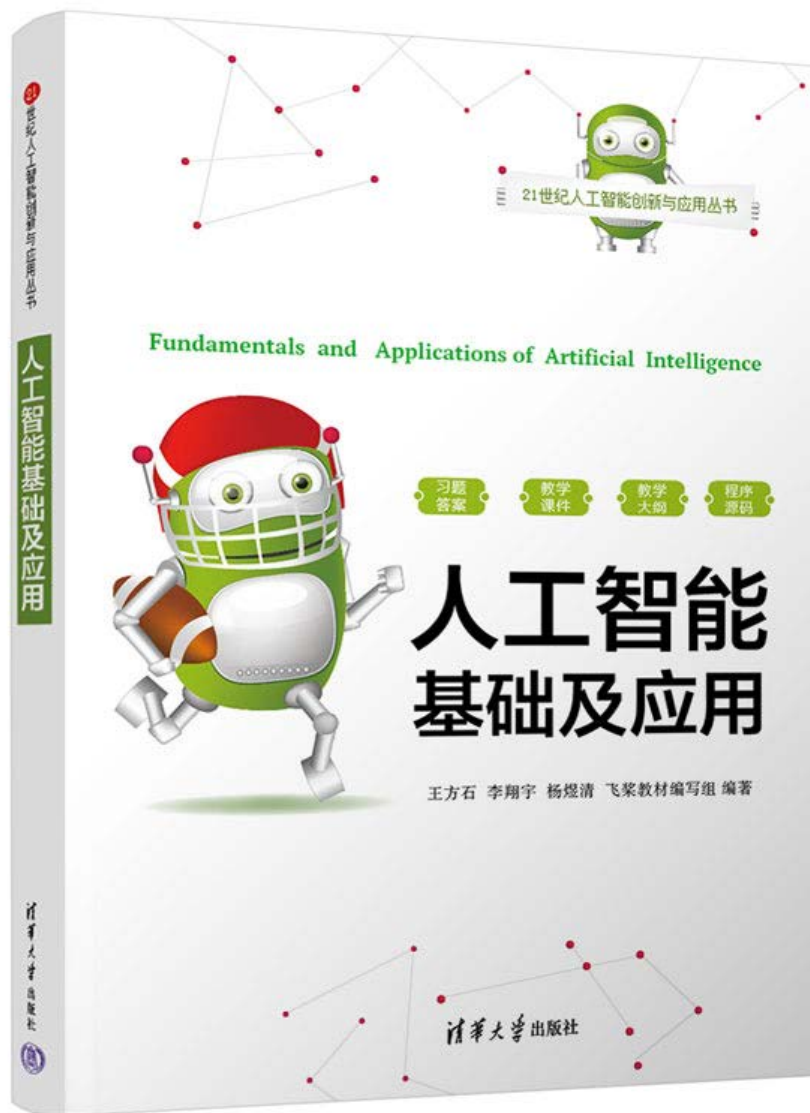


人工智能基础



北京交通大学 软件学院
王方石

Email: fshwang@bjtu.edu.cn

第1章 人工智能概述

- 1.1 人工智能的萌芽与诞生
- 1.2 人工智能的定义
- 1.3 人工智能的发展简史
- 1.4 人工智能的研究流派
- 1.5 人工智能研究的基本内容
- 1.6 人工智能的主要研究领域

本章学习目标

- ◆ 了解人工智能的萌芽与诞生、定义和发展简史。
- ◆ 了解人工智能研究的三个流派及研究方法。
- ◆ 了解人工智能研究的基本内容和主要研究领域。

1.1 人工智能的萌芽与诞生

1.1.1 人工智能的萌芽

- ◆1936年，图灵提出了一种理想计算机的数学模型，即著名的**图灵机**。
- ◆1940-1942年间，美国爱荷华州立大学的**阿塔纳索夫**(Atanasoff)教授和他的研究生**贝瑞**(Berry)装配了世界上第一台电子计算机，命名为阿塔纳索夫-贝瑞计算机 (Atanasoff-Berry Computer, **ABC**)，为人工智能的研究奠定了物质基础。人们熟知的1946年诞生的ENIAC并非世界上第一台计算机。

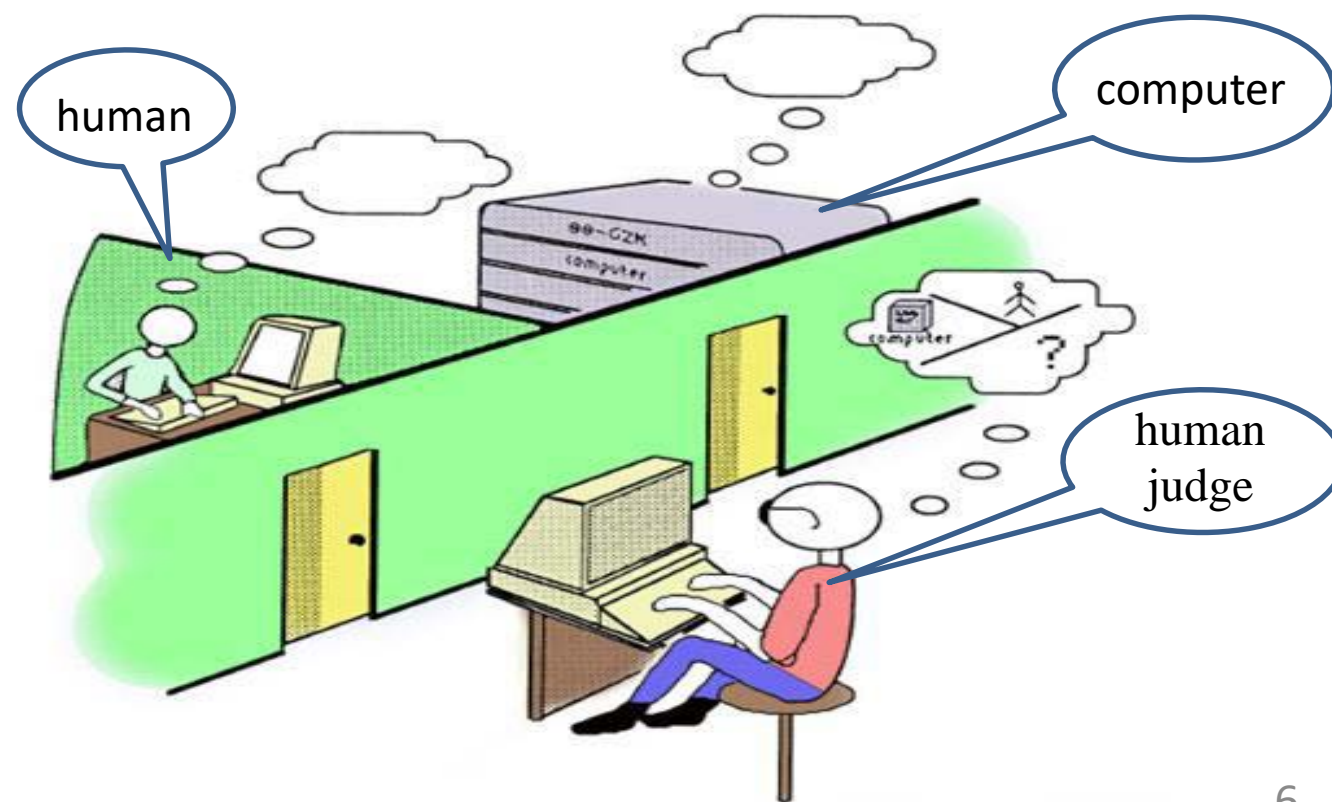
1.1.1 人工智能的萌芽

- ◆ 1943年，美国神经生理学家沃伦·麦卡洛奇(McCulloch)与数理逻辑学家沃尔特·皮茨(Pitts)提出了**人工神经元**的概念，建立了第一个**神经网络的数学模型(M-P模型)**，证明了本来是纯理论的图灵机可以由人工神经元构成，为后续人工神经网络的研究奠定了基础。
- ◆ 1949年，加拿大心理学家赫布(D.Hebb)提出了一种学习规则，用来对神经元之间的**连接强度**进行更新。
- ◆ 1950年，图灵发表名为“计算机器与智能”的文章，论述并提出了著名的“**图灵测试**”，以此作为判断机器是否具有智能的标准，作为一个判断的**充分条件**。

1.1.2 图灵测试与中文屋实验

(1) 图灵测试

- ◆ 分别让人和机器位于两个房间，看不见彼此。
- ◆ 人类书面提问后，若无法分辨这些书面回答究竟是来自于人还是一台计算机，则认为计算机通过了该测试。
- ◆ 图灵预言，到**2000年**，一个普通的人类测试者在5分钟的提问后，能正确判断对方是机器还是人的**概率可能只有70%**。



1.1.2 图灵测试与中文屋实验

(1) 图灵测试

- ◆ 在相当长一段时间内，**“图灵测试”被认为是人工智能水平的标准测试模型。**
- ◆ 从1991年起，每年举办图灵测试，结果都是：程序还不能欺骗经验丰富的仲裁者。
- ◆ **直到2014年6月**，为纪念图灵逝世60周年，聊天机器人尤金•古斯特曼成功地使参加该活动的**33%的人类测试者**认为它是一个**13岁的男孩**，因此组织者认为古斯特曼已经通过了图灵测试，成为历史上**第一个通过图灵测试的人工智能软件。**

Alan Turing（阿伦·图灵）

1912年6月23日—1954年6月7日

- ◆ 计算机科学理论的创始人
- ◆ 英国数学家、逻辑学家、计算机科学家和密码学家
- ◆ 计算机科学之父，人工智能之父

除了**图灵测试**，图灵还提出了著名的**图灵机模型**，它是冯诺依曼计算机的理论模型，为现代计算机的逻辑工作方式奠定了基础。





- ◆ 2015年上映的电影《模仿游戏》，改编自《阿伦·图灵传》。
- ◆ 这部电影展现了图灵作为密码学家在二战中的杰出贡献，主要讲述了他协助盟军破译德国密码系统，从而扭转二战战局的经历。

Turing Award（图灵奖）

- ◆ 1966年，为纪念图灵的杰出贡献，ACM（国际计算机协会）设立图灵奖，专门奖励那些对计算机科学作出重要贡献的个人。
- ◆ 图灵奖对获奖者的要求极高，评奖程序极严，一般每年只奖励一名计算机科学家，只有极少数年度有两名或3名在同一方向上做出贡献的科学家同时获奖。
- ◆ 每年3月下旬颁发上一年的图灵奖。
- ◆ “图灵奖”的奖金数额从20万-25万-100万美元，它是计算机界最负盛名的奖项，有“**计算机界诺贝尔奖**”之称。
- ◆ 图灵奖的奖金在设奖初期为20万美元，1989年起增到25万美元。奖金通常由计算机界的一些大企业提供（通过与ACM签订协议），目前图灵奖奖金由Google公司赞助，为100万美元。
- ◆ 从1966年至**2023年3月**，共计有**76名**科学家获此殊荣，其中美国学者最多。
- ◆ 目前，获此殊荣的**华人仅有1位**，他是**2000年**图灵奖得主**姚期智**（台大毕业，现在清华大学交叉信息研究院、香港中文大学）。

(2) “中文屋” 思想实验 (Chinese Room)

- ◆现在仍有许多人将**图灵测试**作为**衡量机器智能的准则**，但也有许多人认为：即使通过了图灵测试，也不能认定机器就有智能，因为没有涉及思维过程。
- ◆针对图灵测试，美国哲学家约翰·塞尔勒在1980年设计了“**中文屋**”**思想实验**，也被称为塞尔勒的中文屋论证。
- ◆这一观点发表在《行为与大脑科学》杂志上的“智力、头脑与规划”论文中。
- ◆试图揭示计算机绝不能描述为有“智力”或“知性”，不管它多么智能。



John R. Searle

约翰·塞尔勒

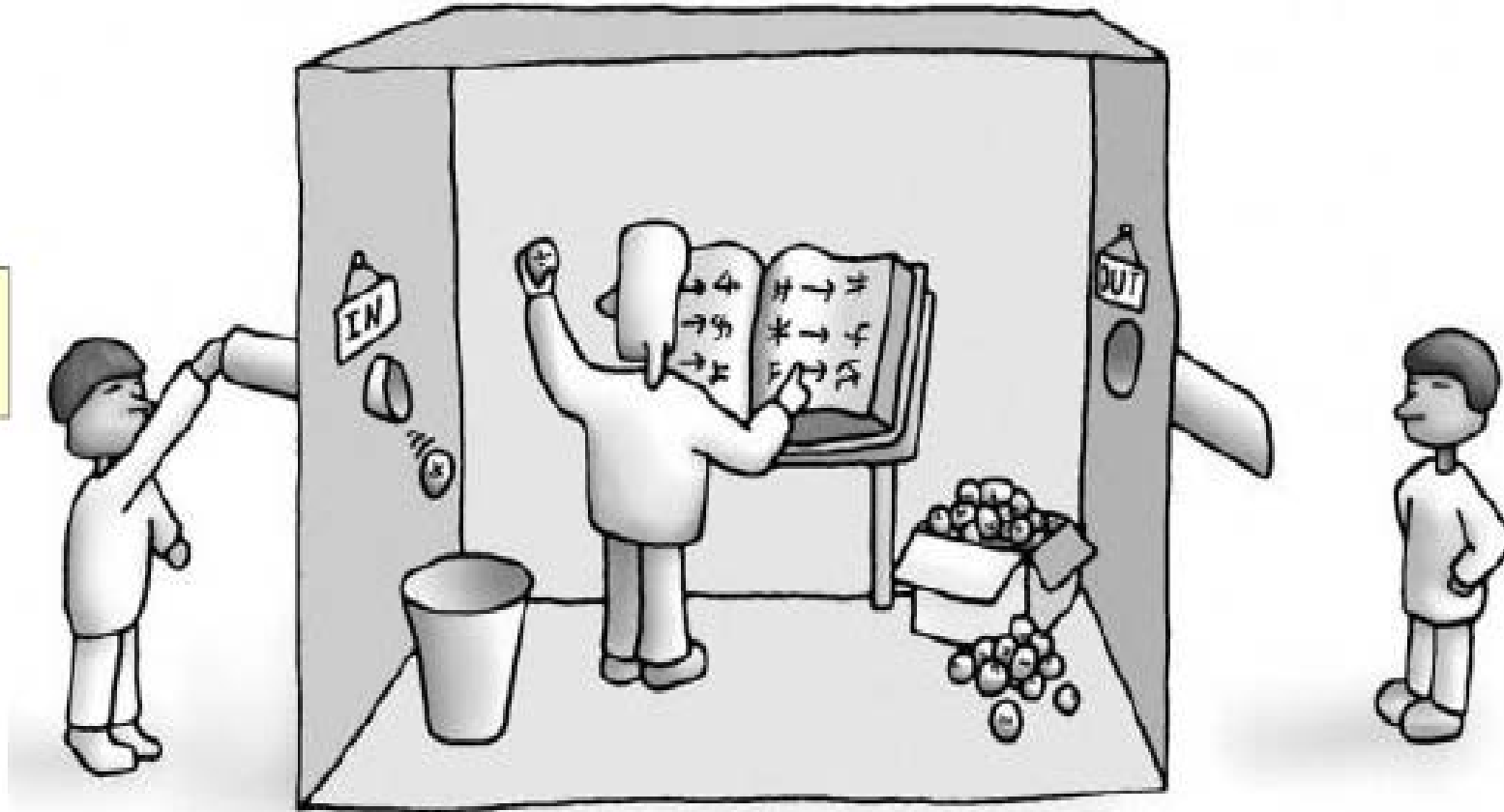
Slusser Prof. of
Philosophy, UC Berkeley.

“斯拉瑟” 哲学教授

塞尔勒的中文屋子Chinese Room

我抢到了红包！

I got the red envelope.



恭喜发财！

May you be happy and prosperous!

不管是谁、无论如何，在房间里的一定是个聪明的精通中文的人！
通过了图灵测试，就具有智能了吗？

1.1.3 人工智能的诞生

- ◆ 1956年的“达特茅斯夏季人工智能研究计划”会议，是被广泛公认的人工智能诞生地。
- ◆ 达特茅斯学院（Dartmouth College）位于美国新罕布什尔州 汉诺威市（New Hampshire , Hanover）。
- ◆ 洛克菲勒基金会资助每个代表1200美元以及外地代表往返车票。
- ◆ 达特茅斯会议历时长达两个多月，会上经麦卡锡提议，正式采用“人工智能”这一术语，标志着人工智能学科正式诞生。麦卡锡因而被称为人工智能之父。

10名与会者

◆ 4 名发起人：

- J. McCarthy：达特茅斯学院数学助理教授
- M. L. Minsky：哈佛大学数学与神经学初级研究员
- N. Rochester：IBM信息研究经理
- C.E. Shannon：贝尔电话实验室信息部数学研究员，信息论的创始人

◆ 6名参会者：A. Newell and Herbert. A. Simon：卡内基.梅隆大学

T. More, A. Samuel：IBM公司

O. Selfridge, R. Solomonff：麻省理工学院

◆ 这些青年学者的研究专业包括数学、心理学、神经生理学、信息论和计算机科学，分别从不同的角度共同探讨人工智能的可能性。

◆ A. Samuel 于1959年编写了第一个电脑跳棋程序。

达特茅斯研讨会的三个亮点

- ◆ 明斯基提出用计算机模拟神经元及其连接，通过“人工神经网络”模拟智能，构建了第一个神经网络模拟器SNARC（Stochastic Neural Analog Reinforcement Calculator）；
- ◆ 麦卡锡带来的是“状态空间搜索法”；
- ◆ 纽厄尔和西蒙展示了“逻辑理论家”（Logic Theorist, LT）程序，

达特茅斯会议10位与会者

1956 Dartmouth Conference: The Founding Fathers of AI



John MacCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



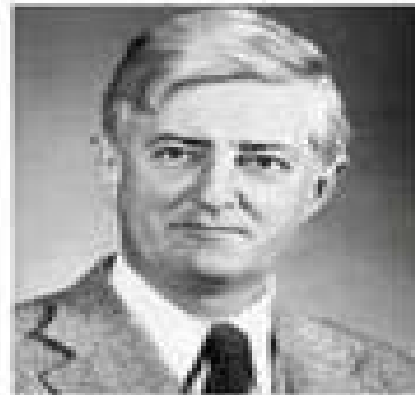
Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

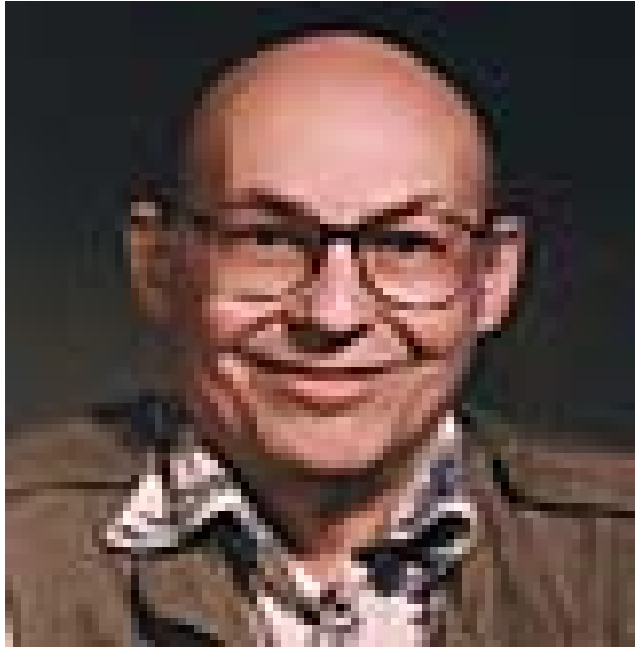
Founding fathers of AI. Courtesy of scienceabc.com

达特茅斯会议50年后，5位与会者合影



2006年重聚，左起：**摩尔**、麦卡锡、明斯基、**塞弗里奇**、所罗蒙夫

图灵奖获得者



Marvin Lee Minsky
1969



John McCarthy
1971



Herbert A. Simon
1975



Allen Newell
1975

人工智能之父

AI领域图灵奖获得者

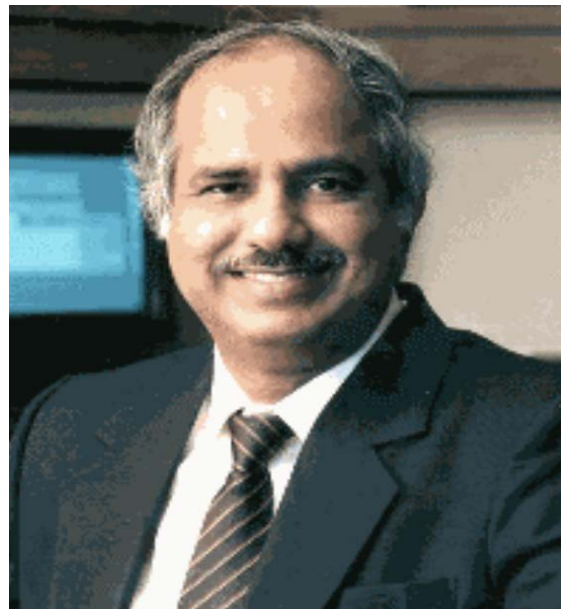
- ◆ 1985年, **Richard M. Karp**, 提出了分支限界法, 解决了旅行商问题 / 汉密尔顿回路,
成就: 算法理论, 尤其是NP-完全性理论。
- ◆ 1994年, **Edward Feigenbaum + Raj Reddy** (麦卡锡的学生、李开复的导师),
成就: 大规模人工智能系统
- ◆ 2011年, **Judea Pearl**, 成就: 提出概率和因果性推理演算法, 彻底改变了人工智能最初基于规则和逻辑的方向。



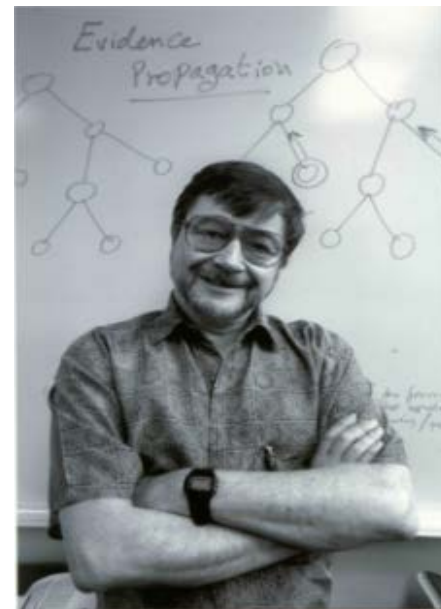
1985



1994



1994



2011

2018年图灵奖获得者：DNN的三大奠基人、巨头

- ◆ 2019年3月27日，ACM宣布，深度学习的三位创造者Yoshua Bengio, Yann LeCun, 以及Geoffrey Hinton 获得了2018年的图灵奖。
- ◆ 奖励他们提出了深度学习的基本概念，在实验中发现了惊人的结果，也在工程领域做出了重要突破，帮助深度神经网络获得实际应用。

2019-3-27 图灵奖获得者：DNN的三剑客、三巨头



Geoffrey Hinton (1947-12-6)

出生于英国

人工智能教父

神经网络之父

深度学习鼻祖

谷歌大脑团队神级人物

多伦多大学向量学院首席科学顾问

美国人工智能协会AAAI院士

Yann LeCun (60.7.8生)

杨立昆 / 燕乐存

出生于法国，居里大学博士

卷积神经网络之父

多伦多大学Geoffrey Hinton

实验室的博士后

美国工程院院士、**Facebook**

人工智能研究院院长、

纽约大学Sliver教授

Yoshua Bengio (64年生)

出生于巴黎

在加拿大蒙特利尔长大，

任教于蒙特利尔大学

Bengio坚持与他的神经网络为伍

最后一个学术纯粹主义者

1.2 人工智能的定义

- ◆ 在计算机科学中，**人工智能**（AI），有时也称为**机器智能**，是用人工的方法在机器（计算机）上实现的智能；或者说是人们使机器具有类似于人的智能。
- ◆ 麦卡锡等提出：对于人工智能**预期目标的设想**是“精确地描述学习的每个方面或智能的任何其他特征，从而可以制造出一个机器来模拟学习或智能”（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）。该预期目标曾被认为是人工智能的定义，对人工智能的发展起到了重要的作用。

三个最常见的人工智能定义

- ◆ 明斯基提出：“人工智能是一门科学，是使机器做那些人需要通过智能来做的事情”。
- ◆ 尼尔森提出：“人工智能是关于知识的科学——怎样表示知识以及如何获得知识并使用知识的科学”，所谓“知识的科学”就是研究知识表示、知识获取和知识运用的科学。
- ◆ 温斯顿教授提出：人工智能就是研究如何使计算机去做过去只有人才能够做的智能工作”。

- ◆ 人工智能学科就是研究人类智能活动的规律，研究如何应用计算机的软硬件制造出智能的机器或系统，使之具有智能的行为，来模拟人类的某些智能活动。
- ◆ 人工智能是机器或软件所展现的智能，**是对人的意识、思维过程的模拟。**
- ◆ 虽然人工智能不是人的智能，但能像人那样思考，甚至在有些方面已经超过了人的智能。
- ◆ 作为计算机科学的一个分支，人工智能是一门融合了自然科学、社会科学和技术科学的交叉学科，它涉及的学科包括认知科学、神经生理学、哲学、**数学、心理学**、控制论、计算机科学、信息论、仿生学与社会结构学等。

谭铁牛院士的描述

◆人工智能是研究开发能够模拟、延伸和扩展人类智能的理论、方法、技术及应用系统的一门新的技术科学。

◆研究目的是促使智能机器：

- 会听（语音识别、机器翻译等）、
- 会看（图像识别、文字识别等）、
- 会说（语音合成、人机对话等）、
- 会思考（人机对弈、定理证明等）、
- 会学习（机器学习、知识表示等）、
- 会行动（机器人、自动驾驶汽车等）。

人工智能分类

人工智能大致分为两大类：**弱人工智能**和**强人工智能**。

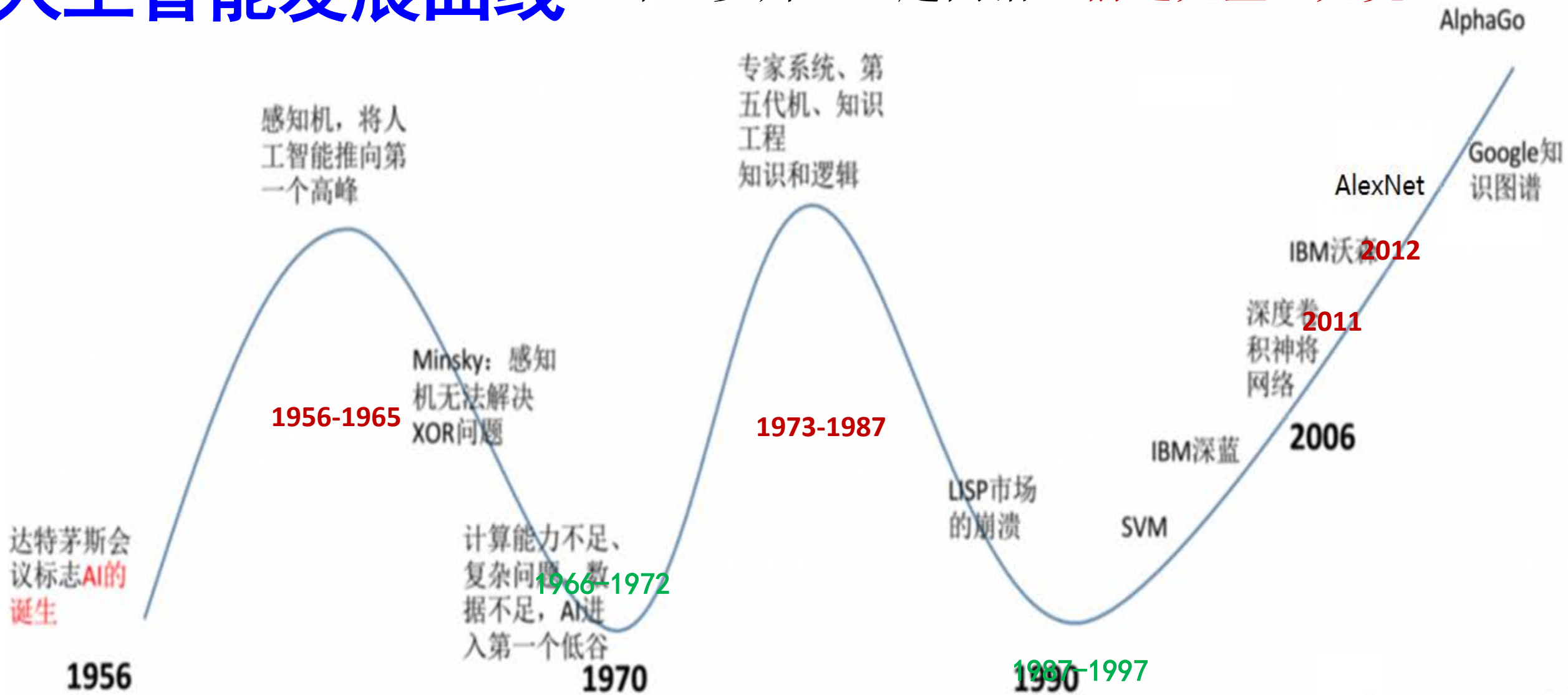
- ◆ **弱人工智能**(weak artificial intelligence)是能够完成某一特定领域中某种特定具体任务的人工智能。
- ◆ **强人工智能**(strong artificial intelligence)也称为**通用人工智能**，是具备与人类同等智慧，或超越人类的人工智能，能表现正常人类所具有的所有智能行为。
- ◆ 人工智能在某些方面很容易超越人类，例如数学计算、博弈、知识记忆（强记博闻）等，但它只能解决一两个特定问题，无法像人类一样能解决各种各样的问题。目前人工智能的研究及应用均聚焦于这类弱人工智能（大语言模型/SORA）。

1.3 人工智能的发展简史

- (1) 1930s-1956: 萌芽期
- (2) 1956年: 诞生
- (3) 1950年代中期—60年代中期: 黄金期 (通用方法时代)
- (4) 1960年代后期—70年代初: 第一个寒冬期
- (5) 1970年代中期—80年代后期: 繁荣期 (专家系统与知识工程时代)
- (6) 1980年代末— 1990年代中期: 第二个寒冬期
- (7) 1997-2011: 复苏期
- (8) 2012至今: 蓬勃发展期

人工智能发展曲线

命运多舛！三起两落！前途无量？无亮？



NIPS 2017, Judea Pearl 的《机器学习的理论障碍》 第三个寒冬是否就要到来？

1950年代中期-1960年代中期：黄金期

- (1) 在定理证明方面
- (2) 在机器学习方面
- (3) 在模式识别方面
- (4) 在计算机博弈方面
- (5) 在问题求解方面
- (6) 在人工智能语言方面
- (7) 在专家系统方面

早期的研究者对这些人工智能项目寄予了巨大的热情和期望。

1960年代后期-1970年代初：第一个寒冬期（1）

人们发现：采用统计模型的机器翻译很容易产生歧义。

(1) **The spirit is willing but the flesh is week.** （心有余而力不足）

The vodka is strong but meat is rotten. （伏特加酒虽然很浓，但肉是腐烂的）
出错的原因是一词多义：Spirit: 1) 精神 2) liquor (烈性酒)

(2) **Out of sight, out of mind.** “眼不见，心不烦”

Blind and insane. “又瞎又疯”

(3) **Time flies like an arrow.**

“光阴似箭” vs. “苍蝇喜欢箭”

◆错误在于仅字面翻译，并非理解了。

◆结论：必须理解才能翻译，而理解需要知识。

1960年代后期-1970年代初：第一个寒冬期（2）

- ◆ 1957年，美国心理学家Rosenblatt（罗森布拉特）提出著名的**感知机**。
- ◆ **感知器的局限性**：1969年明斯基与帕普特（Papert）出版了专著《感知机：计算几何导论》，指出：感知器本质上是一种线性模型，只能处理线性分类问题，**无法解决线性不可分问题**，就连最简单的XOR（亦或）问题都无法正确解决。**计算量过大**，而且感知器上的学习算法将失效。
- ◆ 1970年，**连接主义遭到遗弃**，人工神经网络研究成为小众。
- ◆ 1974年，哈佛的保罗·韦伯斯(Paul Werbos) 在其博士论文《超越回归：行为科学中预测与分析的新工具》中证明了“利用**误差反向传播**（BP）算法训练多层人工神经网络，可以解决异或问题”，但当时**并未引起重视**。
- ◆ 1971-1974年，英美均削减了人工智能的学术研究经费。

感知机的局限性

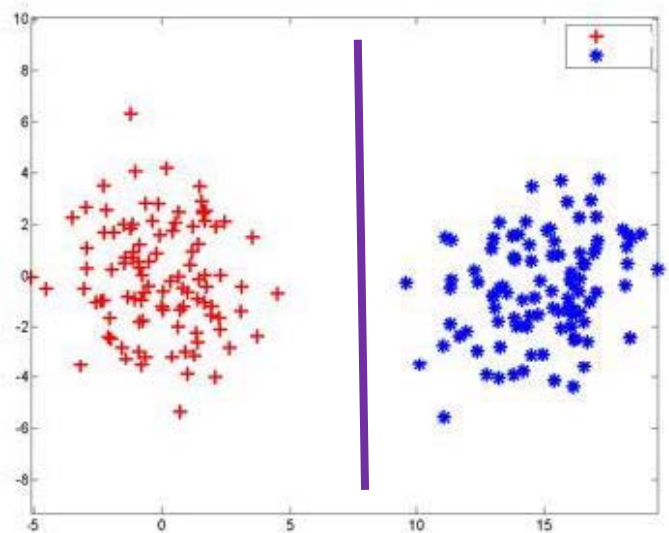
◆ 感知机(单层神经网络)无法解决异或问题

异或：简称XOR，用数学符号 \oplus 表示，在计算机中用 \wedge 表示。

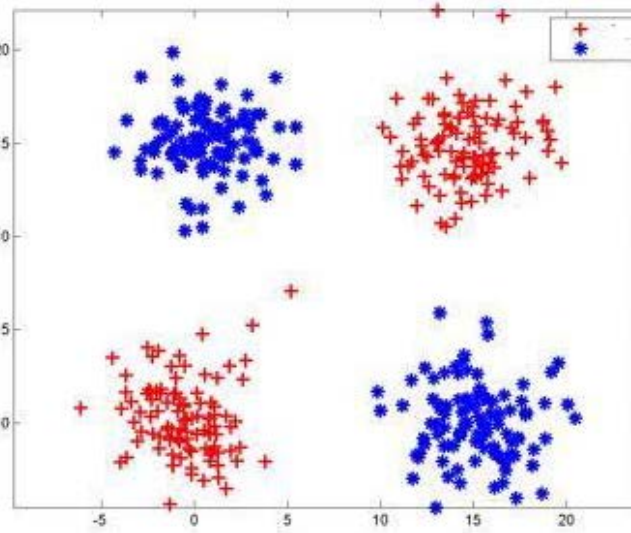
异或运算指二进制中： $0 \oplus 0 = 0$ ， $0 \oplus 1 = 1$ ， $1 \oplus 0 = 1$ ， $1 \oplus 1 = 0$

即：两个值相同得0，不同得1。

◆ 感知机无法解决线性不可分问题



(a) 线性可分



(b) 线性不可分

如右图，无论怎样在平面上画一条直线，都无法将蓝点和红点完全分开。

1970年代中期-1980年代后期: 繁荣期

20世纪70年中期，以**专家系统**为主的人工智能研究逐步进入了繁荣期。有代表性的**应用成功的专家系统**如下。

- ◆ 1965年，斯坦福大学成功研制了化工领域的DENDRAL系统；
- ◆ 1971年，麻省理工学院研制了数学领域的MYCSYMA系统；
- ◆ 1972年，特里·维诺格拉德（Terry Winograd）研制了自然语言理解系统SHRDLU；
- ◆
- ◆ 1975年，开发了医学诊断专家系统 MYCIN；
- ◆ 1976年，美国斯坦福研究所研制了地矿勘探专家系统PROSPECTOR；
- ◆ 1977年，费根鲍姆提出了“知识工程”的概念，人类从此进入了知识工程时代，知识表示与推理取得了突破。
- ◆ 我国也开始开展了人工智能的研究工作，研制出了若干专家系统。

1980年代末-1990年代中期 :第二个寒冬期 (1)

随着人工智能的应用规模不断扩大，专家系统存在的问题逐渐暴露出来，其局限性如下。

- (1) 应用领域狭窄，缺乏常识性知识。
- (2) 知识获取困难，因为领域专家人数少，而且，有限数量的专家的知识不足以涵盖所有领域知识。
- (3) 知识发生冲突，不同专家对同一问题的理解不同，会导致结论不同。
- (4) 当知识发生动态变化时，知识更新不及时，且知识库难以与已有的数据库兼容。
- (5) 推理方法单一，缺乏分布式功能。
- (6) 人工建设专家系统的效率低、成本高，效果逐渐跟不上需求。

至此，专家系统技术陷入瓶颈，抽象推理不再被继续关注。

1980年代末-1990年代中期 :第二个寒冬期 (2)

基于符号处理的模型（即符号主义）遭到反对，表现如下：

- (1) 1987年，LISP 机的市场崩溃。
- (2) 1988年，美国政府的战略计算促进会取消了新的人工智能经费。
- (3) 1992年，日本的“第五代计算机系统”项目未能达到其初始目标，研制计划宣布失败。但随后启动RWC计划（Real World Computing Project）。
- (4) 1993年，专家系统缓慢走向衰落。

1980年代末-1990年代中期 :第二个寒冬期 (3)

- ◆ 与此同时，用BP算法训练深度神经网络的效果也不好。
- ◆ 1986年鲁梅尔哈特和辛顿等人重新提出BP算法。但用BP训练深度神经网络的效果不好，原因是出现了**梯度消失或梯度爆炸的问题**，人们**对连接主义也再次失去信心**。
- ◆ 至此，无论是**符号主义**方法还是**连接主义**方法，都进入了研究的瓶颈期，致使人工智能的发展全面陷入**第二次寒冬期**。

1997–2011：复苏期（1）

- ◆ 1997年，**深蓝**战胜了卫冕国际象棋冠军加里·卡斯帕罗夫。这一事件又掀起了AI热潮。
（1957年，西蒙曾预测：10年内计算机可以击败人类世界冠军，但实际上晚了30年）
- ◆ 1998年,老虎电子公司推出了第一款用于家庭环境的**人工智能玩具**--菲比精灵(Furby)。
一年后,索尼公司推出了**电子宠物狗**AIBO。
- ◆ 2000年,麻省理工学院推出了**拥有面部表情的机器人** Kismet。
- ◆ 2002年,美国iRobot公司推出了**智能真空吸尘器**Roomba。
- ◆ 2004年，美国国家航空航天局探测车“勇气号”和“机遇号”在火星着陆。
- ◆ 2005年，斯坦福大学研制的**自主机器人车辆**Stanley赢得了DARPA无人驾驶汽车挑战赛冠军。
- ◆ 2006年，CMU研制的**无人驾驶汽车**Boss 赢得了城市挑战赛冠军，Boss 安全通过了临近空军基地的街道，能遵守交通规则，并会避让行人和其他车辆。

1997–2011：复苏期（2）

- 2006年，自从棋王克拉姆尼克(V.Kramnik)被**国际象棋**软件“深弗里茨”(Deep Fritz)击败后，**人类象棋棋手再也没有战胜过计算机**。
- 2006，辛顿教授与其学生R. Salakhutdinov在《科学》上发表了文章，业界公认：该论文开启了**深度学习**发展的浪潮。至此，**连接主义再度兴起**。
- **2009年**，微软AI首席科学家邓力**利用深度学习大大降低了语音识别错误率**，已达到商业应用水平。此后的数年，语音识别迅速衍生出巨大的商业价值。
- 2011年，谷歌启动了深度学习项目——**谷歌大脑**，作为 Google X 项目之一。谷歌大脑是由1万6千台计算机连成的一个集群，致力于模仿人类大脑活动的某些方面。它通过**1千万张数字图片**的学习，已成功地**学会识别一只猫**。
- **2011年**，IBM的**Watson**在智力竞赛《**危险！**》中**战胜了人类**，又掀起了AI研究的热潮。

2012至今：蓬勃发展期（1）

- ◆ 2012年的ImageNet大规模视觉识别比赛(ImageNet Large Scale Visual Recognition Competition, ILSVRC)比赛中，**AlexNet** 一举夺魁，**top5 预测的错误率为15.3%**，远超第二名(26.2%)，2011年的最好成绩是25.8%，从此**深度学习**得到了业界的广泛关注。
- ◆ 2014年6月，俄国**聊天机器人尤金·古斯特曼**，在纪念图灵逝世60周年的一个比赛上，被该活动33%的评委认为古斯特曼是人类，因此组织者认为它已经**通过了图灵测试**。
- ◆ 2015年，百度在2015百度世界大会上推出了一款机器人助理—**度秘**（DUER），可以为用户提供秘书化搜索服务。
- ◆ 到2015年中期，谷歌公司**无人驾驶汽车**的车队已经累计行驶超过150万公里，仅发生了14起轻微事故且均不是由无人驾驶汽车本身造成的。
- ◆ **2016年**3月8日至15日，谷歌 DeepMind团队研制的基于深度学习的围棋程序“阿尔法狗”（**AlphaGo**）以4:1战胜了围棋世界冠军李世石。

2012至今：蓬勃发展期（2）

- ◆ **2017年**5月27日，围棋世界冠军柯洁以 0:3的成绩输给 **AlphaGo**。
- ◆ 2017年10月18日，DeepMind 团队发布了AlphaGo的升级版**AlphaGo Zero**（**阿尔法元**）。
- ◆ **两者的区别**在于：**AlphaGo** 是在学习了人类的3000万个棋局后才打败人类，而**AlphaGo Zero** 则是从零开始自学，在没有任何人类棋谱和先验知识、只输入围棋规则的情况下，完全依靠**强化学习**，用3天时间自己互博了490万个棋局（4个TPU，Tensor Processing Unit），最终以百战百胜的成绩完胜 AlphaGo 。
- ◆ 2017年12月，DeepMind 公司又发布了**AlphaZero**，它不仅可以下围棋，还可以下国际象棋、将棋等棋类。
- ◆ **AlphaZero 的核心思想**是：用蒙特拉洛树搜索算法生成对弈数据，将其作为神经网络的训练数据。

2017年，从 AlphaGo 到 AlphaZero

- ◆ 在完全没有人类棋谱的情况下，在仅经过若干小时的训练后，**AlphaZero** 表现出惊人的能力：
 - 战胜最强 **国际象棋 AI Stockfish**：28胜，0负，72平；
 - 战胜最强 **将棋 AI Elmo**：90胜，2平，8负；
 - 战胜最强 **围棋 AI AlphaGo Zero**：60胜，40负
- ◆ 至此，半个多世纪以来使用的**博弈搜索方法彻底被打败**。

注：

- ◆ **AlphaGo** 和 **AlphaGo Zero** 只是**围棋AI 程序**，**AlphaGo Zero**是 **AlphaGo** 的升级版；
- ◆ **AlphaZero** 是从围棋向其他类似棋类游戏的拓展版，不仅可以下围棋，还可以下国际象棋、将棋等棋类。

DeepMind

- ◆ 2010年，人工智能程序师兼神经科学家戴密斯·哈萨比斯(Demis Hassabis)等人联合创立了DeepMind，是前沿的人工智能企业，位于英国伦敦。
- ◆ 该公司结合机器学习和系统神经科学的最先进技术，建立了强大的通用学习算法。
- ◆ 2014年1月，谷歌斥资4亿美元收购DeepMind。
- ◆ 2015年2月，谷歌DeepMind公司在Nature杂志上发表了Deep Q-Network，通过**深度强化学习**达到人类水平的操控。
- ◆ 2015年12月，谷歌DeepMind公司的程序**AlphaGo**打败了欧洲围棋冠军樊麾，成绩5战5胜。
- ◆ 这个消息直到2016年1月27日才宣布，目的是与描述所用算法的论文在《自然》杂志发表的时间同步。
- ◆ **深度学习软件第一次击败了人类职业围棋棋手。**

2012至今：蓬勃发展期（3）

- ◆ **大语言模型（large language model, LLM）**：GPT（Generative pre-trained transformers）是一种基于 transformer 架构的大语言模型。第一个 GPT 于 2018 年由美国的 OpenAI 公司推出，命名为**ChatGPT**，它是基于 GPT 实现的一个落地应用。
- ◆ **ChatGPT 技术有 2 个特征：(1)大，即参数海量、模型很深；(2)通用性。**目前，它可以：
 - 回答问题：某个主题的基本知识、某个问题的答案
 - 提供建议：根据您的需求给出相关的建议和指导，如健康、财务、职业等
 - 语言翻译：将语言翻译成另一种语言
 - 生成文章：生成符合您需求的文本
 - 解题：理解题目，并给出解题过程和结果
 - 写代码：理解需求，产生一段计算机代码
 - 根据文字指示，进行原画创作.....

2012至今：蓬勃发展期（4）

目前，国内已发布了200多种大语言模型，知名的有：

- 2023.02.20 复旦大学的大语言模型**MOSS**
- 2023.8.31 北京智普华章公司(简称智谱AI)的智谱ChatGLM大模型，基于它开发了生成式AI助手——“**智谱清言**”。
- 2023.03.16 百度**文心一言**
- 2023.04.07 阿里云**通义千问**
- 2023.04.09 360**智脑**大模型
- 2023.04.10 商汤科技的“**日日新**”大模型
- 2023.05.06 科大的“**讯飞星火认知**大模型”
- 2023.9.7 腾讯**混元**大模型
- 中科院的“**紫东太初**大模型”
- 字节跳动自研了通用大模型“**云雀**”，基于云雀大模型的AI对话产品称为“**豆包**”。
- ...

2012至今：蓬勃发展期（5）

- ◆ 文生视频模型：美国时间**2024-2-15**，[OpenAI](#)公司发布了**SORA**。
- ◆ Sora可以根据用户的文本提示创建最长**60秒**的逼真视频，可以深度模拟真实物理世界，能生成具有多个角色、包含特定运动的复杂场景。
- ◆ Sora这一名称源于日文“空”（**そら sora**），即天空之意，以示其无限的创造潜力。
- ◆ Sora是在OpenAI的文本到图像生成模型**DALL-E**基础上开发而成的。
- ◆ OpenAI表示，Sora存在不成熟之处，可能难以理解因果关系。
- ◆ 多位人工智能领域人士表示，该问题可能因其概率模式的逻辑存有“硬伤”。加大训练量、增加训练数据与**物理逻辑**可改善该问题，但无法根治。想要真正突破最底层逻辑上的问题，**因果关系**是一条必经之路。

Watching Public Course Video

- ◆ Google Deepmind behind AlphaGo--- Self-learning artificial intelligence
- ◆ 观看网易公开课视频: [阿尔法狗背后的谷歌DeepMind: 能自我学习的人工智能](http://open.163.com/movie/2016/5/G/E/MBMDNOFJB_MBMDOBQGE.html)----14 mins
http://open.163.com/movie/2016/5/G/E/MBMDNOFJB_MBMDOBQGE.html

1.4 人工智能研究的流派

人工智能按研究流派主要分为三类，分别是

- ◆ 符号主义(Symbolism)
- ◆ 连接主义(Connectionism)
- ◆ 行为主义(Behaviorism)

符号主义和连接主义研究流派，在人工智能发展的 60-70 年中，此消彼长，轮流占上风。目前，连接主义是业界关注的焦点。

符号主义流派

- ◆ **符号主义**又称为逻辑主义（Logicism）、心理学派（Psychologism）或计算机学派（Computerism）。
- ◆ 该学派认为：人工智能源于数理逻辑，其主要理论基础是物理符号系统假设。符号主义是基于符号逻辑的方法，用逻辑表示知识和求解问题。
- ◆ **基本思想**：用一种逻辑把各种知识都表示出来；当求解一个问题时，就将该问题转变为一个逻辑表达式，然后用已有知识的逻辑表达式的库进行推理来解决该问题。
- ◆ 符号主义**认为**：只要在符号计算上实现了相应功能，那么在现实世界就实现了对应的功能，这是智能的充分必要条件。
- ◆ 这一流派的代表人物是西蒙和纽厄尔，他们提出了物理符号系统假设，**图灵测试就是符号主义流派的思想实验**。

符号主义流派

有三个本质问题制约了符号主义的发展：

- ◆ **逻辑体系问题：**难以找到一种简洁的符号逻辑体系，能表述出世间所有的知识。
- ◆ **知识提取问题：**即使限定在某一特定领域，仍无法将该领域内的所有知识都提取出来，并全部用逻辑表达式记录下来。
- ◆ **求解复杂度问题：**逻辑求解器的时间复杂度很高。理论上不存在一种通用方法，能在有限时间内判定任意一个谓词逻辑表达式是否成立。

连接主义流派

- ◆ **连接主义**又称为仿生学派（Bionicsism）或生理学派（Physiologism），其**主要原理**为神经网络及神经元间的连接机制与信息传播算法。
- ◆ 连接主义试图发现人类大脑的结构及其处理信息的机制，在计算机上用人工神经网络来模拟人类大脑的神经元及其连接机制，进而实现机器的智能。
- ◆ 连接主义方法起始于1943年，建立了**MP神经元模型**，这是最早的人工神经网络。
- ◆ 事实上，连接主义方法大幅度地抽象、简化了人脑生物神经元细胞网络，一个人工神经元可以将 N 个输入进行加权汇总后，再通过一个非线性函数得到该神经元的输出。

行为主义流派

- ◆ **行为主义**又称为进化主义（Evolutionism）或控制论学派（Cyberneticsism）。
- ◆ 该学派的理论基础是控制论，其**核心思想**是基于控制论构建感知-动作型控制系统。
- ◆ 行为主义学派认为：智能行为是在现实世界中与周围环境交互而获得并表现出来的，人工智能可以像人类的智能一样逐步进化，所以又称为**进化主义**。
- ◆ 行为主义还认为智能取决于感知和动作，无需知识、表示和推理，只需要表现出智能行为即可，**强化学习就属于这一流派**。
- ◆ 行为主义学派的学者对传统的人工智能提出了批评和挑战，否定智能行为来源于逻辑推理及其启发式的思想，认为AI研究的重点不应该是知识表示和编制推理规则，而应该是在复杂环境下对行为的控制。这种思想对早期占据主流的符号主义学派是一次冲击和挑战。

1.4 人工智能研究的流派

- ◆ 纵观人工智能发展的历史，**符号主义流派**在很长时间内都处于人工智能研究的主流地位。
- ◆ 近年来，**深度学习**使得**连接主义流派**占据了人工智能研究的主流地位。
- ◆ 虽然深度学习在处理感知、识别和判断等方面表现突出，但在模拟人的思考过程、处理常识知识和推理，以及理解人的语言方面仍然举步维艰，
- ◆ 在这方面，**从专家系统（符号主义）发展起来的知识图谱**已表现出了其发展的潜力和优势。

1.4 人工智能研究的流派

- ◆ 仅遵循某单一学派的研究思路和方法，是不足以实现人工智能的，很多时候**需要综合各个学派的技术**。
- ◆ 例如，围棋系统**AlphaGo综合使用了三种学习方法**：蒙特卡罗树搜索、深度学习和强化学习
 - 蒙特卡罗树搜索属于符号主义
 - 深度学习属于连接主义
 - 强化学习属于行为主义
- ◆ 可见，AlphaGo 用到了三个人工智能流派的方法和技术。
- ◆ 目前，人工智能的各个研究流派依然在发展，但各个流派的融合发展已是大势所趋。

1.5 人工智能研究的基本内容

1.知识表示

- ◆ 知识是人类进行一切智能活动的基础。
- ◆ 人工智能对问题的求解都是以知识为基础的，因此需要将知识以计算机能够处理的形式表示出来，才能存储到计算机中，并被有效地利用，所以知识表示是人工智能研究的一个基本内容。
- ◆ 目前学术界尚未彻底掌握人类知识的结构与机制，也未建立起知识表示的理论体系和统一规则。但学者们还是总结出了一些知识表示的方法。
- ◆ 常见的知识表示方法包括：一阶谓词逻辑表示法、产生式表示法、框架表示法、语义网络表示法、状态空间表示法、神经网络表示法、过程表示法等，这些知识表示方法在不同的专业领域和应用背景下发挥着各自的作用。

1.5 人工智能研究的基本内容

2.机器感知

- ◆ 机器感知是使计算机系统模拟人类通过其感官与周围世界联系的方式具有解释和理解外部信息的能力。
- ◆ 机器感知外界信息需要通过照相机、摄像机、雷达、红外成像仪等设备，它研究如何使机器具有像人类一样的感知或认知能力：机器视觉、机器听觉、机器触觉、机器嗅觉等。
- ◆ 目前，人工智能研究领域中的机器感知主要聚焦于机器**视觉**和机器**听觉**。
- ◆ 机器听觉是让机器能识别并理解语言、声音等，包括语言识别、语音识别、自然语言理解等；
- ◆ 机器视觉是让机器能够识别并理解文字、图像、场景等，包括文字识别与理解、图像识别、图像理解、模式识别等。

1.5 人工智能研究的基本内容

3. 机器思维

- ◆ 又称为计算机思维（Computer Thinking），是研究如何使计算机能像人类一样进行思维活动，要研制会自主思考的机器。
- ◆ 现有的计算机需要人发出指令、编写程序，才能操作，智能水平不高。
- ◆ 需要研究更聪明的、思维能力更强的智能机器。
- ◆ 正如人的智能来自于大脑的思维活动，机器智能也主要是通过机器思维实现的。
- ◆ 人工智能最关键的难题：机器自主创造性思维能力的建立与提升。
- ◆ 机器思维是人工智能研究中最重要、最核心的内容。

1.5 人工智能研究的基本内容

4. 机器学习

- 人类是**通过学习具有智能**的，计算机也必须学习，才能具有真正的智能。
- 机器学习专门研究如何使计算机模拟或实现人类的学习行为，使计算机能获取新知识和新技能，并在实践中不断地提升性能，实现自我完善。
- 机器学习是AI领域中最具智能特征、最前沿的研究热点。
- 机器学习的研究领域涵盖范围广，它与概率论、数理统计学、认知科学、神经心理学、机器视觉、机器听觉等都有密切联系。
- 近年来，随着大数据、算法、算力的迅猛发展，机器学习研究已取得了长足进步，尤其是深度学习的研究成果在很多领域已有了成熟的应用，**但并未从根本上解决问题**。

1.5 人工智能研究的基本内容

5. 机器行为

- 指机器具有人工智能的行为，或者说，机器能模拟、延伸与扩展人的行为。
- 目前的机器智能还远远达不到人的水平。
- 现阶段，机器行为仅局限于计算机的识别和表达能力，即听、说、读、写、画等能力，智能机器人还应具有人的四肢的功能，即能走路、抓取、踢球、肢体互动、操作工具等。

1.6 人工智能的主要研究领域

1. 深度学习
2. 自然语言理解
3. 计算机视觉
4. 智能机器人
5. 自动程序设计
6. 数据挖掘与知识发现

人工智能基础的课程思政

◆ 在当前社会

- 人工智能不再是一个单独的学科，而是普遍方法、通用方法。
- 人工智能不再是科学问题，而是而是国家战略问题。
- **人工智能**不再是一国之事，而是多国协作竞争，**关乎国运的事**。

◆ 大学生从现在开始，有必要了解人工智能的基本思想，需要对人工智能的发展有正确的认识。

◆ 尤其是计算机和软件工程相关专业的学生，要钻研算法、立足需求、自主研发，国家才能逐步从**加工大国**走向**科技强国**，**才能扭转被“卡脖子”的窘境**。

1.7 本章小结

- ◆ 人工智能已发展八九十年了，由于其涵盖的研究内容和学科门类众多，至今还没有公认的、确切的定义。人工智能的发展过程也并非一帆风顺，期间经历了三起两落。
- ◆ 根据不同的研究方法和技術，科学家们将人工智能分为符号主义、连接主义和行为主义3个主要研究流派，它们各领风骚若干年，都在其各自的辉煌时期作出了巨大的贡献。
- ◆ 人工智能研究的基本内容包括：知识表示，机器感知（听、说、看），机器思维，机器学习、机器行为等方面。
- ◆ 人工智能的研究领域十分广泛，现阶段，其主要研究领域包括：深度学习、自然语言理解、计算机视觉、智能机器人、自动程序设计、数据挖掘与知识发现等方面。