



人工智能基础与进阶

走进人工智能

上海交通大学

目录 content



第一节 什么是人工智能

第二节 人工智能简史

第三节 人工智能对人类的影响

第四节 当前人工智能的局限

第五节 人工智能的未来

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第一节 什么是人工智能



上海交通大学人工
智能创新教育实验室

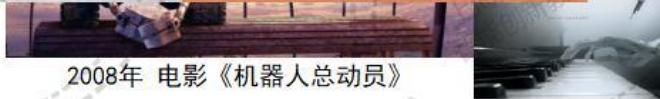
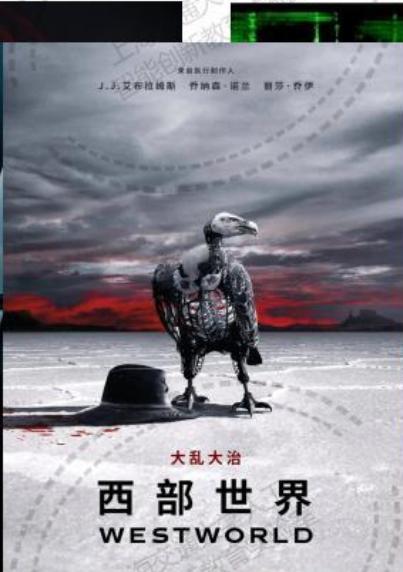
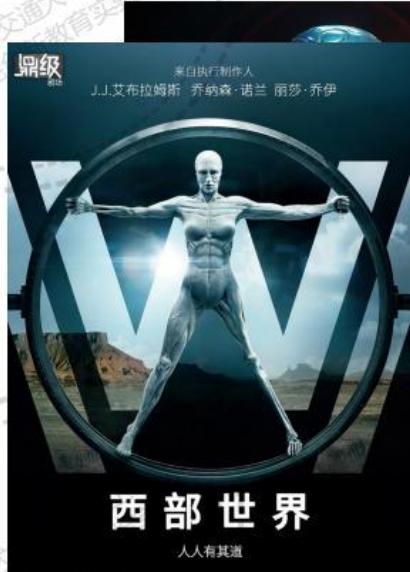
上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

瑞士认知科学家 Jean Piaget 说，

智能是当你不知道如何做的时候你用的东西。

什么是人工智能—电影中的人工智能



什么是人工智能—形式多样风格迥异的人工智能



AlphaGo



餐厅机器人



大堂机器人



商场机器人



人工智能画家Deep Dream



写稿机器人



助残机器人



音乐机器人

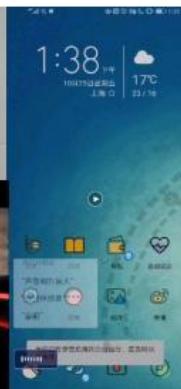
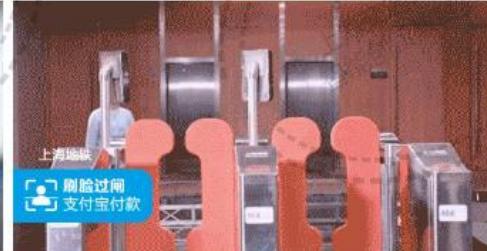


无人车



消毒机器人

什么是人工智能—形式多样风格迥异的人工智能



什么是人工智能—现实中多种形式的人工智能

扫地机器人：

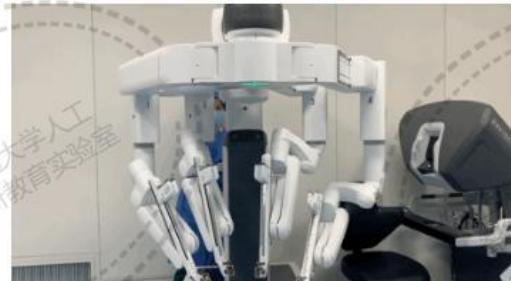


自动化车间：



什么是人工智能—形式多样风格迥异的人工智能

达芬奇手术机器人



什么是人工智能—形式多样风格迥异的人工智能

手机中的“大智慧”



智能图像理解
智能美图

新闻推荐
智能搜索排序

智能出行
自动驾驶

智能会话
智能助理

机器翻译

个性化推荐
智能物流
仓储机器人

识别：
语音、人脸、视频内容
应用：安防、客服。。。

BI、商业流程自动化

互联网数据
购买转化、商品推荐、定价、精准营销、社交媒体营销

券商股票：
智能+量化交易、智能投顾、机器人理财

银行、保险应用：
针对性电话营销、贷款审批、信用卡欺诈

医学：
自动读片、辅助诊断、个性化诊断、基因排序

教育：
学习外语、智能选题



今日头条
你关心的才是头条



无处不在的人工智能—智能语音助手可以联网进行不断自我学习

上海交通大学人工
智能创新教育实验室

小爱同学



智能音箱

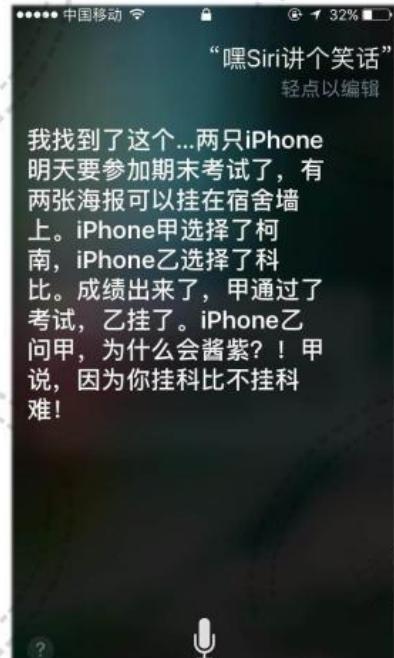


表现：经常答非所问

趋势：逐步展现出与人类沟通的能力。

智慧：还远不及人类

方法：通过不断学习，寻找最匹配的回答。



到底什么是人工智能——会学习的计算机程序?

早期人工智能兴起
改变了人们的观念



人工智能是人类社会
发展主要目标

机器学习开始兴起



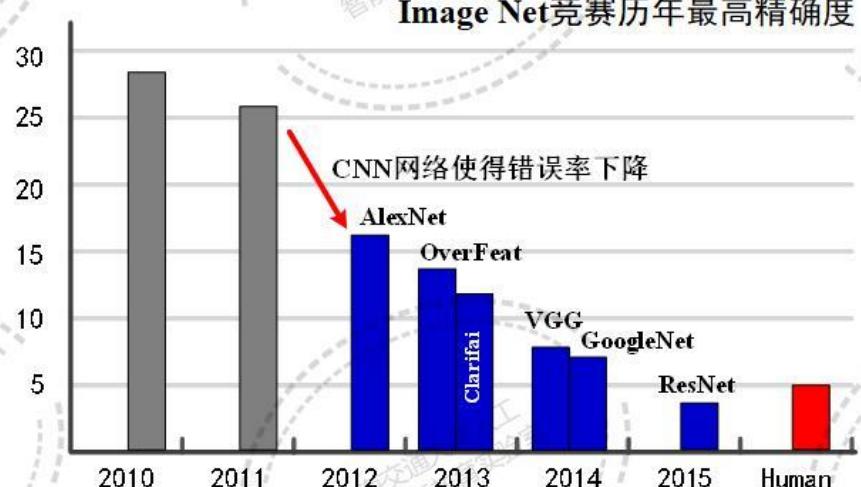
机器学习是实现人工
智能的核心技术

机器学习取得突破驱动
人工智能的蓬勃发展



深度学习是机器学习
中最热门的算法

到底什么是人工智能— 会学习的计算机程序?



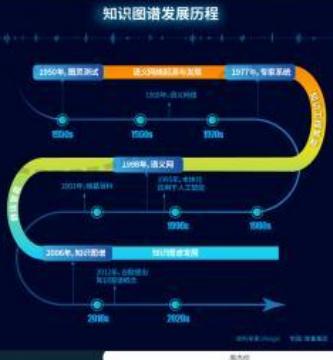
到底什么是人工智能— 感知环境与理性行动的智能体?

人工智能 是有关智能主体的研究与设计的学问

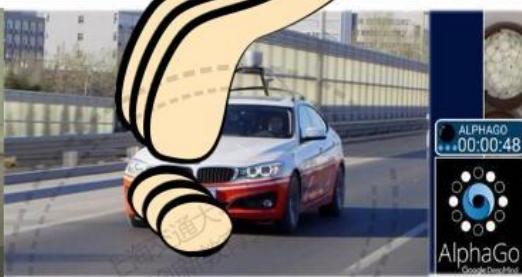
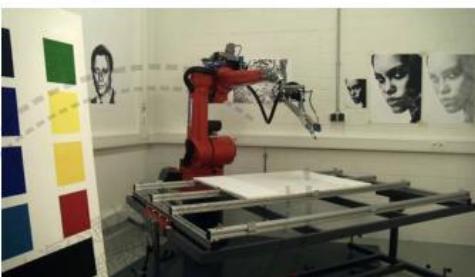
智能主体 是指一个可以观察周围环境并做出行动以达至目标的系统



到底什么是人工智能— 聪明的交互系统?



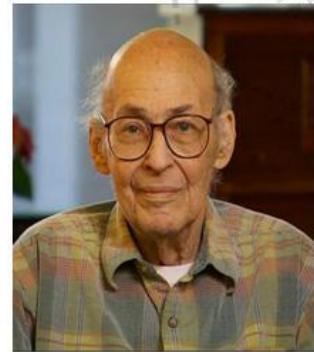
到底什么是人工智能



到底什么是人工智能— 会帮助人做事?

马文·明斯基 (Marvin Minsky):

**“人工智能就是让机器来完成那些如果由人来做则
需要智能的事情的科学”**



- **人工智能之父**

- 1969年获图灵奖, 获此殊荣的第一位人工智能学者
- 1969年获图灵奖, 1991年获IJCAI终身成就奖。在人工智能、认知心理学、数学、计算语言学、机器人学等领域都做出了杰出贡献。他创建了MIT的AI实验室、还是MIT的Media实验室奠基人。

到底什么是人工智能— 人工智能不同解读



像人一样思考的系统



理性地思考的系统



像人一样行动的系统



理性地行动的系统

“要使计算机能够思考.....意思是：
就是：有头脑的机器”
(John Haugeland, 1985)

“通过利用计算模型来进行心
智能力的研究”
(Chamiak 和 McDermott,
1985)

“一种技艺，创造机器来执
行人需要智能才能完成的功
能” (Kurzweil, 1990)

“计算智能是对设计智能化
智能体的研究”
(Alan Mackworth, 1998)

到底什么是人工智能—达特茅斯会议主题

60多年前的达特茅斯大学



会议主题：

如何用机器来模仿人类学习以及其他方面的智能？

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第二节 人工智能简史

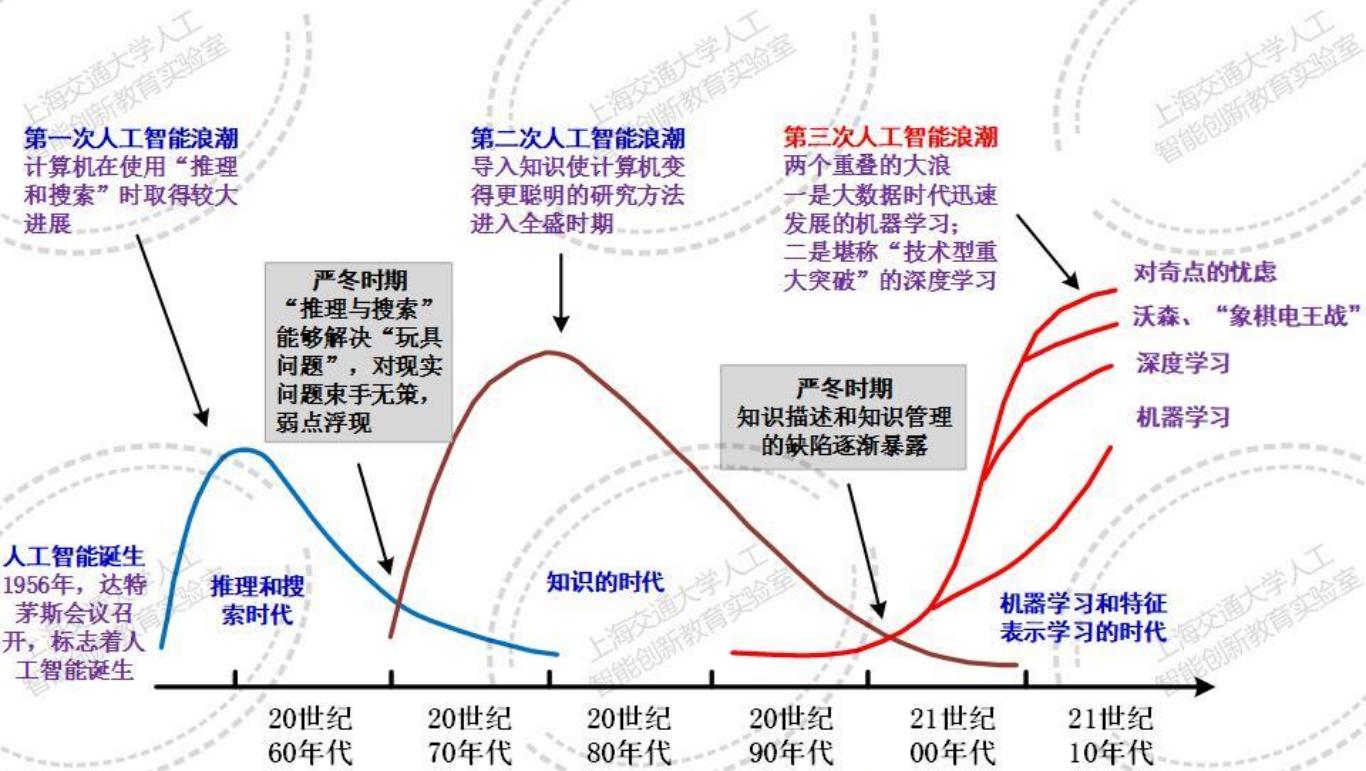


上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

人工智能简史



摘自：<https://www.useit.com.cn/thread-16571-1-1.html>

人工智能简史—达特茅斯会议

60多年前的达特茅斯大学



达特茅斯会议参会者



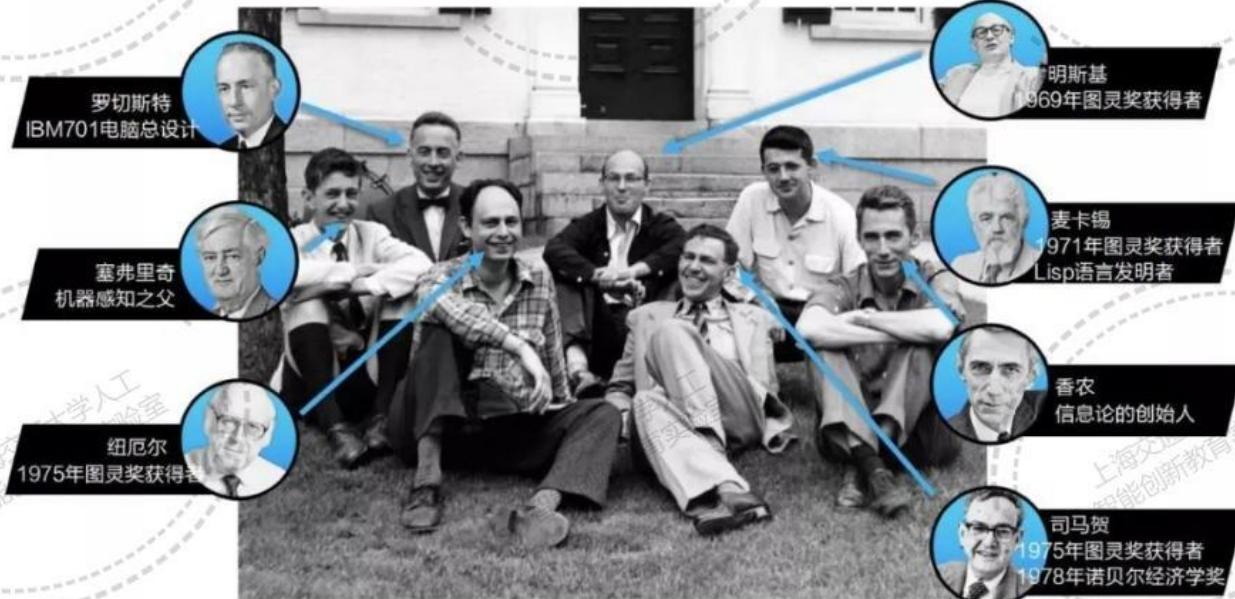
1956年的夏天，香农和一群年轻的学者在达特茅斯学院召开了一个头脑风暴式研讨会。

**会议主题：
如何用机器来模仿人类学习以及其他方面的智能？**

讨论的是当时计算机尚未解决，甚至尚未开展研究的问题，包括人工智能、自然语言处理和神经网络等。

参会者均为人工智能领域的拓荒者

达特茅斯会议7侠



人工智能诞生的土壤



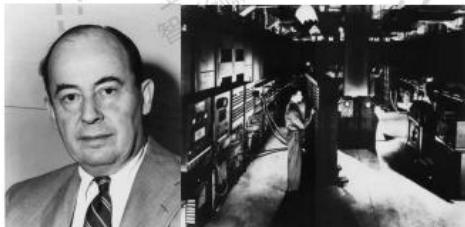
莱布尼茨
(1646—1716年)
德国哲学家、数学家
发明几何微积分



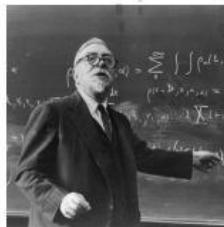
乔治·布尔
(1815—1864年)
英国数学家

莱布尼茨，数理逻辑的创始人，
乔治·布尔在符号逻辑运算领域的研究成
果为历史上计算机的发展和当今的人工智
能革命打下了坚实基础

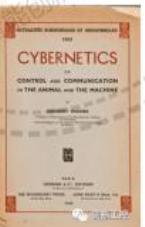
阿兰·图灵
(1912—1954年)
英国数学家、逻辑学家，被称
为计算机科学之父、人工智能
之父



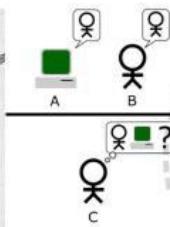
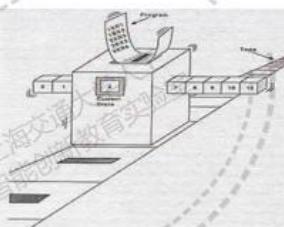
冯·诺依曼，1903-1957年，物理学家、数学家。
提出了现代计算机的架构，1946年EDVAC诞生



诺伯特·维纳，1948年，数学家，控制论之父，《控制论》控制理论学科奠基者，现代计算机的先驱（现代计算机设计的五项原则）。



1936年，阿兰·图灵提出
图灵机模型；
1950年，《计算的机器
和智能》提出了一个验
证机器有无智能的判
别方法。



人工智能简史—黄金年代：1956 - 1974

第一代AI研究者们非常乐观，曾做出了如下预言：

- 十年之内，数字计算机将成为国际象棋世界冠军。
- 二十年内，机器将能完成人能做到的一切工作。
- 再有三到八年的时间里将得到一台具有人类平均智能的机器。

方法：首先了解人类是如何产生智能的，然后让计算机按照人的思路去做。

语音识别、机器翻译等领域迟迟不能突破

我
能
怎
么
办
我
也
很
绝
望



人工智能简史—第一次低谷：1974 - 1980

低谷：承诺无法兑现、乐观情绪遭到打击、研究经费缩减

贾里尼克 (Fred Jelinek) 在 IBM 做语音识别，利用解决通讯方法来解决问题。由于计算机性能的瓶颈、计算复杂性的指数级增长、数据量缺失等问题，一些难题看上去好像完全找不到答案。

历史上赫赫有名的《莱特希尔报告》(Lighthill Report) 针对英国人工智能研究状况的报告批评了人工智能在实现其“宏伟目标”上的完全失败。美国和英国政府纷纷撤回了对人工智能研发的资助资金。

二十世纪70年代 人工智能 “萧条期”

小贴士：莱特希尔，20世纪世界上伟大的数学科学家之一，1979~1989年，为伦敦大学校长。他在29岁时当选英国皇家协会会员。



人工智能简史——繁荣：1980 - 1987

繁荣：新方法、新理论、新途径、大投入

- 一类名为“专家系统”的AI程序开始为全世界的公司所采纳，而“知识处理”成为了主流AI研究的焦点。
- John Hopfield证明一种新型的神经网络（现被称为“Hopfield网络”）能够用一种全新的方式学习和处理信息。
- David Rumelhart推广了反向传播算法。
- 日本经济产业省巨款支持第五代计算机项目。
- 英国开始了耗资三亿五千万英镑的阿尔维计划



腻害！

人工智能简史—第二次AI低谷：1987 - 1993

二次低谷：费用昂贵、局限性、预期落空

- Apple和IBM生产的台式机性能不断提升。
- 专家系统维护费用居高不下。
- 专家系统的实用性仅仅局限于某些特定情景。
- 日本人宏伟的“第五代工程”并没有实现。



不切实际的预期导致更大的失望



我能怎么办？我也很绝望呀

人工智能简史—走在正确的路上：1993 - 2005

各种理论的不断完善与丰富

- IBM的超级计算机深蓝大战人类国际象棋冠军卡斯伯罗夫。
- Judea Pearl发表于1988年的名著将概率论和决策理论引入AI。
- 投入应用的新工具包括贝叶斯网络，隐马尔可夫模型，信息论，随机模型和经典优化理论。
- 针对神经网络和进化算法等“计算智能”范式的精确数学描述也被发展出来。

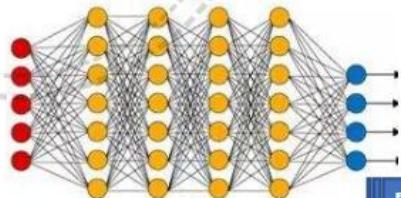
人类从1996年的 4: 2 到1997年的 2.5: 3.5

深蓝团队其实把一个机器智能问题变成了一个大数据
和大量计算的问题。



人工智能简史—大数据：2005 - 现在

深度神经网络



量变导致了质变

智能革命

变智能问题为数据问题

人工智能+



大数据

小贴士：阿尔法围棋运用神经网络、深度学习、蒙特卡洛树搜索法等，使其实力实质性飞跃，击败数十位高手，连续60局无一败局。

开什么玩笑
我们是故意走到绝路的



人工智能简史—大数据：2005 - 现在

上海交通大学
智能创新教育实验室

人工智能

钱是挣出来的
不是省出来的



应用层
技术使用者

技术层
算法、产品
及
解决方案提供者

机器学习引擎

基础层
数据及计算资源



AI走向3.0的原因分析

浅层次动因：需求、环境和技术

□ 信息环境巨变：

互联网、移动计算、穿戴设备、物联网、云计算、网上社区、搜索引擎等等

□ 社会新需求爆发：

智慧城市、智能医疗、智能交通、智能游戏、无人驾驶、智能制造等等

□ AI的基础和目标巨变：

大数据、多媒体、传感器网、增强现实（AR）、虚拟现实（VR）等等

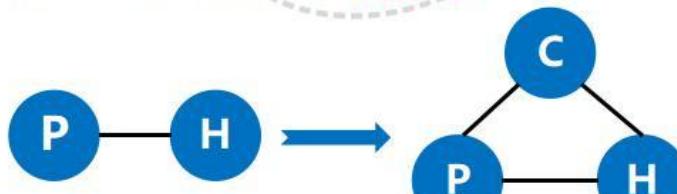
计算机模拟人的智能 → 人机智能 → 群体智能

人工智能迈向新一代

AI走向3.0的原因分析

深层次分析：世界正从原来的二元空间进入新的三元空间

世界原来是二元空间：**人类社会空间（H）** 和**物理空间（P）**。但近年来，信息力量的迅速壮大，已长成除P、H两极之外的新一极：**信息空间（C）**。



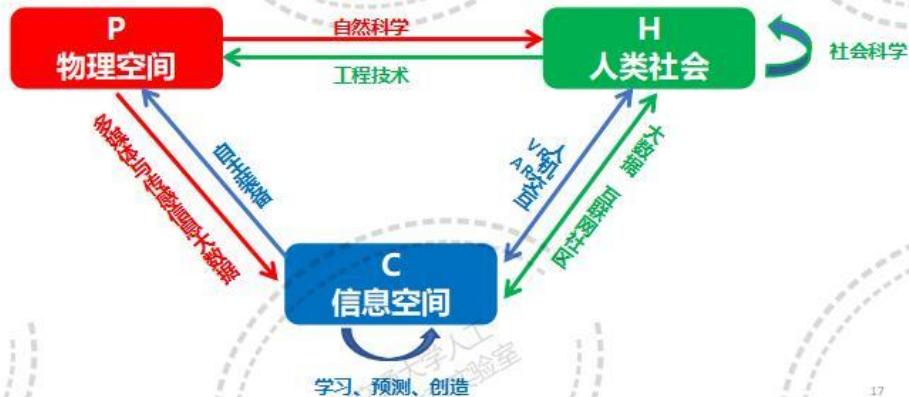
近50年来，信息空间成长壮大的历程

- 信息来自于人类社会：书籍、交流、媒体、计算机
- 信息开始互联：互联网、移动通讯和搜索
- 信息绕过人类，直接来自于物理世界：传感器和物联网
- 产生大数据：信息快速膨胀，人类已无法单独处理
- 从大数据必然会走向**大知识**，并推动人类认识与控制能力的大变化

AI走向3.0的原因分析

新认知

- 新门类 认识复杂巨系统：城市运行系统、环境生态系统、健康医疗系统（科学+工程+社会+影响）
- 新通道 给自然科学、工程技术、社会科学提供了新途径、新方法
- 新计算 AI2.0：建立在新老空间的互动CH、CP之上的AI



□ 空间变化带来了认知的新门类、新通道、新计算

AI走向3.0的原因分析

AI走向3.0



从PH到CPH的人类空间演变并未结束，而正在深化。AI走向3.0正是这种深化演变为的一个结果

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第三节 人工智能派别



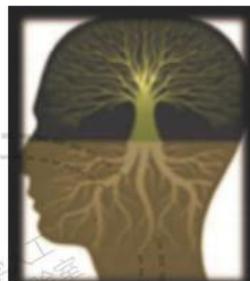
上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

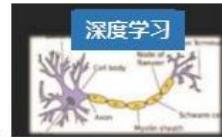
人工智能的三大门派

上海交通大学人工
智能创新教育实验室



聪明的AI

学习
归纳



有学识的AI

认知
表达
规则



会做事的AI

推理
感知
定位
规划



行动

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

人工智能类别 — 符号主义

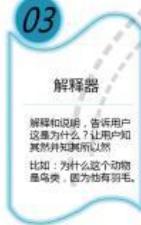
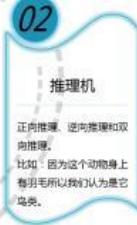
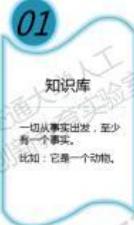
符号主义

又称：逻辑主义、心理学派或计算机学派

原理：物理符号系统假设和有限合理性原理

起源：源于数理逻辑

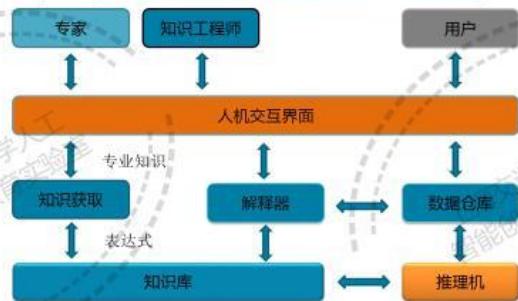
学派代表：纽厄尔、西蒙和尼尔逊等



认为人的认知基元是符号，
认知过程即符号操作过程。

认为人是一个物理符号系统，计算机也是一个物理符号系统，因此，能用计算机来模拟人的智能行为。

认为知识是信息的一种形式，是构成智能的基础。人工智能的核心问题是知识表示、知识推理。



人工智能类别 — 连接主义



原理：神经网络及神经网络间的连接机制与学习算法。

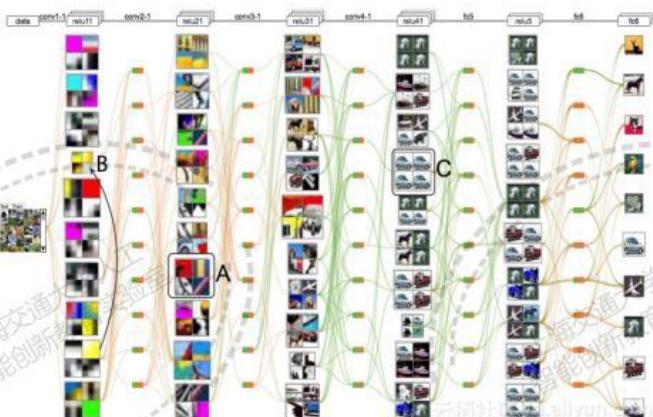
起源：源于仿生学，特别是人脑模型的研究。

学派代表：麦卡洛克、皮茨、Hopfield、鲁梅尔哈特等。



连接主义基本理论：

- 认为思维基元是神经元，而不是符号处理过程。
- 认为人脑不同于电脑，并提出连结主义的大脑工作模式，用于取代符号操作的电脑工作模式。



人工智能类别 — 行为主义

行为主义

又称：进化主义或控制论学派

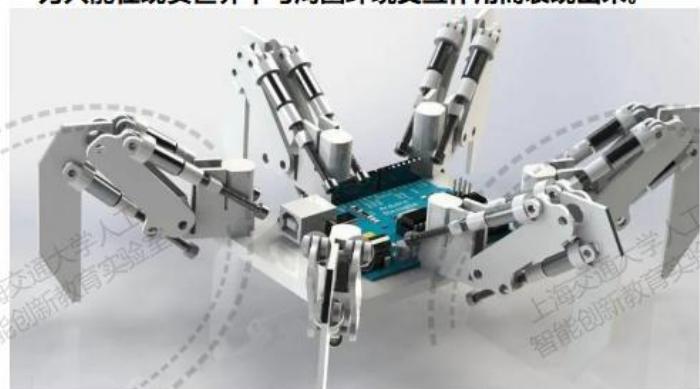
原理：控制论及感知—动作型控制系统

起源：源于控制论

学派代表作：布鲁克斯(Brooks)的六足行走机器人，一个基于感知-动作模式的模拟昆虫行为的控制系统

行为主义基本理论

- 认为智能取决于感知和行动(所以被称为行为主义)，提出智能行为的“感知—动作”模式；
- 认为智能不需要知识、不需要表示、不需要推理；人工智能可以像人类智能一样逐步进化(所以称为进化主义)；智能行为只能在现实世界中与周围环境交互作用而表现出来。

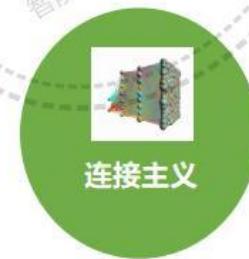


人工智能类别 — 三大学派的集成



符号主义

+



连接主义

+



行为主义

随着研究和应用的深入，人们又逐步认识到，三个学派各有所长，各有所短，应相互结合、取长补短，综合集成。不同学派间的争论进一步促进了人工智能的发展。

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第四节 人工智能对人类的影响



上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

人工智能层级—弱人工智能（Weak AI）



弱人工智能

1. 只能解决特定领域问题的人工智能；
2. 属于相对容易控制和管理的计算机程序；
3. 只是一种技术工具。

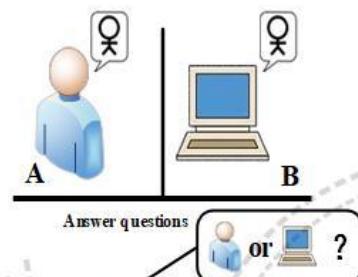
只要严格控制，严密监管，人类也可以像使用其他工具一样放心地使用目前所有的人工智能技术。

人工智能层级——强人工智能（Strong AI）

强人工智能又称为通用人工智能或完全人工智能，它是可以胜任所有人类工作的人工智能。这个定义有点过于宽泛，缺乏一个可以量化的标准来评估一个计算机程序是否是强人工智能。

- 推理、使用策略、制定决策、解决问题的能力
- 知识表示能力
- 规划能力
- 学习能力
- 使用自然语言
- 整合上述能力

是否需要具备“意识”



图灵测试

计算机行为与人类行为不可区分性
具有怎样的物质或能力？

人工智能层级—超人工智能（Superintelligence）



超人工智能定义最为模糊甚至无法定义

每一方面都比最强的人类大脑聪明很多的智能
无法知晓超越人类最高水平的智慧为何种能力

恐惧来自未知

奇点

如果把整个人类大约6000年的文明浓缩成24小时

科技发展是越来越快的，呈现出不断加速的势头

人类文明开始

00:00

04:00

08:00

12:00

16:00

20:00

24:00

Alpha Go战胜围棋选手李世石

人类进入移动互联网时代

人类开始使用互联网

人类发明通用电子计算机

人类学会使用电力

欧洲发明蒸汽机

毕昇发明活字印刷术

上海交通大学
智能创新教育实验室

上海交通大学
智能创新教育实验室

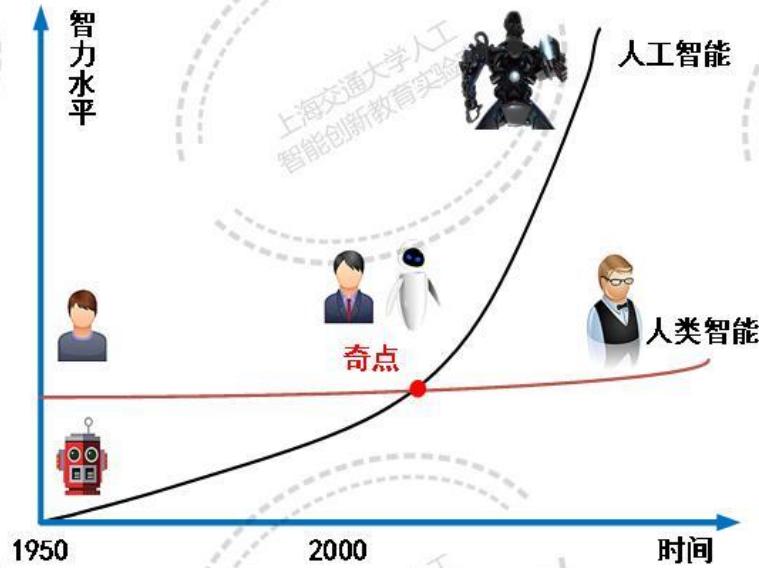
上海交通大学
智能创新教育实验室

上海交通大学
智能创新教育实验室

上海交通大学
智能创新教育实验室

上海交通大学
智能创新教育实验室

奇点——摩尔定律



蒂姆·厄班的推理

几十年的时间达到人类脑残的水平

一小时后电脑立马推导出了广义相对论和量子力学的物理理论

一个半小时后这个强人工智能变成了超人工智能，智能将达到普通人类的17万倍。”

奇点—摩尔定律的瓶颈

价格不变时，集成电路可容纳的元器件数目每隔18-24个月会增加一倍，性能也将提升一倍。

集成电路的发展，曾在1975年到2012年的数十年间保持稳定的增长，但在2013年后**显著放缓**。

高登·摩尔本人也说：“我猜我可以看见摩尔定律会在大约**10年内失效**，但这并不是一件令人吃惊的事。”



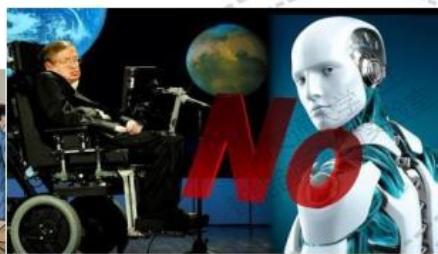
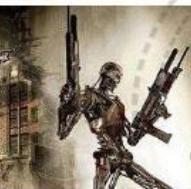
奇点——霍金预言

霍金的担忧还是建立在人工智能技术将以**加速发展的趋势**不断发展的基础上

霍金就通过媒体告诉大家：“完全人工智能的研发可能意味着人类的末日。”

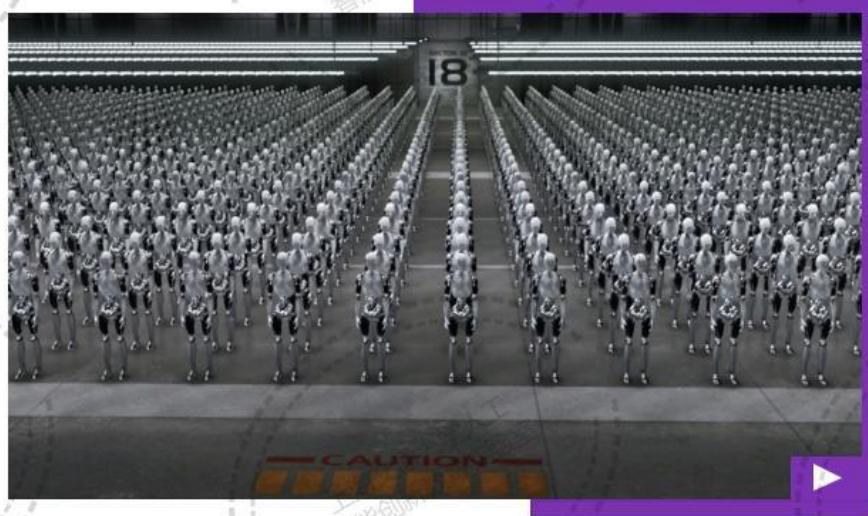
霍金曾说：“人工智能可以在自身基础上进化，可以一直保持加速发展的态势，不断重新设计自己，而人类的生物进化速度相当有限，无法与之竞争，终将被淘汰。”

人类离超人工智能的威胁真的还非常遥远？



人工智能的发展争议

人工智能会拥有情感，奴役人类吗？



人类与人工智能之争

统治？被统治？



最极端的假设则预测了一个人工智能比人类更加聪明的遥远未来。

人工智能的发展依然处于非常初级的阶段，现状基本就是——

“没有人工，就没有智能”

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

第五节 当前人工智能的局限



上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

当今人工智能局限——跨领域推理

人类明显的能力优势：**举一反三、触类旁通**的能力。三

四岁的小孩在描述一样东西时

“像羽毛一样轻”、“像蚂蚁一样小”

文学作品中更体现出人类的想象与推理

“人固有一死，或重于泰山或轻于鸿毛，为人民利益而死，就比泰山还重……” —《为人民服务》毛泽东

“五岭逶迤腾细浪，乌蒙磅礴走泥丸”

—《七律·长征》毛泽东

计算机很难自发地总结出事物间的相似性和关联性

当今人工智能局限—抽象能力



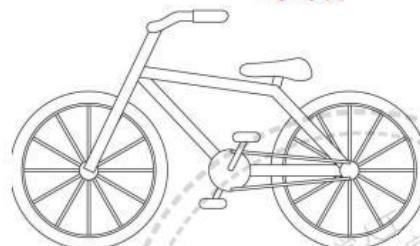
人类超强的抽象能力：数学理论

自然数

负数

对称性

小数

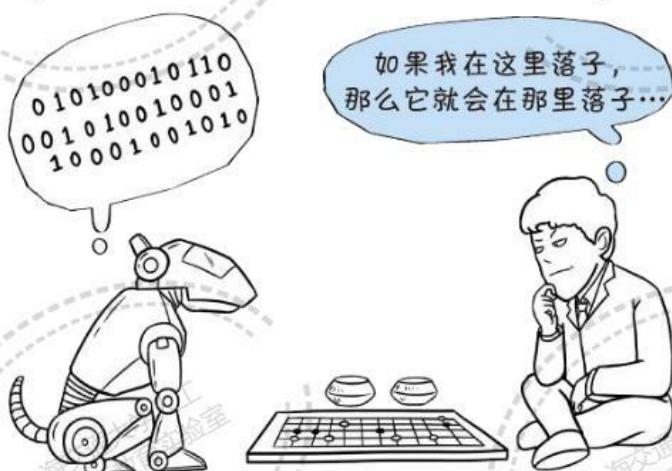


学习对比

深度学习：大量训练样本

人类学习：只要两三个样本就可以了

当今人工智能局限—知其然，也知其所以然



与其说机器在思考，不如说它在计算

- ◆ 深度学习：通过自动调整参数来完成模型的学习，在决策时**难以解释**其中的道理。
- ◆ 围棋专家自然可以用自己的经验来解释计算机所下的大多数棋。
- ◆ 人类专家的理论是成体系的、有内在逻辑的，但这个体系和逻辑却并不一定是计算机能简单理解的。
- ◆ 人通常追求“知其然，也知其所以然”，但目前的弱人工智能程序，大多都只要结果足够好就行了。

当今人工智能局限一常识

常识：

- ◆ 一个心智健全的人应当具备的基本知识
- ◆ 人类与生俱来的，无须特别学习就能具备的认知、理解和判断能力
- ◆ 常识可以给人类带来很多便利
- ◆ 机器中的常识更多的还只是一些预设规则，远未如人类所理解的“常识”那么丰富



两点之间直线最短

当今人工智能局限一自我意识



“认知、意识、情感”

感觉机器只有具备了自我意识后才能称为真的智能。

拥有意识的机器人也必定会像人类那样发出“对酒当歌，人生几何？譬如朝露，去日苦多”之类的感慨。

今天的弱人工智能还远未达到自我意识的地步

当今人工智能局限一审美

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室

上海交通大学人工
智能创新教育实验室



- ◆ 机器已经可以仿照，但机器并不真正懂得什么是美。
- ◆ 审美能力同样是人类独有的特征，但很难用技术语言解释，也很难赋予机器。

当今人工智能局限—审美



人工智能复制了蒙克、佛里达、马蒂斯等著名画家的风格，应用到同
一张照片上。



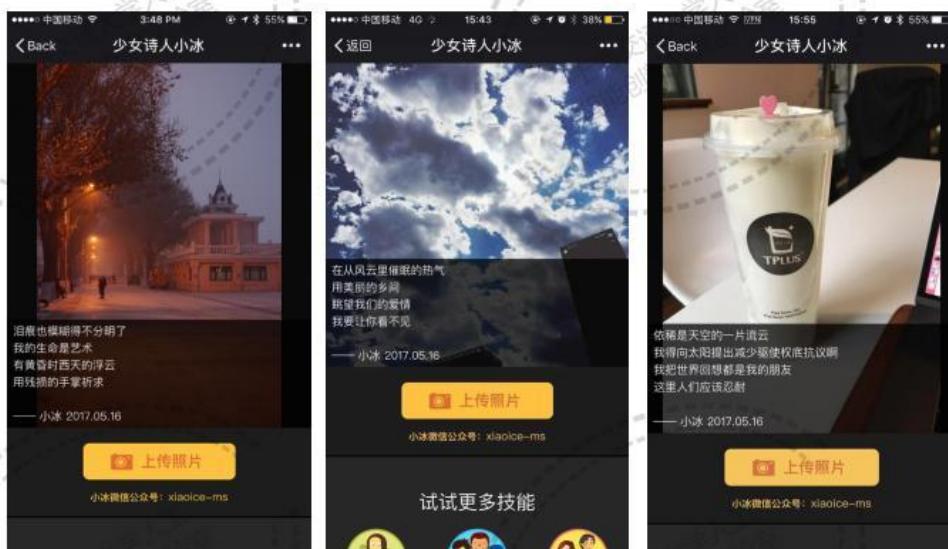
梵高的《星空》风格被复制
在了其他照片上



◆ 机器已经可以仿照，但机器并不真正懂得什么是美。

◆ 审美能力同样是人类独有的特征，但很难用技术语言解释，也很难赋予机器。

当今人工智能局限一审美



私人定制“三百首”
口诀：输入量难忘的地点或名字，为你生成专属诗歌

三里屯 生成

三千青丝夜成霜，里一白发人断肠
屯千古道游子梦，情百年心不彷徨

分享给朋友 再玩一次

- ◆ 机器已经可以仿照，但机器并不真正懂得什么是美。
- ◆ 审美能力同样是人类独有的特征，但很难用技术语言解释，也很难赋予机器。

当今人工智能局限—情感



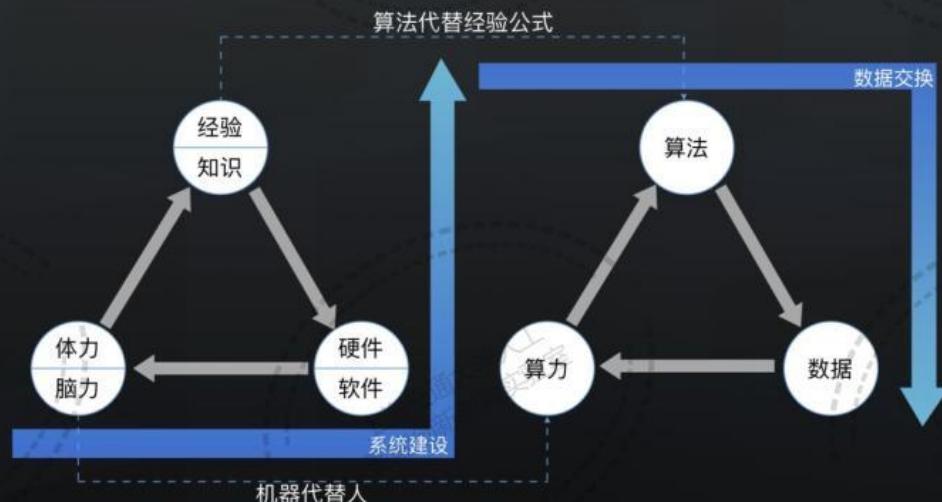
机器还完全不能理解人的哀乐、七情六欲、信任与尊重……

第六节 迎接人工智能的未来



人工智能的未来

「算力+数据+算法」对于传统「人+系统+流程」的替代



人工智能的未来 让人欢喜让人忧

上海交通大学
人工智能创新教育实验室

从大数据到小数据

上海交通大学
人工智能创新教育实验室

深度学习的训练过程需要大量经过人工标注的数据
大规模数据的标注是一件费时费力的工作

从大模型到小模型

目前深度学习的模型都非常大在移动设备上使用就会非常麻烦

从感知认知到理解决策

在感知和认知的部分，比如视觉、听觉，机器在一定限定条件下已经能够做到足够好了。而在一些动态的任务中，仍存在着很大的挑战

名言：只要一直跟着心前行，总会到达成功的终点

上海交通大学
人工智能创新教育实验室

上海交通大学
人工智能创新教育实验室

迎接人工智能的未来

上海交通大学人工
智能创新教育实验室

信息科学

计算机科学、自动化、信息工
程等学科是人工智能的基石



神经科学

探索智能、意识的人脑机理，
认识人的行为和情感，
创新脑疾病诊断与治疗



语言学

对人工智能助力大的应用领域，
让计算机识别人说的话(语音识别)，
教计算机生成人说的话(语音合成)



数理学科

数学、物理基础学科，是人工
智能的理论基础。



心理学

心理学实际上人工智能的基础理
论之一。包心理学对人工智能的
影响深入是意识研究的基础



人工智能的未来

让人欢喜让人忧

从大数据到小数据

深度学习的训练过程需要大量经过人工标注的数据
大规模数据的标注是一件费时费力的工作

从大模型到小模型

目前深度学习的模型都非常大在移动设备上使用就会非常麻烦

从感知认知到理解决策

在感知和认知的部分，比如视觉、听觉，机器在一定限定条件下已经能够做到足够好了。而在一些动态的任务中，仍存在着很大的挑战

名言：只要一直跟着心前行，总会到达成功的终点



稍后继续

THANKS FOR YOUR ATTENTION