的人，是因为你打算如此？

No.
不是。

Why not?
为什么不是？

My life is nonsense.
我的生活毫无意义。

Maybe your life has something to do with this.
也许可以在你的生活中找到根源。

I don't think so.
我不觉得。
U:*** *doctor* Top (42,10) (Doctor +3 Fill)
```

从直观感受上，ELIZA感觉已经非常像一个真实的人了。

然而ELIZA的程序逻辑却非常简单，就是在一个有限的话题库里，用关键字映射的方式，根据病人的问话，找到自己的回答。比如病人说“你好”，ELIZA就会说“我很

好，跟我说说你的情况。”此外，ELIZA还会用引导性的句子，比如“为什么？”，“请详细解释一下”之类引导性的句子，来使得整个对话持续下去。此外，ELIZA还使用一个小技巧，使用人称替换和句式替换来重复用户的句子。比如，用户说“我感到孤独和难过”，ELIZA会说“为什么你感到孤独和难过？”这样，虽然并不理解用户说了什么，但是ELIZA通过这些小技巧“装作”自己可以理解自然语言的样子。

ELIZA可以看作是人们第一次实现聊天机器人（Chatbot）的尝试。[1]

各个方向的进展使得在20世纪50年代到60年代，人们对人工智能普遍持过分乐观的态度。

1957年，达特茅斯会议的发起人之一Herbert Simon预言10年内计算机就会在国际象棋上击败人类。到了10年后，1968年达特茅斯的另一位发起人J. McCarthy又和国际象棋大师David Levy打赌说10年内程序下棋会战胜Levy，结果输掉了2000美元。事实上，计算机在国际象棋上击败人类要等到30年后，1997年IBM深蓝击败卡斯波罗夫了。

达特茅斯会议的另一位发起人明斯基在1968年库布里克的电影2001: A Space Odyssey的新闻发布会上大胆预言30年内机器智能将赶上人类。他在1989年又预言20年可以解决自然语言处理问题。

过分乐观却迟迟没有实质进展，质疑的声音开始越来越多。在1965年伯克利大学的欧陆派哲学家Hubert Dreyfus就发表了一篇文章《炼金术与人工智能》，后来演变成了著名的（或者说“臭名昭著”的）《计算机不能干什么》一书，其批判的对象就是人工智能。

到了70年代，AI进入了瓶颈期。当时的主要问题就是当时计算机的处理速度和内存不足以解决许多实际的AI问题。例如，在自然语言处理方向上，内存只能容纳20个单词的词汇表，只能应付表演。另外，学界在理论上也证明，AI有关的很多问题都与时间指数相关，这意味着复杂问题的解决在当时的计算能力下几乎需要无限长的时间。[5]

1973年，著名的莱特希尔报告（Lighthill report）出炉。Lighthill report是1973年James Lighthill发表在Artificial Intelligence: a paper symposium上的论文“Artificial Intelligence : A General Survey”的简称。该报告的主要目的是为英国科学研究委员会

评估AI领域的学术研究。报告的结论是：“AI领域的任何一部分都没能产出人们当初承诺的有主要影响力的进步。”该报告出炉后，英国政府停止了除三所大学之外的全部AI相关研究的资助。各国政府也纷纷效仿，AI进入了第一次寒冬期（AI Winter）。

日本的野心与第二次AI热潮

第二次AI热潮是20世纪80年代到90年代。

当时突破性的进展是多方面的。在商业应用上，有知识工程和专家系统（expert system）；在学界，以反向传播算法为代表的多层神经网络被研发成功（这是第三次AI热潮的基础）；政府方面，各国政府也纷纷推出政府主导的AI新计划。日本推出了号称投资8.5亿美元的五代机计划，英国推出了耗资3.5亿英镑的Alvey计划，美国有DARPA计划，其对AI的投资1984到1988年翻了3倍。

专家系统可以看作是一类具有专门知识和经验的计算机智能程序系统，一般采用人工智能中的知识表示和知识推理技术来模拟通常由领域专家才能解决的复杂问题。

[6]

一般来说，专家系统=知识库+推理机，因此专家系统也被称为基于知识的系统。一个专家系统必须具备三要素：

1. 领域专家级知识
2. 模拟专家思维
3. 达到专家级的水平

斯坦福大学的教授Edward Feigenbaum（获1994年图灵奖）和诺贝尔奖得主Joshua Lederberg等人合作，于1965年就开发出了世界上第一个专家系统程序DENDRAL，能够从几千种可能的分子结构中挑选出一个正确的。

1980年，卡内基梅隆大学DEC公司设计了名为XCON的专家系统。到了1985年，企业在专家系统中投入超过10亿美元，催生出了一批IT新公司。

不过专家系统也有其局限性，它智能针对专家预先考虑过的状况来准备对策，并没有自行学习的能力。

而各国政府的AI计划中，日本的5代机计划声势最大。

第五代电脑（第五世代コンピュータ）是日本通商产业省（现称经济产业省）于1982年的一个大型研发计划，其目的为开发一部划时代的电脑，利用大量平行计算，使它拥有超级电脑的运算效能和可用的人工智能能力。名字中的“第五代”用于指明它将会是具划时代意义的电脑。[8]

1978年，日本通产省（Ministry of International Trade and Industry, 简称MITI）委托日本计算机界大佬、时任东京大学计算机中心主任的元岡達(Tohru Moto-Oka)研究下一代计算机系统。

通产省（MITI）大概相当于中国原来的信息产业部，稍带科技部加中科院，2001年已重组为经济产业省，简称METI，大概相当于中国现在的工信部。就像所有其它东亚国家，日本政府对私营企业有重大影响力。通产省决定了日本的科技和工业政策。

当时的计算机工业按照电路工艺划分计算机的发展：第一代计算机是电子管，第二代是晶体管，第三代是集成电路，第四代是超大规模集成电路(VLSI)。通产省决定三年后开始建造第五代计算机，这是日本雄心勃勃的从制造大国到经济强国转型计划的一部分，用今天的话说：他们认为日本已经进入“无人区”，日本必须搞“双创”。首创第五代计算机，可以建立日本在全球信息产业的领导地位。

日本MITI对五代机的自信来自DRAM存储芯片的成功，70年代日本半导体工业在MITI的协同下，组织了业界协会，在很短时间内，DRAM研发全面赶超美国，日本在计算机硬件制造方面由此对美国构成威胁，MITI不满足于跟随美国，产生更大野心，要在整个IT领域设立自己的标准。

三年后的1981年，元岡達为首的委员会认真提交了一份长达89页的报告，他们认为第五代计算机不应再以硬件工艺为划分，更应看重体系结构和软件。这份报告题目就是《知识信息处理系统的挑战：第五代计算机系统初步报告》。[7]

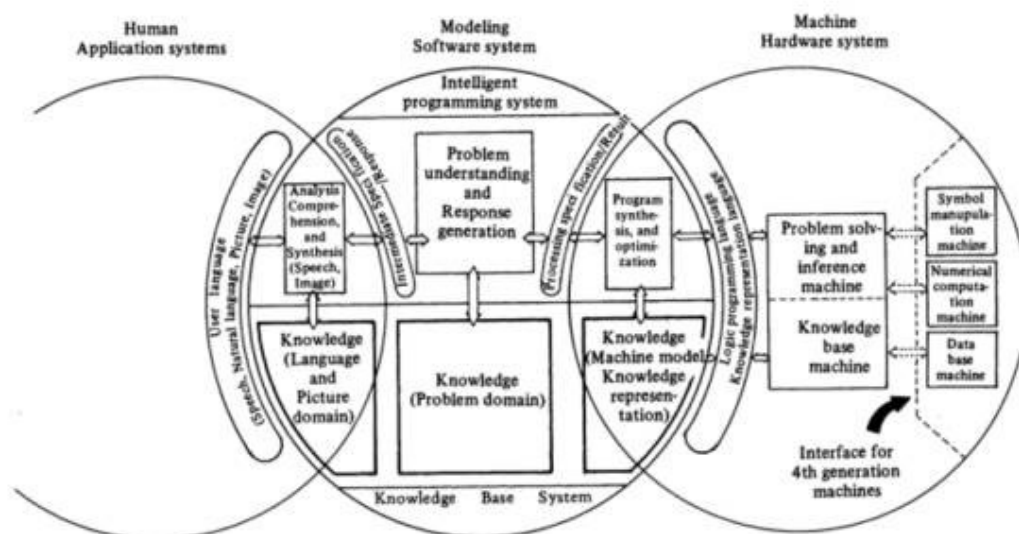


Fig. 1-1 Conceptual diagram of the fifth generation computer systems

报告提出了六种先进体系结构：1) 逻辑程序机，2) 函数机，3) 关系代数机，4) 抽象数据类型机，5) 数据流机，和6) 冯诺依曼机上的创新。

1981年在日本信息处理开发中心（JIPDEC）召开了第一届第五代计算机会议，对外公布元岡達委员会的报告。

前文已经提到，AI界一直有“结构与功能”之争。这两派后来分别演化为“语义”和“语法”之争。“语义”派重“知识”，代表是专家系统，而“语法”派重逻辑。

日本五代机的领军人物、“新一代计算机技术研究所”（Institute of New Generation Computer Technology,简称ICOT）所长渊一博（Kazuhiro Fuchi）在会议上发表的论文明确强调了逻辑程序和Prolog语言的重要性。

英国人沃伦（David Warren）1982年在AI Magazine上发表的论文分析了渊一博为什么选择了Prolog：在五代机的所有应用中，Prolog能覆盖的面最广。五代机的终极目标是知识信息处理，在当时的语境下特指专家系统和自然语言理解。

当时的自然语言理解还是以“语法”为主流，Prolog语言具有优势。另外Prolog也不是美国人发明的（当时美国人工智能研究者的标准语言是LISP），日本更有机会掌控并领先，这也更符合日本的初衷。

通产省为五代机制定的十年计划是4亿5千万美元，第一期头三年通产省独资4500万到5000万美元，后两期参与的公司会有一比一匹配资金。所以总投入预期会到8亿5千万美元。而日本历来工业界对政府项目的匹配都会超过政府出资，所以通产省的小算盘是整个项目有可能达到10亿美元投入。虽然照美国标准，这不是巨资，IBM 1982年一年的研发经费就是15亿美元，但在当时的日本已经是史无前例了。

然而五代机虽然野心勃勃，最后却成为了大杂烩。Carl Hewitt认为，大部分五代机的工作都是试着去用逻辑程序去解决其它手段早就解决的问题，而不是去解决其它手段不能解决或解决得不好的问题。他还认为逻辑编程和并行性天生就是一对不可调和的矛盾。

1992年日本的五代机计划正式宣布失败。英美的AI计划也没能取得预想的结果。

这股AI热潮到了80年代后期开始降温。与第一次AI热潮不同，第二次AI热潮并没有一个衰落的标志性事件，而是到了90年代，随着个人电脑和互联网产业的迅速兴起，AI逐步淡出视线，进入了第二次低潮期。

第三次AI热潮

2006年，深度学习的泰斗Geoffrey E. Hinton与Simon Osindero合作发表了一篇论文“A fast learning algorithm for deep belief nets”（《一种深度置信网络的快速学习算法》）。在业界很多人将此作为第三次AI热潮的开端。以此为标志，2006-2010年一系列关键论文相继发表，推动了此轮AI热潮的兴起。

不过深度学习本身却并非新鲜事物。在1943年通用计算机诞生之前，神经科学家Warren McCulloch和Walter Pitts就提出了一种假说，描述人类神经节沿着网状结构传递和处理信息的模型。这一假说一方面被神经科学家用于研究人类的感知原理；另一方面则被计算机科学家们借鉴，用于人工智能的相关研究。后者也被学术界称为人工神经网络。

20世纪40年代，唐纳德·赫布（Donald Hebb）尝试将人工神经网络用于机器学习，创建出早期的“赫布型学习”（Hebbian Learning）理论。1954年，计算机科学家韦斯利·A. 克拉克（Wesley A. Clark）在麻省理工学院尝试在计算机上实现赫布型学

习的基本模型。1958年，弗兰克·罗森布拉特（Frank Rosenblatt）提出了“感知机”（Perceptron）的概念，这是一个基于人工神经网络的两层计算结构，用于简单的模式识别。1965年，A. G. 伊瓦赫年科（Alexey Grigorevich Ivakhnenko）提出建立多层人工神经网络的设想，这种基于多层神经网络的机器学习模型后来被人们称为“深度学习”，伊瓦赫年科有时也被称为“深度学习之父”。[1]

1969年麻省理工学院的图灵奖得主、人工智能大师、人工神经网络的早期奠基人之一Marvin Minsky在这一年和Seymour Papert出版了《感知机》（Perceptrons）一书，书中讨论了当时人工神经网络难以解决的“异或难题”。

到了1975年，Minsky提出的“异或难题”才被学界解决。20世纪80年代人工神经网络的发展回到正轨，此后人工神经网络作为机器学习的一种算法而存在。

2006年的多篇重要论文的发表在基本理论方面取得了重大突破，深度学习开始进入高速发展的全盛期。

深度学习发挥作用有2大前提条件：强大的计算能力和高质量的大数据。这二者都在2010年以后逐渐成熟。

而此轮AI热潮的崛起，除了公众所熟知的AlphaGo之外，就是Google、Facebook以及百度等一大批顶尖科技公司纷纷将“人工智能优先”设定为公司的科技发展战略。

我们可以从Google Brain的发展中一窥此轮AI热潮的兴起脉络。

2011年，吴恩达（Andrew Ng）在Google内部建立的一个项目——Project Marvin（以AI先驱Marvin Minsky命名，该项目最初是属于Google X计划中的项目（该计划中还包含很多著名的项目，比如Google Glass）），用于研究“神经网络”取得了一些不错的成果。很快Google的传奇人物Jeff Dean很感兴趣，也加入了进来，他随后建议吴恩达邀请有神经科学背景的同事Greg Corrado加入。后来Google内部逐渐将该项目称为Google Brain。[9]

Dean 表示，Google Brain 成立后一年左右，开发具有一岁儿童智力的机器的实验取得了巨大的进展。谷歌的语音识别团队将其旧系统的一部分改为神经网络，并且效果得到很大提升，甚至取得了近 20 年中最好的成果。谷歌物体识别系统的能力也提高了一个数量级。

Google Brain成立的第二年，吴恩达离开加入百度，成为百度首席科学家，领导者1300人规模的AI团队。而深度学习的泰斗Geoffrey Hinton加入Google Brain团队。

几个月后，Hinton 和他的两个学生在 ImageNet 大型图像识别竞赛中取得了惊人的成果，让计算机不仅识别出猴子，而且区分蜘蛛猴和吼猴，以及各种各样不同品种的猫。

2012年6月，Google Brain使用了一个拥有16000个CPU的大规模计算机集群，让计算机用深度学习模型“看”了1000万段YouTube上的视频，然后计算机自己“学”会了如何从视频中识别一只猫。

2013年3月收购DNNResearch。

2014年1月，Google收购DeepMind团队，后来正是该团队开发了著名的AlphaGo。

2015年11月，Google Brain团队开发的TensorFlow正式开源发布，目前已经是深度学习领域主流框架之一。

2016年3月，AlphaGo以4:1击败李世乭。

2016年9月，Google推出神经机器翻译（GNMT），不仅提升了机器翻译的质量，还可以直接实现多语言之间的zero-shot翻译，即通过中间语言实现两种没有训练过直接对应翻译的语言的互翻。

2017年5月，强化版AlphaGo战胜柯洁。

.....

事实上，Google Brain的发展史虽然具有标志性，但还难以概括此轮AI热潮的整体画面。从2006年的短短十年，深度学习就完成了从理论突破到商业应用的跨越。而目前深度学习的应用范围之广，很难在简短的篇幅内介绍清楚。

第三次AI热潮与前两次最大的不同在于，第一次AI热潮还停留在学术研究的范畴，而第二次AI热潮的专家系统后来被证明应用范围狭窄，距离实用还有相当距离，而此次AI热潮已经有了大量可实现商业应用的案例。

所以在从时间维度了解了AI的过去之后，下一篇，我们就从空间维度来梳理本轮AI热潮的应用方向。

刘晟西

2018年7月17日

于北京

参考文献

[1] 《人工智能》李开复 王咏刚

[2] [维基百科：图灵测试](#)

[3] [《人工智能的缘起-中国计算机学会信息网》](#)：《中国计算机学会通讯》2016年第3期，《专栏》

[4] [A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE](#) J. McCarthy, Dartmouth College, M. L. Minsky, Harvard University, N. Rochester, I.B.M. Corporation, C.E. Shannon, Bell Telephone Laboratories, August 31, 1955

[5] [AI的酷暑与寒冬 | AI传奇系列之二](#)

[6] [维基百科：专家系统](#)

[7] [野心勃勃的日本第五代计算机，如何一步步走向失败](#)

[8] [维基百科：第五代电脑](#)

[9] [《纽约时报》两万字长文，深度剖析谷歌大脑简史](#)