

# 人工智能

*Artificial Intelligence*

杨育彬

[yangyubin@nju.edu.cn](mailto:yangyubin@nju.edu.cn)

<http://cs.nju.edu.cn/yangyubin>

2019年春季南京大学计算机科学与技术系  
本科三、四年级课程

# 课程简介

- 人工智能研究的起源、定义与发展，以及研究内容和应用领域
- 确定性人工智能的三大技术：知识表示、确定性推理、搜索策略
- 计算智能：神经计算、进化计算、模糊计算
- 不确定性推理
- 传统机器学习、深度学习、强化学习
- 分布式人工智能：Agent技术

# 学习目标

- 领略人工智能思想的精髓，对人工智能的**思想和方法**有比较深刻的认识，从人工智能的角度出发去思考问题，解决问题
  - 了解人工智能的发展历史，国内外人工智能相关领域的发展动态
  - 掌握人工智能的基本理论、技术及其应用方法
  - 讨论一些新的和正在研究中的人工智能方法与技术
  - 能够应用相应的人工智能技术解决实际应用问题

# 课程考核

总成绩 = 期末考试 + 期中测试 + 课堂/平时测验

( 100%        50%        30%        20% )

## 要求

1. 按时上课，不缺勤、不迟到、不早退
2. 按要求及时完成课堂/平时测验
3. 期中测试、期末考试合格

# 主要参考书

- 人工智能—复杂问题求解的结构和策略 (原书第6版)  
George F. Luger著, 史忠植 等译  
机械工业出版社
- 人工智能—一种现代的方法 (第3版)  
罗素 RUSSELL 清华大学出版社
- DeepLearning 深度学习  
Ian Goodfellow / Yoshua Bengio 著  
人民邮电出版社出版

# 信息、知识和智能

- 信息、知识和智能
- 信息：是由数据表达的客观事实
- 知识：是由智力对信息进行加工后所形成的对客观世界规律性的认识
- 智能：是指人类在认识客观世界中，由思维过程和脑力活动所表现出的综合能力
- 三者之间的关系
  - 信息：是形成知识的原料，是智能的加工对象
  - 知识：是信息的关联，是由智能加工后的产品
  - 智能：是信息到知识的一个加工器

# 第1章 人工智能概述

- AI的定义及其研究目标
- AI的产生与发展
- AI研究的基本内容
- AI的主要研究和应用领域
- AI研究的不同学派
- AI近期发展分析

# 1.1 AI的定义及其研究目标

## AI的定义

- 形式化定义
- **AI**是研究理解和模拟人类智能、智能行为及规律的一门学科
- 一般解释
- 人工智能就是用人工的方法在机器（计算机）上实现的智能，或称机器智能
- 无形式化定义的理由
  - 人工智能的严格定义依赖于对智能的定义
  - 但什么是智能？还无严格定义
  - 因此，应先对人类的自然智能进行讨论

# 何谓智能（自然智能）

- 自然智能
- 指人类和一些动物所具有的智力和行为能力
- 人类的自然智能（简称智能）
- 指人类在认识客观世界中，由思维过程和脑力活动所表现出的综合能力。
- 人类大脑是如何实现智能的
  - 两大难题之一：宇宙起源、人脑奥秘
  - 对人脑奥秘知之甚少，有待于进一步认识
- 对人脑奥秘知道什么
  - 结构： $10^{11-12}$ 量级的神经元，分布并行
  - 功能：记忆、思维、观察、分析 等

# 认识智能的观点

- 思维理论
- 智能来源于思维活动，智能的核心是思维
- 知识 = 思维的产物
- 知识阈值理论
  - 智能取决于知识的数量及其可运用程度。
- 进化理论
- 智能取决于感知和行为，取决于对外界复杂环境的适应
- 智能不需要知识、不需要表示、不需要推理，智能可由逐步进化来实现。（emergent）
- 观点不一致，从层次结构上再来认识一下

# 智能的层次结构

- 高层智能
- 以大脑皮层（抑制中枢）为主，主要完成记忆、思维等活动。
- 中层智能
- 以丘脑（感觉中枢）为主，主要完成感知活动。
- 低层智能
- 以小脑、脊髓为主，主要完成动作反应活动。
- 不同观点在层次结构中的对应关系
  - 思维理论
  - 知识阈值理论 }
  - 进化理论
  - 问题：智能到底包含哪些能力？>>技能

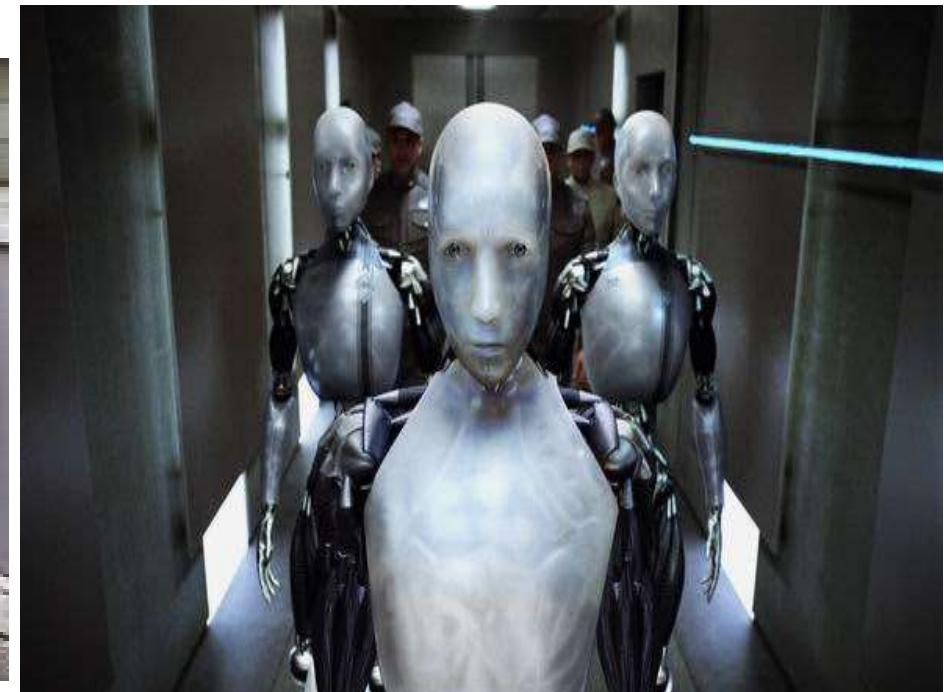
# 智能包含的能力（一）

- 感知能力
  - 感知--动作方式：对简单、紧急信息
  - 感知--思维--动作方式：对复杂信息
- 记忆和思维能力
  - 记忆：对感知到的外界信息和由思维产生的内部知识的存储过程
  - 思维：对已存储信息或知识的本质属性、内部知识的认识过程
  - 思维方式：
    - 抽象思维（逻辑思维）：根据逻辑规则对信息和知识进行处理的理性思维方式。例如，逻辑推理等
    - 形象思维（直感思维）：基于形象概念，根据感性形象认识材料对客观现象进行处理的一种思维方式。例如，图像、景物识别等
    - 灵感思维（顿悟思维）：是一种显意识和潜意识相互作用的思维方式。例如，因灵感而顿时开窍

# 智能包含的能力（二）

- 学习和自适应能力
  - 学习：是一个具有特定目的的知识获取过程是人的一种本能。不同人的学习方法、能力不同
  - 自适应：是一种通过自我调节适应外界环境的过程是人的一种本能。不同人的适应能力不同
- 行为能力
  - 含义：是人们对感知到的外界信息作出动作反应的能力
  - 信息来源：由感知直接获得的外界信息经过思维加工后的信息
  - 实现过程：通过脊髓来控制由语言、表情、体姿等来实现

# 什么是AI？



# 何为人工智能？（一）

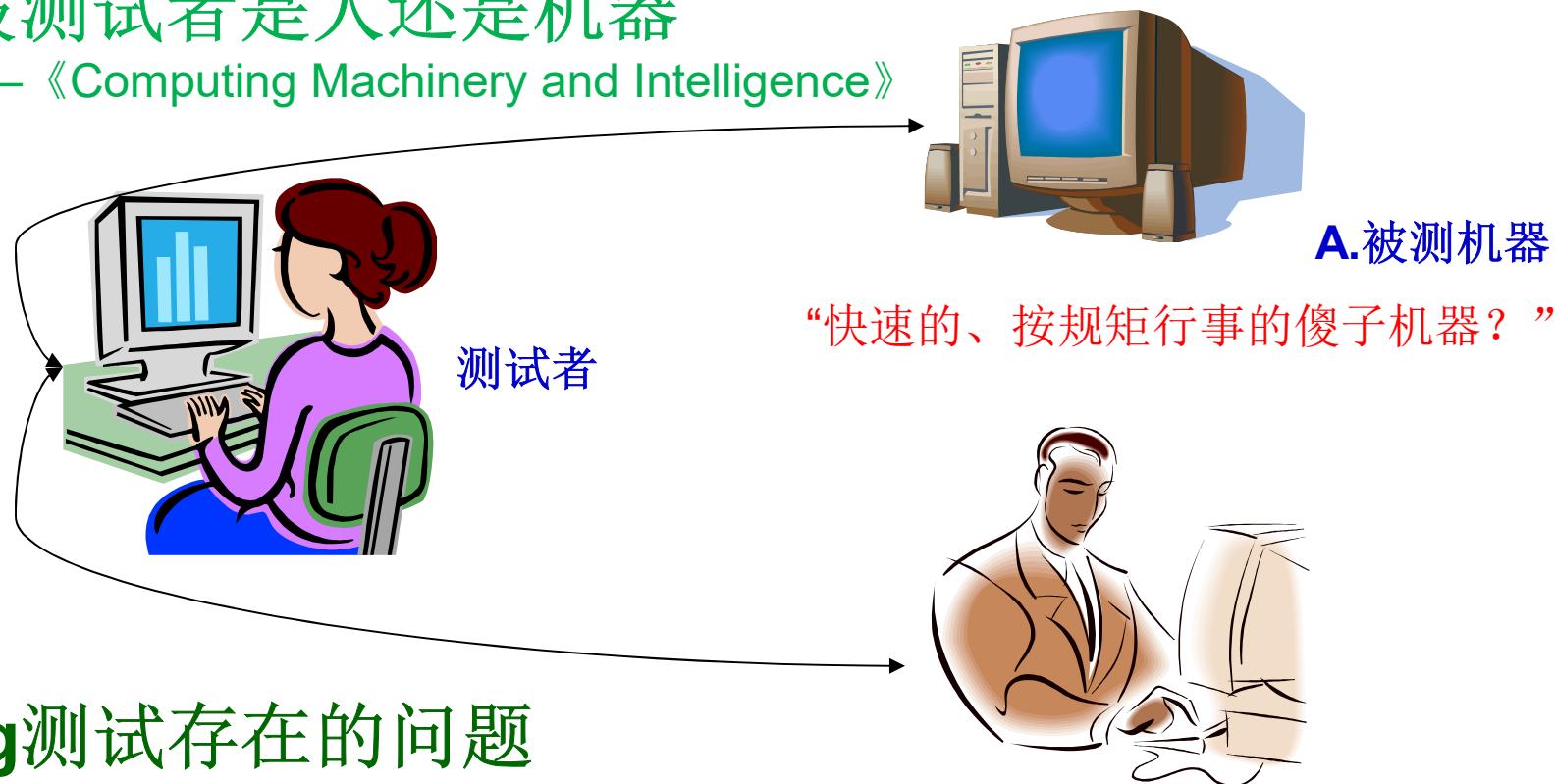
- 综合各种不同观点，可从能力和学科两个方面讨论
- 能力方面
- 人工智能就是用人工的方法在机器（计算机）上实现的智能，或称机器智能
- 计算智能 -> 感知智能 -> 认知智能
- 学科方面
- 是一门研究如何构造智能机器或智能系统，以模拟、延伸和扩展人类智能的学科

# 何为人工智能？（二）

## ----Turing测试

判断标准：超过 30% 的测试者不能确定出被测试者是人还是机器

——《Computing Machinery and Intelligence》

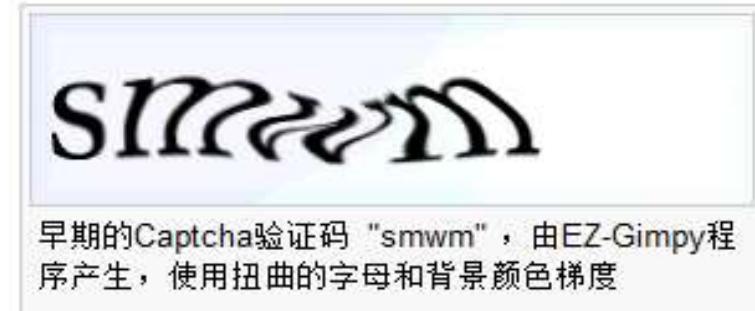


## Turing测试存在的问题

- 仅反映了结果的比较，无涉及思维过程
- 受人为因素干扰太多，严重依赖于裁判者和被测试者的主观判断
- 无法定量分析智能系统的智力水平

# CAPTCHA

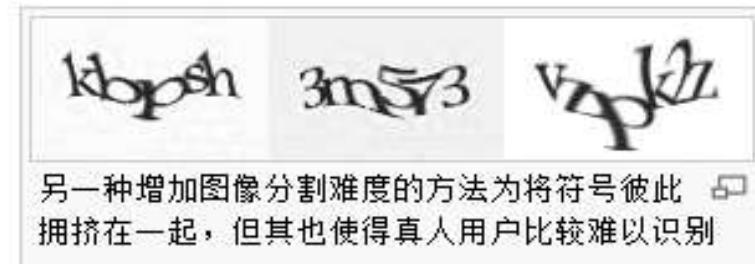
- Completely Automated Public Turing test to tell Computers and Humans Apart
- 验证码，区分用户是计算机和人的公共全自动程序
- 由计算机来考人类，而不是标准图灵测试中那样由人类来考计算机，人们有时称**CAPTCHA**是一种反向图灵测试



早期的Captcha验证码 "smwm"，由EZ-Gimpy程序产生，使用扭曲的字母和背景颜色梯度



一种更现代的CAPTCHA，其不使用扭曲的背景及字母，而是增加一条曲线来使得图像分割（segmentation）更困难。



另一种增加图像分割难度的方法为将符号彼此拥挤在一起，但这也使得真人用户比较难以识别

# 人工智能的研究目标

- 远期目标
  - 揭示人类智能的根本机理，用智能机器去模拟、延伸和扩展人类的智能
  - 涉及到脑科学、认知科学、计算机科学、系统科学、控制论等多种学科，并依赖于它们的共同发展
- 近期目标
  - 研究如何使现有的计算机更聪明，即使它能够运用知识去处理问题，能够模拟人类的智能行为。

## 1.2 AI的产生与发展

### 60多年的发展道路

- 孕育期（1956年以前）
- 形成期（1956---1970年）
- 知识应用期（1970--- 20世纪80年代末）
- 从学派分离走向综合（20世纪80年代末到本世纪初）
- 人工智能综合学科的兴起（本世纪初以来）
- 人工智能产业酝酿形成（2010年左右）
- 人工智能产业初步爆发形成风口（近年来）

# 1.2 AI的产生与发展

## 孕育期（1956年以前）(1/2)

普罗米修斯：人类获取智慧火种的最初努力（渴望智慧会带来恶果？）

- 自远古以来，人类就有用机器代替人们脑力劳动的幻想：公元前900多年我国有歌舞机器人流传的记载；公元前850年古希腊有制造机器人帮助人们劳动的神话传说
- 亚里斯多德（Aristotle，公元前384——322）：古希腊伟大的哲学家和思想家，创立了演绎法。他提出的三段论至今仍然是演绎推理的最基本出发点。
- 莱布尼茨(G.W.Leibnitz，1646——1716)：德国数学家和哲学家把形式逻辑符号化，奠定了数理逻辑的基础（四则运算计算器）
- 图灵(A.M.Turing，1912——1954)：英国数学家，1936年创立了自动机理论，自动机理论亦称图灵机，是一个理论计算机模型。
- 莫克利(J.W.Mauchly，1907——1980)：美国数学家、电子数字计算机的先驱，与他的研究生埃克特(J.P.Eckert)合作，1946年研制成功了世界上第一台通用电子计算机ENIAC.

## 1.2 AI的产生与发展

### 孕育期（1956年以前）(2/2)

- 麦克洛奇(W.McCulloch)和皮兹(W.Pitts): 美国神经生理学家，于1943年建成了第一个神经网络模型(MP模型)。
- 维纳(N.Wiener, 1874—1956): 美国著名数学家、控制论创始人。1948年创立了控制论。控制论向人工智能的渗透，形成了行为主义学派。
- 图灵又于1950年，发表题为《计算机能思维吗？》的著名论文，明确提出了“机器能思维”的观点。
- 可见，在人工智能诞生之前，一些著名科学家就已经创立了数理逻辑、神经网络模型和控制论，并发明了通用电子数字计算机。为人工智能的诞生准备了必要的思想、理论和物质技术条件。
- 同时加深了认识：智慧的奥秘可以被纳入到科学分析中

# 1.2 AI的产生与发展

## 形成期（1956--1970年）(1/3)

- AI诞生于一次历史性的聚会(Dartmouth人工智能夏季研讨会)
- 时间：1956年夏季
- 地点：达特茅斯(Dartmouth)学院
- 目的：为使计算机变得更“聪明”，或者说使计算机具有智能
- 发起人：
  - 麦卡锡(J.McCarthy)，Dartmouth的年轻数学家、计算机专家，后为MIT教授
  - 明斯基(M.L.Minsky)，哈佛大学数学家、神经学家，后为MIT教授
  - 洛切斯特(N.Lochester)，IBM公司信息中心负责人
  - 香农(C.E.Shannon)，贝尔实验室信息部数学研究员
- 参加人：
  - 莫尔(T.more)、塞缪尔(A.L.Samuel)，IBM公司
  - 塞尔夫里奇(O.Selfridge)、索罗蒙夫(R.Solomonff)，MIT
  - 纽厄尔(A.Newell)，兰德(RAND)公司
  - 西蒙(H.A.Simon)，卡内基—梅隆大学
- 会议结果：
  - 由麦卡锡提议正式采用了“Artificial Intelligence”这一术语

# 1.2 AI的产生与发展

## 形成期（1956--1970年）(2/3)

- 心理学小组
- 1957年，纽厄尔、肖(J.Shaw)和西蒙等人的心理学小组研制了逻辑理论机(Logic Theory Machine，简称LT)的数学定理证明程序。
- 1960年研制了通用问题求解(General Problem Solving)程序。该程序当时可以解决11种不同类型的问题，如不定积分、三角函数、代数方程、猴子摘香蕉、河内梵塔、人—羊过河等。

此时机器具有的推理能力，已达到了人类的巅峰水平！

- IBM工程小组
- 1956年，塞缪尔在IBM704计算机上研制成功了具有自学习、自组织和自适应能力的西洋跳棋程序。这个程序可以从棋谱中学习，也可以在下棋过程中积累经验、提高棋艺。通过不断学习，该程序1959年击败了塞缪尔本人，1962年又击败了一个州的冠军。
- MIT小组
- 1958年，麦卡锡建立了行动规划咨询系统。
- 1960年，麦卡锡又研制了人工智能语言LISP。
- 1961年，明斯基发表了“走向人工智能的步骤”的论文，推动了人工智能的发展。

# 1.2 AI的产生与发展

## 形成期（1956--1970年）(3/3)

- 失败的预言：

60年代初，西蒙预言：10年内计算机将成为世界冠军、将证明一个未发现的数学定理、将能谱写出具有优秀作曲家水平的乐曲、大多数心理学理论将在计算机上形成。

- 挫折和教训：仅有逻辑推理能力是不够的

- 在博弈方面：塞缪尔跳棋程序的失败。
- 在定理证明方面：鲁宾逊归结法的能力有限。当用归结原理证明两个连续函数之和还是连续函数时，推了10万步也没证出结果。
- 在问题求解方面：对于不良结构，会产生组合爆炸问题。
- 在机器翻译方面：不容易，甚至会闹出笑话。“Out of sight, out of mind”(“眼不见心不烦”): 翻译成：“又瞎又疯”
- 在神经生理学方面：研究发现人脑有 $10^{11-12}$ 以上的神经元，在现有技术条件下用机器从结构上模拟人脑是根本不可能的。
- 科幻小说和电影激发了过高的期望

- 一些学者指责“人工智能研究不是骗局，也是庸人自扰”

24

- 全世界范围内人工智能研究陷入困境、落入低谷。

# 1.2 AI的产生与发展

知识应用期（1971--80年代末）(1/2)

- 以知识为中心的研究：

- 专家系统的出现，使AI研究出现新高潮。
- 1972年，费根鲍姆开始研究**MYCIN** 疾病诊断和治疗系统专家系统，并于1976年研制成功。从应用角度看，它能协助内科医生诊断细菌感染疾病，并提供最佳处方。从技术角度看，他解决了知识表示、不精确推理、搜索策略、人机联系、知识获取及专家系统基本结构等一系列重大技术问题。
- 1972年法国马赛大学**Clomerauer**发明**Prolog**逻辑程序设计语言
- 1976年，斯坦福大学的杜达(**R.D.Duda**)等人开始研制地质勘探专家系统**PROSPECTOR**
- 1970年**Artificial Intelligence**国际期刊创刊
- 1977年**Feigenbaum**提出知识工程概念，专家系统广泛应用

将人工智能引向了实用化

# 1.2 AI的产生与发展

知识应用期（1971--80年代末）(2/2)

- 新的问题：
- 专家系统本身所存在的应用领域狭窄、缺乏常识性知识、知识获取困难、推理方法单一、没有分布式功能、不能访问现存数据库等问题被逐渐暴露出来。



机器自动去学知识

“机器学习”阶段（90年代至今）

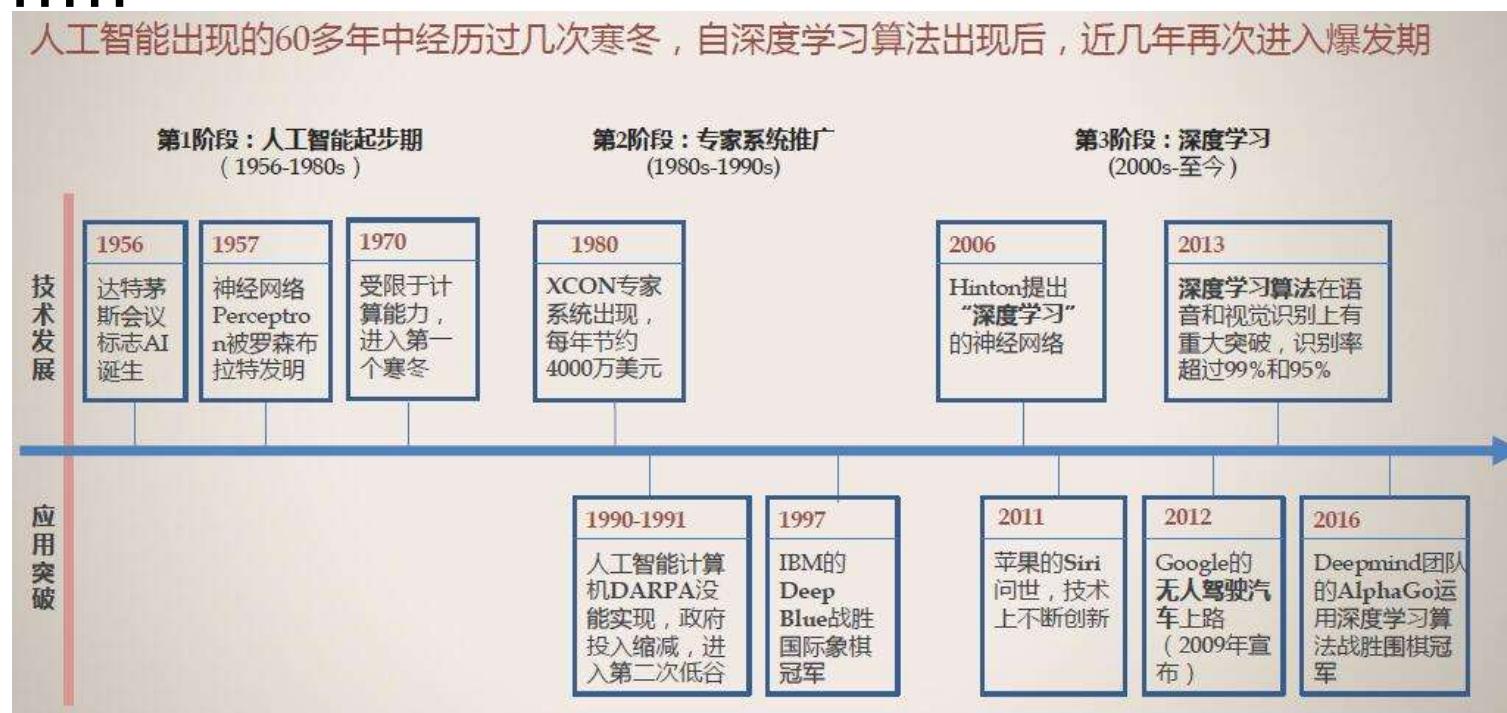
目的：解决知识获取瓶颈。利用经验来改善系统自身的性能。

经验在计算机中以数据的形式存在，主要研究的是如何利用计算机对数据进行分析的理论和方法。

# 1.2 AI的产生与发展

AI工业化, 新技术层出不穷 (20世纪80年代到本世纪初)

- 英国的Alvey Report建议恢复投资AI
- 1986年: 人工神经网络
- 1990年: 分布式人工智能和Agent技术, 机器学习和数据挖掘
- 



# 1.2 AI的产生与发展

- 人工智能研究形成了三大学派：
- 随着人工神经网络的再度兴起和布鲁克（R.A.Brooks）的机器虫的出现，人工智能研究形成了符号主义、连接主义和行为主义三大学派。
  - 符号主义
    - 功能模拟
    - 构造能够模拟大脑功能的智能系统。
  - 连接主义
    - 结构模拟
    - 构造模拟大脑结构的神经网络系统。
  - 行为主义
    - 行为模拟
    - 构造具有进化能力的智能系统。

# 1.2 AI的产生与发展

从学派分立到综合（20世纪80年代到本世纪初）

- 符号主义学派

是指基于符号运算的人工智能学派，他们认为知识可以用符号来表示，认知可以通过符号运算来实现。例如，专家系统等。

- 连接主义学派

是指神经网络学派，人的思维基元是神经元，而不是符号处理过程。罗森布拉特发明感知机，Rumelhart提出反向传播算法(BP)算法，1987年，首届国际人工神经网络学术大会在美国的圣迭戈(San-Diego)举行，掀起了人工神经网络的第二次高潮。进入21世纪后，连接主义卷土重来，提出了“深度学习”的概念。

- 行为主义学派

- 是指进化主义学派，在行为模拟方面，麻省理工学院的布鲁克教授1991年研制成功了能在未知的动态环境中漫游的有6条腿的机器虫。

- 符号主义学派（逻辑主义、心理学派）
  - 主要观点：AI起源于数理逻辑，人类认知的基元是符号，认知过程是符号表示上的一种运算
  - 代表性成果：厄尔和西蒙等人研制的称为逻辑理论机的数学定理证明程序LT
  - 代表人物：纽厄尔、肖、西蒙和尼尔逊(Nilsson)等
- 
- 连接主义学派（仿生学派或心理学派）
  - 主要观点：AI起源于仿生学，特别是人脑模型，人类认知的基元是神经元，认知过程是神经元的连接活动过程
  - 代表性成果：由麦克洛奇和皮兹创立的脑模型，即MP模型
  - 代表人物：麦克洛奇和皮兹
- 
- 行为主义学派（进化主义、控制论学派）
  - 主要观点：AI起源于控制论，智能取决于感知和行为，取决于对外界复杂环境的适应，而不是推理。
  - 代表性成果：机器虫
  - 代表人物： Brooks教授

# 不同学派的理论之争

- 符号主义
- 智能的基础是知识，其核心是知识表示和知识推理
- 知识可用符号表示，也可用符号进行推理，因而可以建立基于知识的人类智能和机器智能的统一的理论体系。
- 连接主义
- 思维的基元是神经元，而不是符号；
- 思维过程是神经元的联结活动过程，而不是符号运算过程
- 反对符号主义关于物理符号系统的假设。
- 行为主义
- 智能取决于感知和行动，提出了智能行为的“感知—动作”模型
- 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那样逐步进化

融合是一种必然趋势

# 1.2 AI的产生与发展

人工智能产业酝酿成形

- 人工智能产业：
  - 大数据的不断丰富
  - 计算速度的日益提高
  - 人机交互的重大突破
  - 深度学习的应用取得突破性进展
  - 政府及产业资本的全力支持与投入
- 美国政府：
  - 《为了人工智能的未来做好准备》
  - 《美国国家人工智能研究与发展战略规划》
- 2018年3月5日，中国《政府工作报告》中再次提到：“加强新一代人工智能研发应用，发展智能产业，拓展智能生活。” 32

# 1.2 AI的产生与发展

## 人工智能产业酝酿成形

截至2016年年初，全球已有957家人工智能公司，其中美国公司就有499家



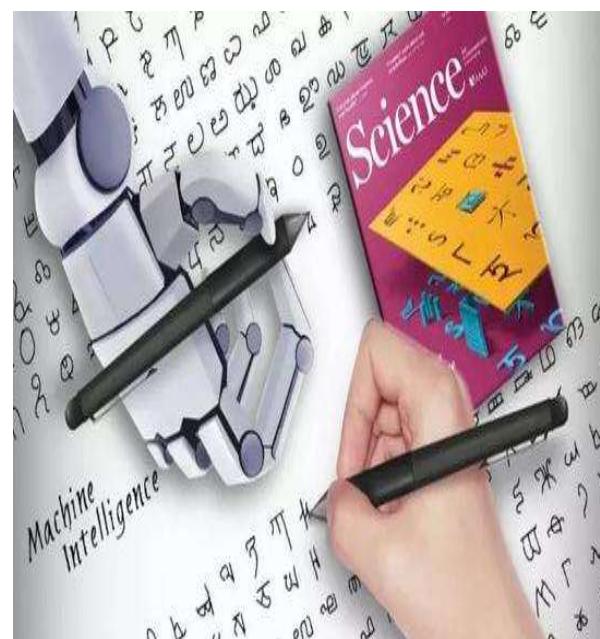
# 发展历程：两起两落，正经历第三次高潮

发展历史：已发展60年，历经两起两落；  
发展现状：2013年起，迎来第三次高潮。

## 人工智能发展历程



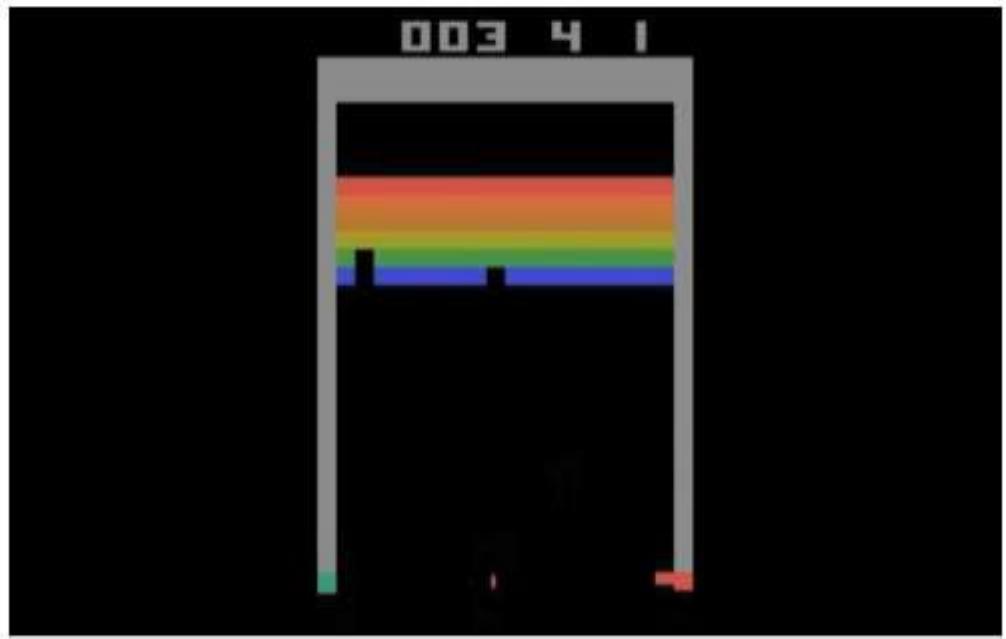
# AI成功应用案例



# DeepMind: 靠不断试错学习最后成为 击败人类专业玩家的游戏AI



Video 2: Demonstration of Learning Progress in the Game Breakout



## 1.2 AI的产生与发展

人工智能产业初步爆发形成风口

- AlphaGo、Master、AlphaZero
- DeepStack、Libratus
- 皮肤癌诊断、Waston医疗应用
- Show&Tell、LipNet、WaveNet
- AI Partnership
- AI 2020竞赛

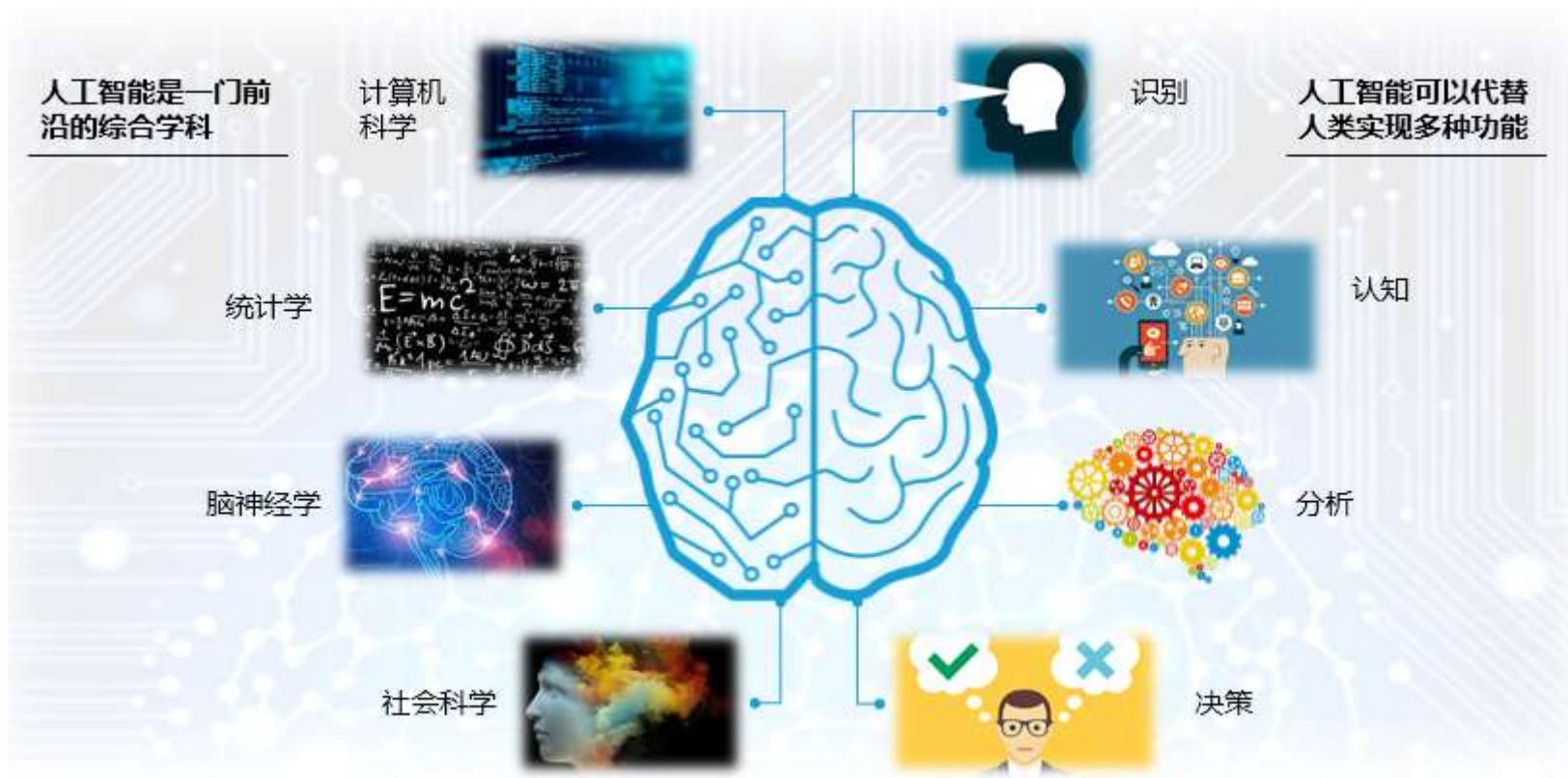


## 1.3 AI研究的基本内容

- 人工智能的学科位置
- 与脑科学和认知科学的交叉研究

# AI 的 学 科 位 置

- 新兴学科，是自然科学与社会科学的交叉学科
- AI的交叉包括：逻辑、思维、生理、心理、计算机、电子、语言、自动化、光、声等
- AI的核心是思维与智能，构成了自己独特的学科体系
- AI的基础学科包括：数学（离散、模糊）、思维科学（认知心理、逻辑思维学、形象思维学）和计算机（硬件、软件）等



## 1.4 AI的主要研究与应用领域

- 采用了基于智能本质和作用的划分方法，即从感知、思维、行为、学习、计算智能、分布智能、智能机器、智能系统、智能应用等方面来讨论。

## 1.4.1 机器思维

- 机器思维：就是让计算机模仿和实现人的思维能力，以对感知到的外界信息和自己产生的内部信息进行思维性加工。
- 包括：推理、搜索、规划等方面的研究。

## 1.4.2 机器感知

- 所谓机器感知，就是要让计算机具有类似于人的感知能力，如视觉、听觉、触觉、嗅觉、味觉。

## 1.4.2 机器感知

### 计算机视觉

- **概念：**用计算机来实现或模拟人类的视觉功能，其主要研究目标是使计算机具有通过二维图像认知三维环境信息的能力。
- **重要性：**在人类感知到的外界信息中，有**80%**以上是通过视觉得到的。
- **视觉系统：**人类视觉系统的功能是通过眼睛与大脑共同实现的。人们视野中的物体在可见光的照射下，先在眼睛的视网膜上形成图像，然后由感光细胞转换成神经脉冲信号，再经神经纤维传入大脑皮层，最后由大脑皮层对其进行处理与理解。
- **视觉，**不仅仅指对光信号的感受，它包括了对视觉信息的获取、传输、处理、存储与理解的全过程。

## 1.4.2 机器感知

### 模式识别

- 模式识别的概念
  - 是指让计算机能够对给定的事物进行鉴别，并把它归入与其相同或相似的模式中。
  - 被鉴别的事物可以是物理的、化学的、生理的，也可以是文字、图像、声音等。
- 模式识别的一般过程：
  - (1) 采集待识别事物的模式信息；
  - (2) 对其进行各种变换和预处理，从中抽出有意义的特征或基元，得到待识别事物的模式；
  - (3) 与机器中原有的各种标准模式进行比较，完成对待识别事物的分类识别；
  - (4) 输出识别结果。

指纹识别，人脸识别  
虹膜识别，掌纹识别...



## 1.4.2 机器感知

### 自然语言处理 Natural Language Processing

- 自然语言处理(**NLP**)主要包括

1. 机器翻译

把一种自然语言翻译成另外一种自然语言

2. 自然语言理解

概念：主要研究如何使计算机能够理解和生成自然语言。

理解的语言类型：声音语言、书面语言。

主要步骤：语音分析、词法分析、句法分析、语义分析。

- 自然语言理解的意义

人工智能如果不能用自然语言作为其知识表示基础，建立起不确定人工智能的理论和方法，人工智能也就永远实现不了跨越的梦想。

## 1.4.3 机器行为

- 机器行为就是让计算机能够具有像人那样地行动和表达能力，如走、跑、拿、说、唱、写画等。
- 机器行为则可看作智能系统的输出部分。
- 主要讨论：智能控制、智能检索和智能机器人等。

## 1.4.3 机器行为

### 智能控制

- 智能控制的概念：是指那种无需或需要尽可能少的人工干预就能独立的驱动智能机器实现其目标的控制过程。它是人工智能技术与传统自动控制技术相结合的产物。
- 智能控制系统：是指那种能够实现某种控制任务，具有自学习、自适应和自组织功能的智能系统。从结构上，它由传 感器、感知信息处理模块、认知模块、规划和控制模块、执行器和通信接口模块等主要部件所组成。
- 智能控制的主要应用领域：智能机器人系统、计算机集成 制造系统（CIMS）、复杂工业过程的控制系统、航空航天控 制系统、社会经济管理系统、交通运输系统、环保及能源系 统等。

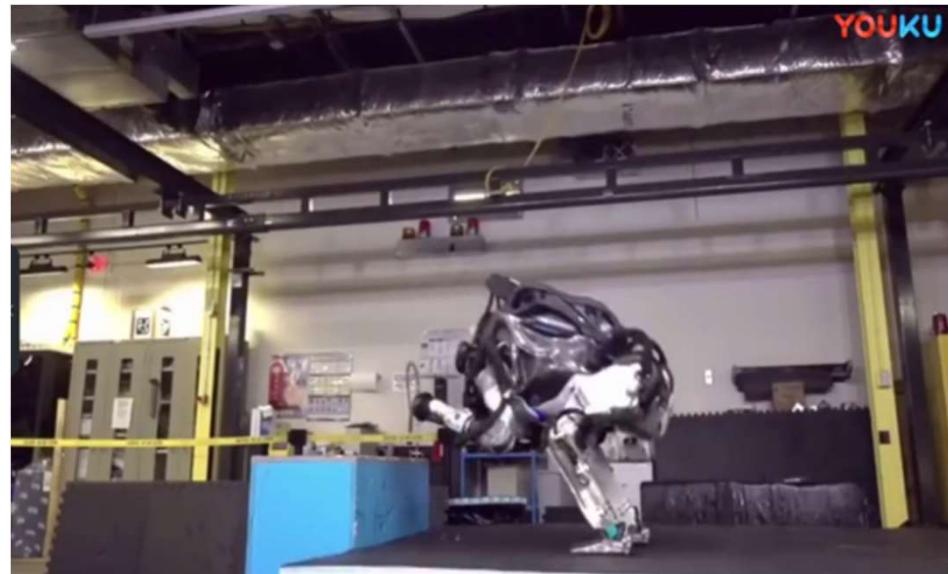
# 波士顿动力(Boston Dynamics)



机器狗SpotMini 自主打开房门



机器人Atlas 轻松后空翻



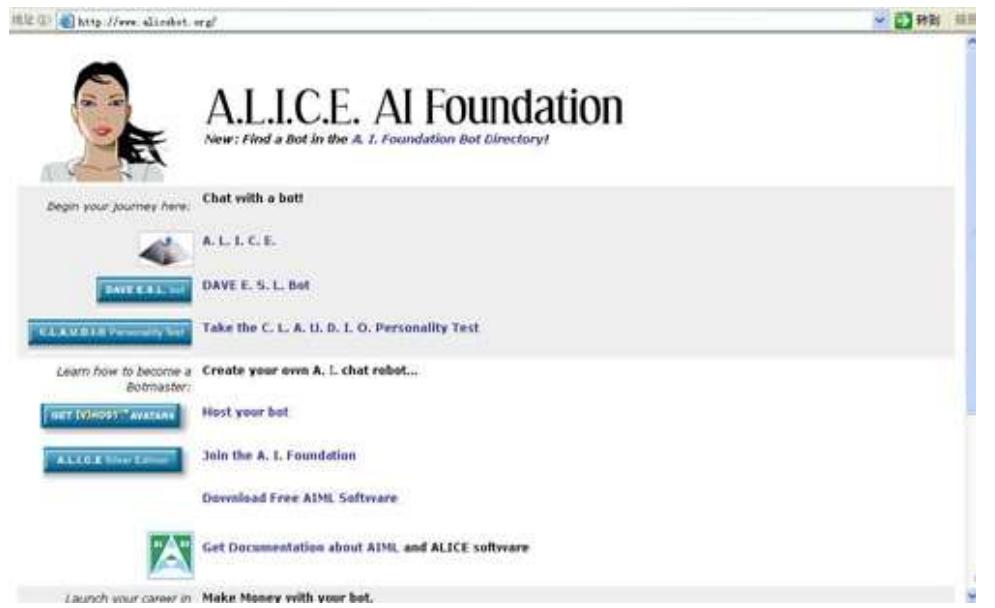
## 1.4.3 机器行为

### 智能检索

- 智能检索的概念：是指利用人工智能的方法从大量信息中尽快找到所需要的信息或知识。
- 智能检索的重要性：目前，在各种数据库中，尤其是互联网上存放着大量的、甚至是海量的信息或知识。面对这种信息海洋，如果还用传统的人工方式进行检索，已很不现实。
- 智能检索系统须解决的主要问题：
  - (1) 具有一定的自然语言理解能力，能理解用自然语言提出的各种询问；
  - (2) 具有一定的推理能力，能够根据已知的信息或知识，演绎出所需要的答案；
  - (3) 系统应拥有一定的常识性知识，以补充学科范围的专业知识。系统根据这些常识，将能演绎出更一般询问的一些答案。

## 1.4.3 机器行为 智能机器人

- **机器人(Robots)和机器人学：**机器人(Robots)是一种可再编程的多功能操作装置。机器人学是在电子学、人工智能、控制论、系统工程、精密机械、信息传感、仿生学、以及心理学等多种学科或技术发展的基础上形成的一种综合性技术学科。
- **机器人研究的意义：**机器人既是人工智能的研究对象，同时又是人工智能的试验场地，人工智能的所有技术几乎都可以在这个领域得到应用。
- **机器人的发展过程：**经历了遥控、程序、自适应、智能机器人、情感机器人。人工智能的主要研究对象是智能机器人和情感机器人。
- **智能机器人具有的能力：**感知能力、思维能力和行为能力的机器人。这种机器人能够主动的适应外界环境变化，并能够通过学习丰富自己的知识、提高自己的工作能力。
- **情感机器人：**是一种具有情感（爱、恨...）和情绪(喜、怒、哀、乐...)功能新一代机器人。



## Siri

Use your voice to send messages, set reminders, search for information, and more.



# chatbot

- 人工无脑 “人工无脑（じんこうむのう）”
  - chatterbot、chatbot, 聊天机器人
  - 与[强AI]产物相反，是[弱AI]的代表
    - 世界上第一个[人工无脑]系统chatterbot, 诞生于1966年. 名为ELIZA
    - 分解问话，再以问话检索事前录入好的答案。从而模拟人机通信，造成计算机与人沟通的[假象]，这也是所有[弱 AI]系统的共同特点
    - ELIZA, PARRY, Ractor, Verbot, A.L.I.C.E., ELLA等
  - [弱AI]: 仅是模拟人机数据交换，而根据简单的[表面现象]使之产生作用的程式，并没有[强AI]那种模拟生命智能的需求
    - 只是应用编码好的程序，把提出的问题再以固定的方式回应
    - 是在暂时无法实现[强AI]时，一种模拟[强AI]的方式
  - 绝大部分游戏编程中，关于[人机交互]这一部分，都是以[人工无脑]的方式来实现的。

# 孰“强”孰“弱”

- “强人工智能”一词最初是约翰·罗杰斯·希尔勒针对计算机和其它信息处理机器创造的，其定义为：“强人工智能观点认为计算机不仅是用来研究人的思维的一种工具；相反，只要运行适当的程序，计算机本身就是有思维的。”（J Searle in *Minds Brains and Programs. The Behavioral and Brain Sciences*, vol. 3, 1980）

# 孰“强”孰“弱”

- 争论要点：如果一台机器的唯一工作原理就是对编码数据进行转换，那么这台机器是不是有思维的？
- 希尔勒认为即使有机器通过了图灵测试，也不一定说明机器就真的像人一样有思维和意识。
- 弱人工智能并非和强人工智能完全对立，也就是说，即使强人工智能是可能的，弱人工智能仍然是有意义的。

# 情感计算(Affective Computing)



# 情感计算(Affective Computing)

- 让计算机具有情感能力是由美国MIT大学Minsky在1985年提出
- 美国MIT媒体实验室Picard教授提出情感计算一词Affective Computing并给出了定义：

“关于情感、情感产生以及影响情感方面的计算”

人工情感和人工心理

人工情感：人工情感(Artificial Emotion)是利用信息科学的手段对人类情感过程进行模拟、识别和理解，使机器能够产生类人情感并与人类进行自然和谐地人机交互的研究领域

人工心理：人工心理(Artificial Psychology)就是利用信息科学的手段，对人的心理活动(着重是人的情感、意志、性格、创造)的更全面再一次人工机器(计算机、模型算法等)模拟，其目的在于从心理学广义层次上研究人工情感、情感与认知、动机与情感的人工机器实现问题

# 情感计算(Affective Computing)

也存在不同的声音：

- “不可能论”

- 情感只有是真时才能起作用,模拟永远是假的

- 电脑不具有任何社会性需求, 因而不可能具有真正的情感

- “不必要论”

- 人类目前的技术手段和思维高度远没有达到要求

- “鸿沟论”

- 人脑绝不是单纯处理0和1的装置

# 对人工智能担忧的声音

- 黑洞中心存在着一个让一切已知物理定律都失效的“奇点”（singularity）
- 奇点理论(《奇点逼近》，库兹韦尔, 2005)
  - 技术的指数级发展将不可避免地导致超人类智能的出现。库茨韦尔甚至预测，21世纪20年代末，电脑就能达到人脑智力水平。2045年，奇点到来，超人类诞生。
- 麻省理工学院的机器人学教授罗德尼·布鲁克斯说：“在越过奇点后，我们和我们的世界将与现在大不一样，对什么是人、什么是机器人的定义也将变得不同。”
- 创办人之一伊利泽·尤德库斯基最担心的是，一些科技怪才会发明一种能够自我进化但却没有道德感的机器人，这将给人类带来灾难。
- 大多数批评的焦点在于，电脑是否真能实现智能化？
- 而这也正是人工智能研究的核心问题

# AI “威胁论”

在一些人眼里，人工智能对人类构成的威胁甚至超过核武器。

The screenshot shows a news article from Sina Tech. At the top left is the Sina logo with the text "sina 新浪科技". The main title of the article is "马斯克称人工智能威胁人类生存" (Elon Musk says AI poses a threat to human survival). Below the title is the publication date "2017年06月01日 22:04" and the source "新浪科技". There is also a small icon for Weibo.

短期内，最直接的威胁是人工智能将取代人类工作。马斯克称，在未来 20 年，驾驶人员的工作将被人工智能所颠覆。之后，全球 12% 至 15% 的劳动力将因为人工智能而失业。

The screenshot shows a news article from Xinhua News. At the top left is the Xinhua logo with the text "新华网 新闻". To the right, it says "新华网 > 军事 > 正文". The main title of the article is "霍金：彻底开发人工智能可能导致人类灭亡" (Stephen Hawking: Fully developing AI could lead to human extinction). Below the title is the publication date "2016年05月19日 07:58:20" and the source "来源：解放军报".

霍金表示，人工智能的威胁分短期和长期两种。短期威胁包括自动驾驶、智能性自主武器，以及隐私问题；长期担忧主要是人工智能系统失控带来的风险，如人工智能系统可能不听人类指挥。

## 1.4.4 计算智能

- 计算智能（**Computational Intelligence, CI**）是借鉴仿生学的思想，基于人们对生物体智能机理的认识，采用数值计算的方法去模拟和实现人类的智能。
- 计算智能的三大基本领域包括神经计算、进化计算、模糊计算。

## 1.4.5 机器学习

- 专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。
- 是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域。
- 机器学习有多种不同的分类方法，如果按照对人类学习的模拟方式，机器学习可分为符号学习和神经学习等
- 多领域交叉，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科

## 1.4.6 分布智能

- **分布智能的概念：** 分布智能主要研究在逻辑上或物理上分布的智能系统之间如何相互协调各自的智能行为，实现问题的并行求解。
- **分布智能的两个主要方向：**
- **分布式问题求解**主要研究如何在多个合作者之间进行任务划分和问题求解，它一般是针对某一问题去创建一个能够进行合作求解的协作群体；
- **多Agent系统**主要研究如何在一群自主的Agent之间进行智能行为的协调，它不限于单一目标，可创建一个能够共同处理单个目标或多个目标的智能群体。
- **多Agent系统的组成与工作：** 它由多个自主Agent所组成，其中的每个Agent都可以自主运行和自主交互，即当一个Agent 需要与别的Agent合作时，就通过相应的通信机制去寻找可以合作并愿意合作的Agent，以共同解决问题。

## 1.4.7 智能系统

- 智能系统可以泛指各种具有智能特征和功能的软硬件系统。
- 人工智能的表现形式，例如智能控制系统、智能制造系统、智能检索系统等。

## 1.4.8 人工生命

- 人工生命（Artificial Life）是美国洛斯·阿拉莫斯（Los Alamos）非线性研究中心克里斯·兰顿（Chris Langton），在研究“混沌边沿”的细胞自动机中于1987年提出的一个概念。
- 他认为：人工生命就是要研究能够展示人类生命特征的人工系统。即研究以非碳水化合物为基础的、具有人类生命特征的人造生命系统。
- 人工生命最关心的是生命的存在形式，即“生命之所能（Life as it could be）”。生命之所能，是人工生命研究所关心的主要问题。
  - 如果能从具体的生活中抽象出控制生命的“存在形式”，并且这种存在形式可以在另外一种物质中实现，那么就可以创造出基于不同物质的另外一种生命——人工生命。
- 人工生命的主要研究内容主要包括计算机进程、细胞自动机、人工脑和进化机器人等。
  - 其中，进化机器人不同于传统意义上的机器人，它是一种利用计算机和非有机物质构造出来的具有人类生命特征的人工生命实体。

# 1.5 人工智能近期发展分析

## 多学科交叉研究

- 人工智能理论基础研究
- 应强调与脑科学、认知科学、心理学、信息科学、生物学、逻辑学、物理学和数学等学科的交叉研究。
- 脑科学为人工智能研究提供人脑神经系统功能的本质和机理；
- 认知科学为人工智能研究提供感知、思维、学习和语言等基本原理；
- 心理学为人工智能研究提供认知、情感、意识等心理过程及联系；
- 生物学为人工智能研究提供自然界生物运行的机制；
- 逻辑学为人工智能研究提供思维规律描述的理论和方法；
- 信息科学为人工智能研究提供模拟的物质基础和技术手段；
- 数学为人工智能研究提供各种有效的计算模型和方法。

# 1.5 人工智能近期发展分析

- 高盛发布中国AI报告《**China's Rise in Artificial Intelligence**》：全面剖析**BAT**三巨头，**2017-09-06**
  - 国内AI公司包括：
    - 三巨头(**BAT**)，百度、阿里巴巴、腾讯
    - 主要AI公司，美团点评、滴滴出行、科大讯飞、海康威视等
    - AI创业公司，出门问问、Rokid 等
  - AI创造价值的4个驱动力：人才、数据、基础设施和计算能力。
  - 中国有着全球第二大AI生态系统，仅次于美国。
  - 中国在AI领域的投资额，在2016年前两个季度达到26亿美元，另外中国拥有AI领域的专利15700个，AI公司710家。

# 1.5 人工智能近期发展分析

- 斯坦福大学出品的**AI指数2018年度报告(AI Index 2018 Report)**发布：
  - 中国AI追赶速度惊人，在多个领域已经具备独特优势
  - AI学术论文方面，中美实力势均力敌。在Scopus的AI学术论文中，产出数量最多的地区是欧洲，占有28%；其次是中国，占25%；而美国则占17%。但美国AI论文的引用率保持最高。
  - AI论文的主导力方面，中美存在明显差异。由中国政府主导的AI论文增长了400%，而由企业主导的AI论文增长了73%；而在美国，2017年由企业主导的AI论文数量是中国的6.6倍。
  - AI的研究方向也出现了差异。欧洲的AI研究一直以来都较为全面，能涵盖多个领域；中国则开始侧重于在农业科学、工程和技术上的AI研究；相比之下，美国的侧重点也发生了改变，逐渐注重在人文科学，医学和健康上的AI研究。

# 1.6 人工智能当前研究热点

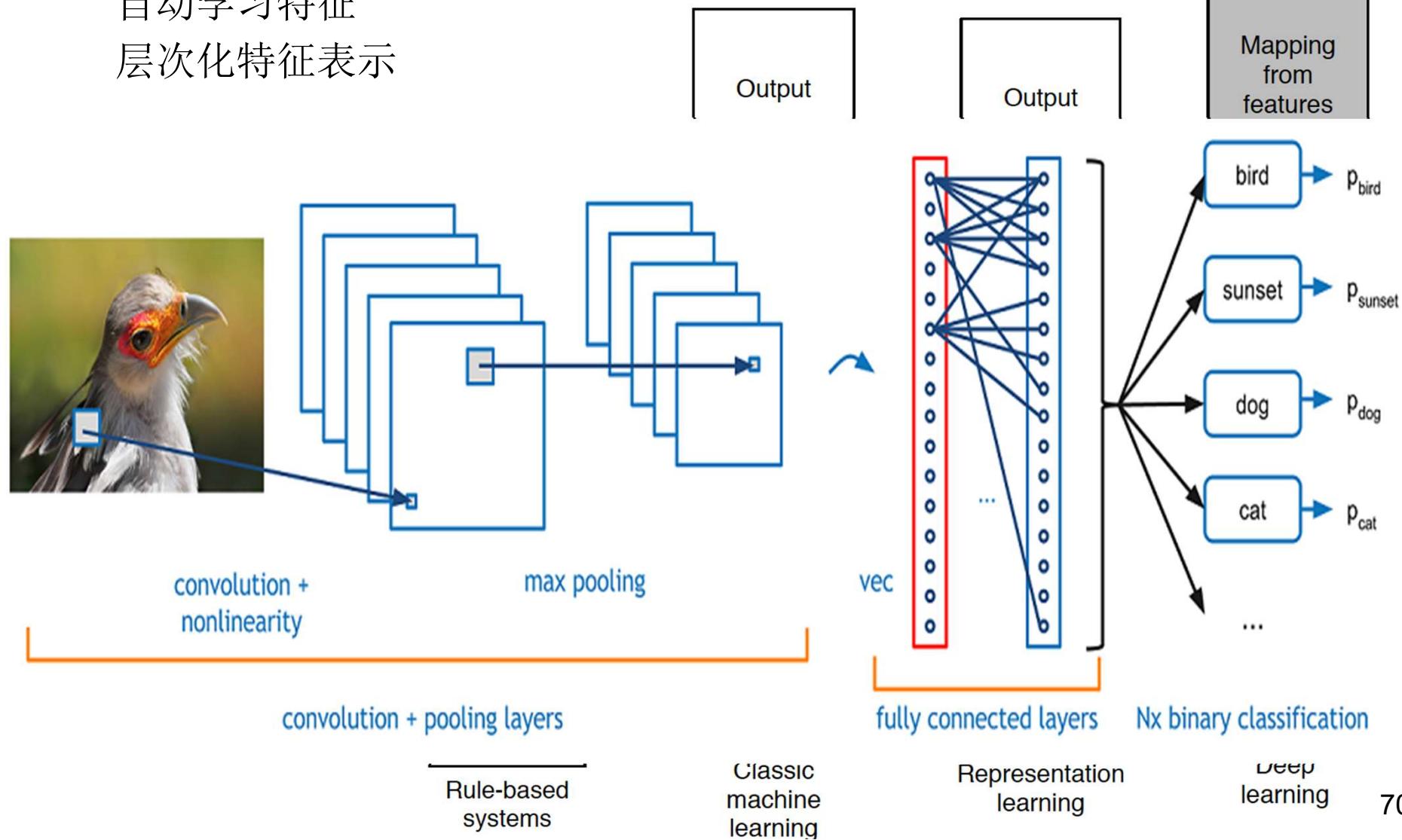
- 图像处理（深度学习）
  - 自动提取特征
- 强化学习
  - 尝试过程中学习较优策略
- 自然语言处理
  - 计算机分析和理解自然语言
- 智慧医疗
  - 制定更有效的医疗方案
- 认知计算（类脑计算）
  - 基于生物学的脑认知模式
- 智能人机交互
  - 实现设备与用户之间的自然交互

# 1.6 人工智能当前研究热点

- 深度学习

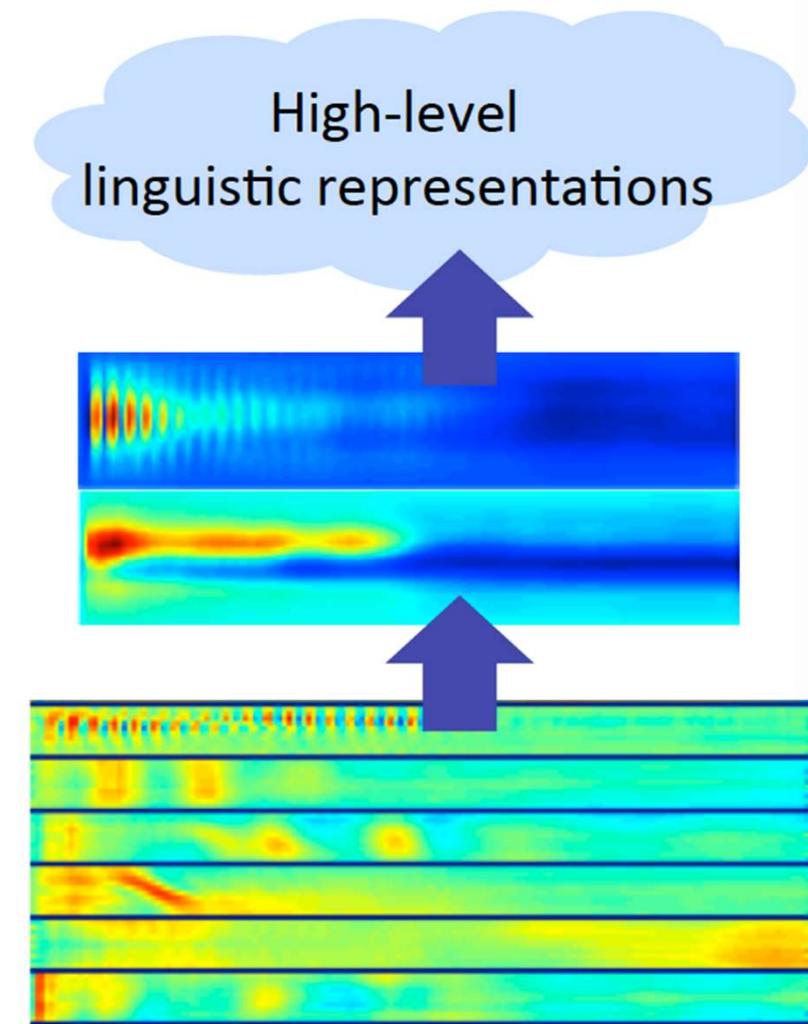
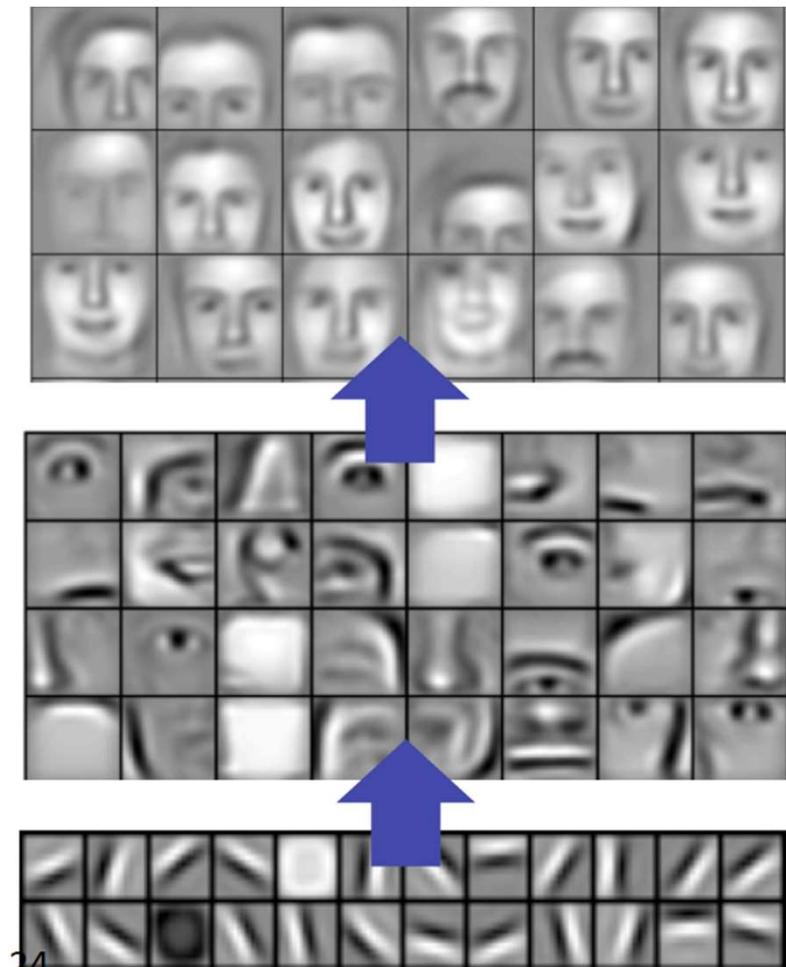
自动学习特征

层次化特征表示



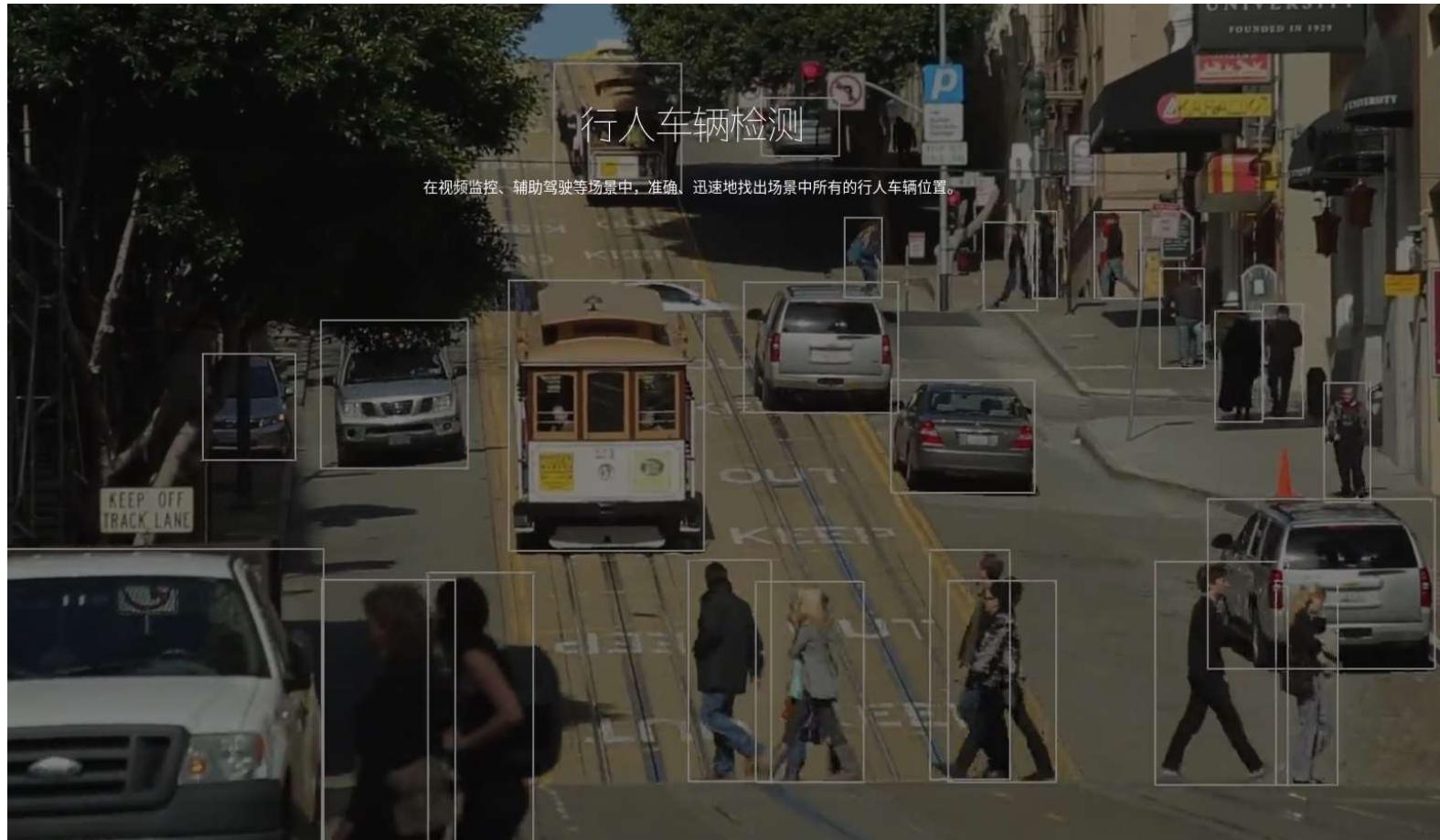
# 1.6 人工智能当前研究热点

- 层次化特征表示



# 1.6 人工智能当前研究热点

- 深度学习应用场景



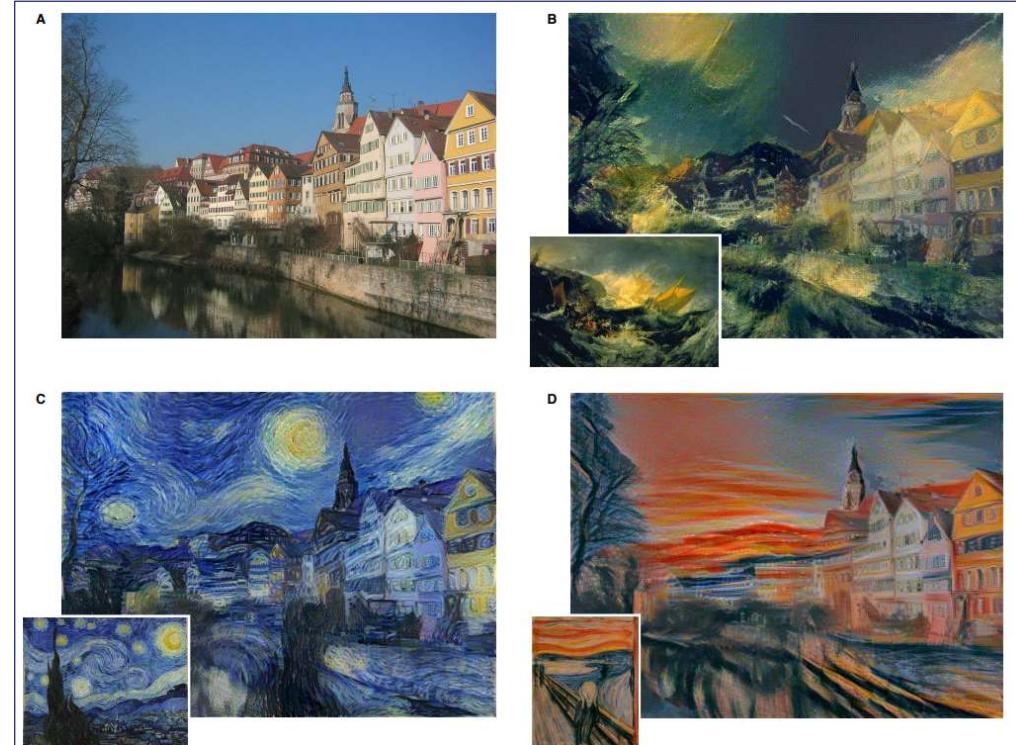
自动驾驶—行人车辆检测

# 1.6 人工智能当前研究热点

- 深度学习应用场景



自拍贴纸特效  
人脸识别检测



“拍”出名画  
图像风格迁移

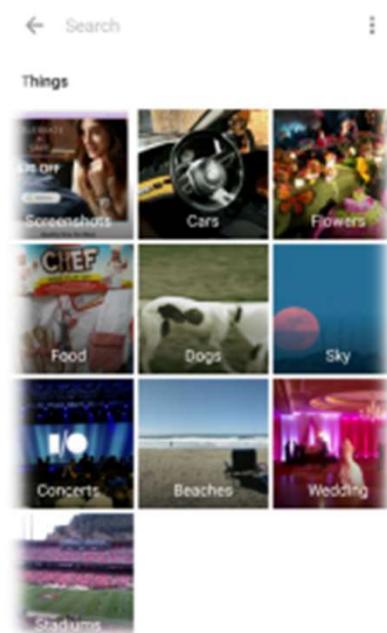
# 1.6 人工智能当前研究热点

- 计算机视觉

图像识别



图像分类



图像理解



# 1.6 人工智能当前研究热点

- 计算机视觉

自2014年由Ian J. Goodfellow等人提出GAN之后，无监督的图像生成算法成为了学术界和工业界的一大研究热点。

2018年，BigGANs的提出将图像生成任务提升到了一个全新的高度，算法可以生成几乎让人无法区分的虚拟图片。



# 1.6 人工智能当前研究热点

- ## 计算机视觉

计算机视觉任务往往需要使用大型的网络结构，耗费大量的计算资源才能得以实现。因此，通过网络压缩等一系列手段，减小这些负担也是一大研究热潮：

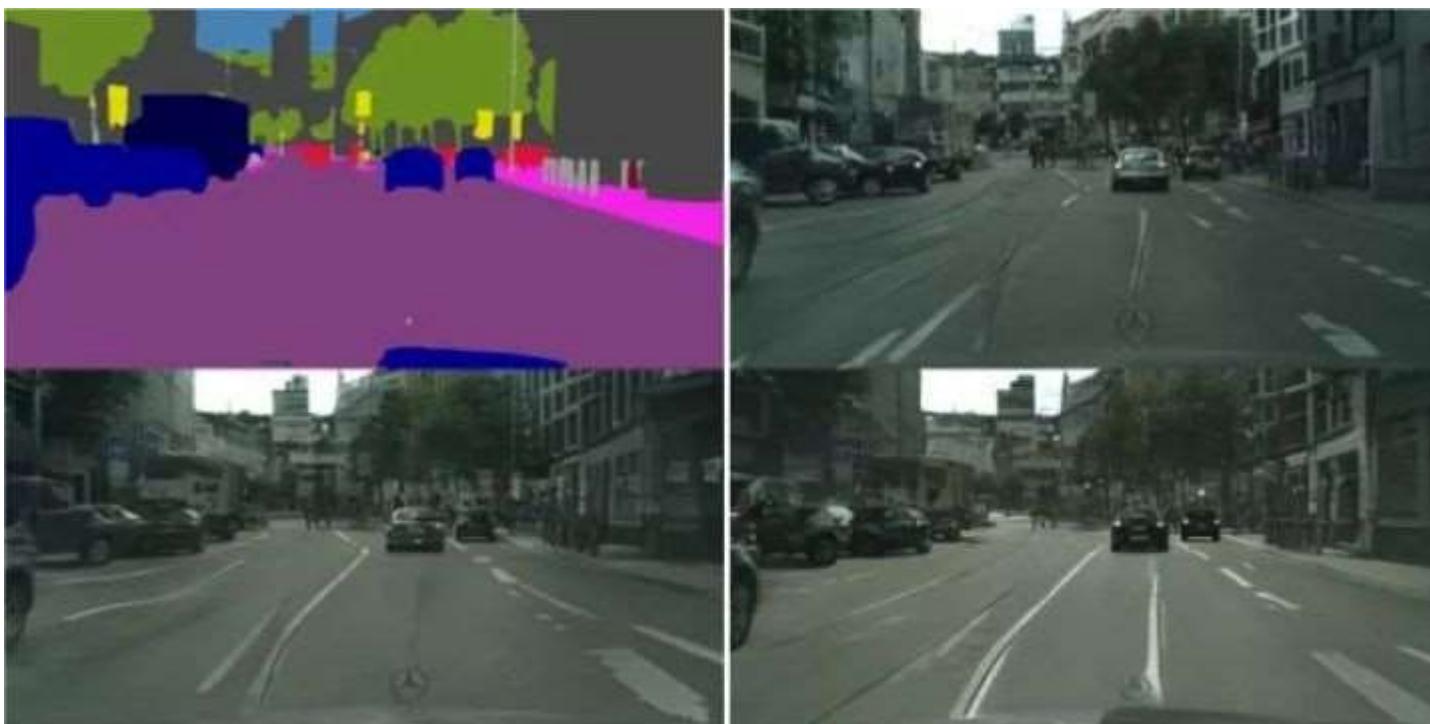
- Fast.ai在ImageNet上训练仅用18分钟便得到了93%的top5准确率，而硬件成本仅为40美元。
- Sony在2176块Tesla V100上使用ResNet-50训练ImageNet，耗时224s，top1准确率达到了75.03%。
- 腾讯也成功在6.6分钟内完成了训练

# 1.6 人工智能当前研究热点

- 计算机视觉

在过去的4-5年里，图像处理已经实现了跨越式的发展。但从静态框架转换为动态框架却比多数人想象得要困难一些。

2018年，NVIDIA的vid2vid以令人难以置信的精度实现了对视频的预测，这也标志着CV领域在从图像到视频的跨越中取得了重要的进展。



# 1.6 人工智能当前研究热点

- 深度学习网络+蒙特卡罗搜索树： AlphaGo (2016-03)



利用 3 万多幅专业棋手对局的棋谱进行训练。

- 2016年3月, AlphaGo4:1战胜职业九段李世石
- 2017年5月,乌镇围棋人机大赛, AlphaGo 以 3:0 完败人类排名第一的围棋手柯洁。

- 基于强化学习算法： AlphaGo Zero (2017-10)

采用一种新颖的强化学习算法，即自我对抗(self-play)的竞争性训练。无需提供数据、指导和规则。Zero 是迄今为止最强大，最具效率，最有通用性的 AlphaGo 版本。

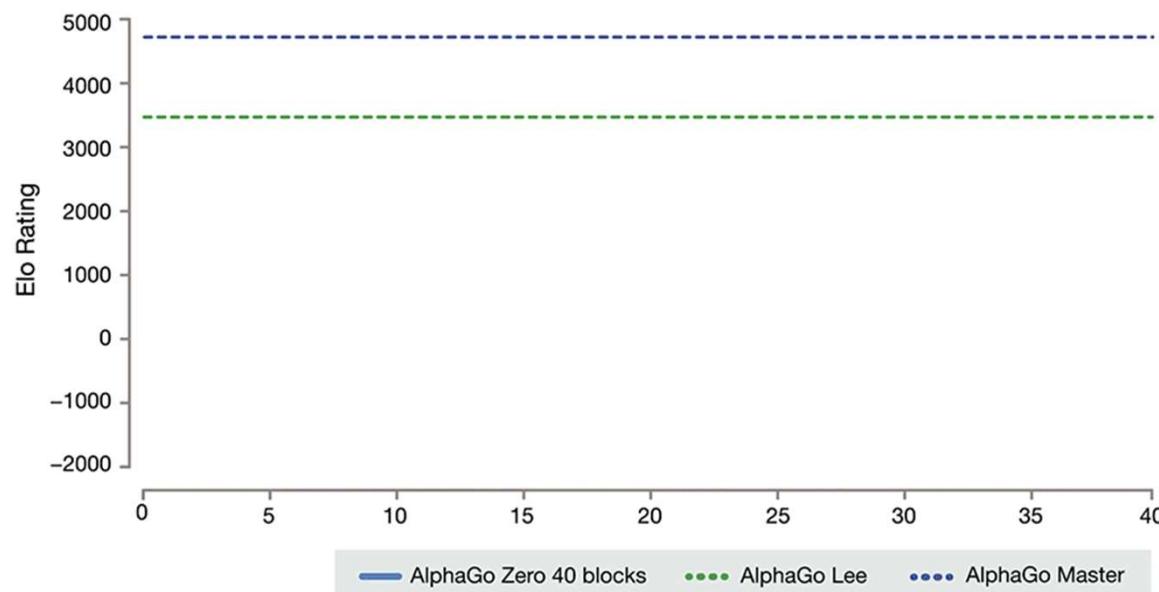
- AlphaGo Zero经过短短 3 天的自我训练，轻松地以 100 比 0 的成绩击败李世石版本的 AlphaGo
- 经过21 天训练达到了之前击败柯洁的 AlphaGo Master 的水平。

# 1.6 人工智能当前研究热点

- **AlphaGo Zero:** 从零开始学起，做自己的老师

AlphaGo Zero 会成为一个从复杂数据中进行搜索提炼的机器。其并不仅仅是为了理解围棋，它可以应用于至其它领域，比如药物的发现，蛋白质折叠，量子化学，粒子物理以及材料设计等等。”

——Demis Hassabis (DeepMind 创始人)



AlphaGo Zero 经过一定时间训练后，Elo等级分变化情况

# 1.6 人工智能当前研究热点

- OpenAI Five: 在Dota2比赛中战胜人类选手

2018年8月6日，人类和OpenAI Five展开决战。人类代表队是由半职业Dota高手组成的战队，赛制为三局两胜。经过三轮的对决，高下立判而且毫无悬念，OpenAI Five轻松赢下头两盘。经过估计，OpenAI Five的水平至少是6600以上。



# 1.6 人工智能当前研究热点

- **AlphaStar:** 在星际争霸 2中战胜人类选手

2019年1月25日，DeepMind公布了其录制的 AI 在《星际争霸 2》中与2位职业选手的比赛过程：AlphaStar 5:0 战胜职业选手TLO，5: 0 战胜 2018 年 WSC 奥斯汀站亚军 MaNa。与两位人类对手的比赛相隔约两周，AI 自学成才，经历了从与 TLO 对战时的菜鸟级别，进化到完美操作的过程。



# 1.6 人工智能当前研究热点

- 强化学习(Reinforcement Learning)

以上述例子为代表的强化学习算法目前也是人工智能领域的一大研究热点。但其应用场景绝不仅限于游戏，它在医疗、投资、控制等众多领域都有巨大的价值。大家熟知的自动机器学习(Auto ML)也是基于强化学习算法实现自动调参的功能。

尽管RL算法在各类游戏中大杀四方，但在学术领域，近一年来强化学习并未有突破性的进展。它依然存在以下问题：

- 数据效率过低，需要海量的样本才能完成训练
- 如何更好地泛化和迁移，即训练任务A，但却测试相关任务B
- 利用分层化的RL自动分解子目标，实现更高效的学习

事实上，这些都是最前沿的研究热点。此外，为了促进相关的研究，社区内开源了典型的强化学习资料。OpenAI有spinning up in Deep RL; Google有Dopamine。

# 1.6 人工智能当前研究热点

- 行业决策系统

决策系统的发展是随着棋类问题的解决而不断提升，从 80 年代西洋跳棋开始，到 90 年代的国际象棋对弈，机器的胜利都标志了科技的进步，决策系统可以在自动化、量化投资等系统上广泛应用。

游戏AI



自动化



量化投资

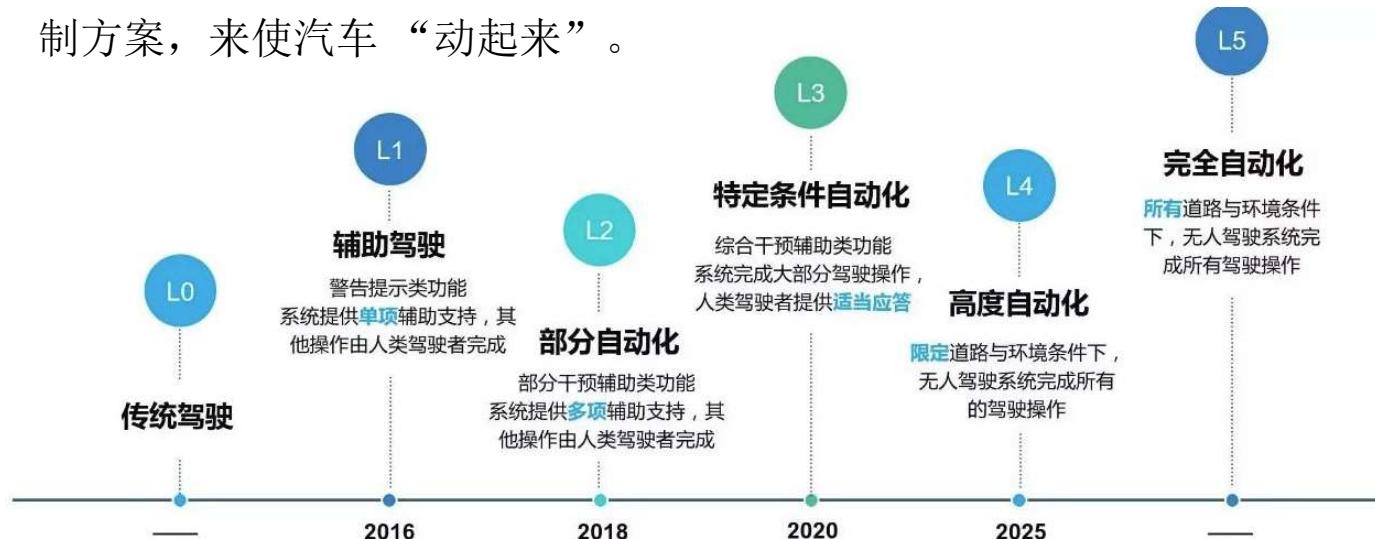


# 1.6 人工智能当前研究热点

## • 无人驾驶

### 三大核心技术

- 对外的环境感知：主要由摄像头、激光雷达、IMU 惯性传感、高精度地图等外界的数据
- 对行驶的决策：由各种算法组成的无人驾驶系统来完成，如计算机视觉、深度学习等，通过数据的训练实现越来越精准的决策。
- 对行动的控制：基于人工智能控制的直接控制方案或基于规划跟踪的间接控制方案，来使汽车“动起来”。



业界普遍认为 2020 年将实现 L3 级别无人驾驶汽车（但需要驾驶员随时能够接管）的量产。

# 1.6 人工智能当前研究热点

- 无人驾驶

全球首例自动驾驶致死案背后，“谁来担责”的法理探讨

2018年03月20日 22:21:00



国内首例特斯拉车祸致死案，终确认事发时处于“自动驾驶”状态

2018-04-20 23:00:13

特斯拉 / 自动驾驶

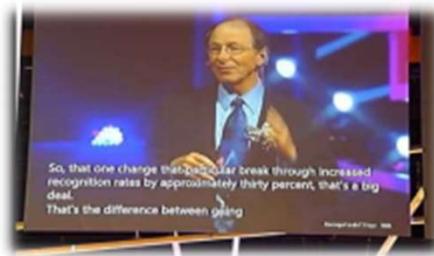
进入18年以来，自动驾驶导致事故的新闻屡见不鲜，新技术的推广收到了众多反对的声音。这不仅为AI从业者提出了新的技术难关，同样给人们带来了道德上的质问。随着AI研究的深入，我们将要面临的不仅仅是技术上的挑战。

# 1.6 人工智能当前研究热点

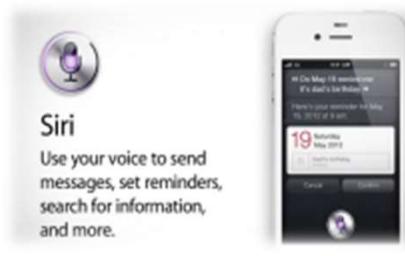
## • 自然语言处理 (Natural Language Processing)

运用计算机对自然语言进行分析和理解，更好地理解人类语言。自然语言处理技术在自动问答、机器翻译等领域得到了广泛的应用。百度度秘，微软小冰，苹果 siri 都是自然语言处理技术的结晶。

语音识别



语义理解



语音合成



# 1.6 人工智能当前研究热点

- 自然语言处理 (**Natural Language Processing**)

过去，机器翻译的结果经常会让人大吃一惊，但2018年却成为了NLP的分水岭。一个又一个显著的突破，ULMFiT，ELMo，OpenAI的Transformer和Google的BERT等，使NLP迈上了新的台阶：

- ULMFiT是年初第一个NLP迁移学习框架，几乎可以应用于任何NLP任务，从此无需从头开始训练模型。
- ELMo是语言模型嵌入的简称。ELMo使用语言模型来获取每个单词的嵌入，同时还会考虑单词适用的句子或段落上下文。
- Google提出BERT标志着NLP的新时代，它真正击败了所有竞争对手并在11个任务上取得了最优效果。自此，基于迁移学习，预训练的语言模型嵌入将无处不在。

# 1.6 人工智能当前研究热点

## • 大数据应用

可以通过你之前看到的文章，理解你所喜欢的内容而进行更精准的推荐；分析各个股票的行情，进行量化交易；分析所有的像客户的一些喜好而进行精准的营销等。机器通过一系列的数据进行判别，反馈出最适合的策略。



# 1.6 人工智能当前研究热点

- 智慧医疗

AI 应用潜力巨大：

1. 减少药物研发周期，降低研发成本；
2. 利用虚拟助手为病患提供服务；
3. 提高医疗影像诊断的准确性，优化诊疗程序。



图像识别算法应用在医疗领域，用于诊断皮肤癌。将结果与21位皮肤科医生进行对比，发现深度神经网络的诊断准确率在91%以上，与人类医生不相上下。

机器学习技术可针对其他数百万病人的病程进行的研究来预测这个病人的未来，以此帮助医疗专业人士做出更好的决策。

# 1.6 人工智能当前研究热点

- 受人脑结构的启发（并非模拟人脑）

目前的 AI 系统是在介观层面上模仿了人类，远未触及智能的核心。我们还需要向神经科学学习和合作，使AI with Understanding。

AI Systems based Data Driven vs. Human Perception

performance	Brain	AI
power consumption	20W	Hundreds of thousand times
robustness	strong	weak
generalization capacity	strong	weak
big mistake	no	yes
the size of learning samples	small	big
explain ability	yes	no

脑神经中的 Feedback Connection、Lateral Connections、Sparse Firing、Attention Mechanism、Multi-model、Memory 等机制，都值得注意和学习。

Re: 张钹院士，《AI 和神经科学》，清华大学脑与智能实验室学术研讨会

2017-12-27

90

# 1.7 人工智能展望

- AI 正式成为国家的战略发展方向

时间	行政机关	相关政策	政策内容
2015年5月	国务院	《中国制造2025》	提出“以推进智能制造为主攻方向”
2015年7月	国务院	《国务院关于积极推进“互联网+”行动的指导意见》	将人工智能作为“互联网+”的十一个重点布局领域之一
2016年3月	国务院	《国民经济和社会发展第十三个五年规划纲要（草案）》	提出要“重点突破新兴领域人工智能技术”
2016年5月	发改委	《“互联网+”人工智能三年行动实施方案》	培育发展人工智能新兴产业，推进重点领域智能产品创新，提升终端产品智能化水平
2016年7月	国务院	《“十三五”国家科技创新规划》	提出“重点发展大数据驱动的类人智能技术方法”
2017年3月	国务院	《政府工作报告》	人工智能首次被写入《政府工作报告》
2017年7月	国务院	《新一代人工智能发展规划》	提出了“三步走”的战略目标，宣布举全国之力在2030年抢占人工智能全球制高点
2017年12月	工信部	《促进新一代人工智能产业发展三年行动计划（2018-2020年）》	促进人工智能产业发展，提升制造业智能化水平，推动人工智能和实体经济深度融合

# 1.7 人工智能展望

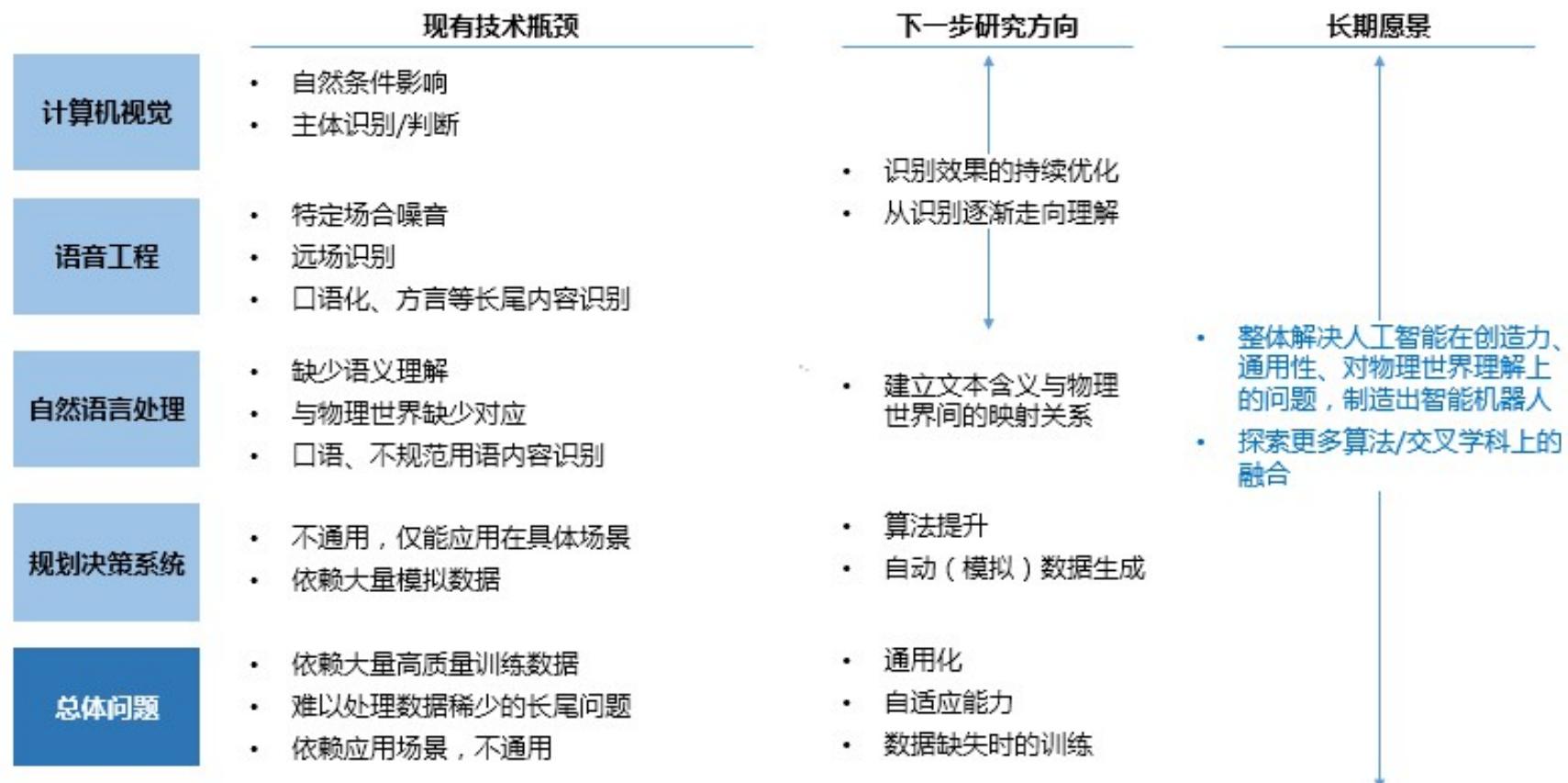
## • AI 助力六大实体经济领域



- 健康医疗，从辅诊到精准医疗      ➤ 智能零售，实体店加速升级
- 智慧城市，为城市安装智慧中枢      ➤ 智能服务业，提供个性化服务
- 智能制造，自动化的下一站      ➤ 智能教育，面向未来“自适应”教育

# 1.7 人工智能展望

## • AI 目前存在的技术瓶颈



# 1.7 人工智能展望

- **AI 本身亟待解决的问题**

AI在具有以下4个特征的任务上具有较好的表现：

1. 有充足的数据（或知识）
2. 完全信息
3. 确定性
4. 单领域

当某问题未满足其中任意一个条件时，计算机处理将变得十分困难。比如打桥牌（不完全信息博弈），无人车（不确定性，不完全信息和多领域），自然语言理解（多领域）等。

- **深度学习的可解释性和可理解性**

深度学习存在的最大问题是：不可解释和不可理解，缺乏推广能力。当环境动态变化、信息不完全、存在干扰与虚假信息时，AI的性能显著下降。

此外，由于不可理解，无法实现人机交互，无法与人类协同工作与和谐相处。人工智能在适应动态变化环境以及解决算法的鲁棒性问题，依然任重道远。

# 1.7 人工智能展望

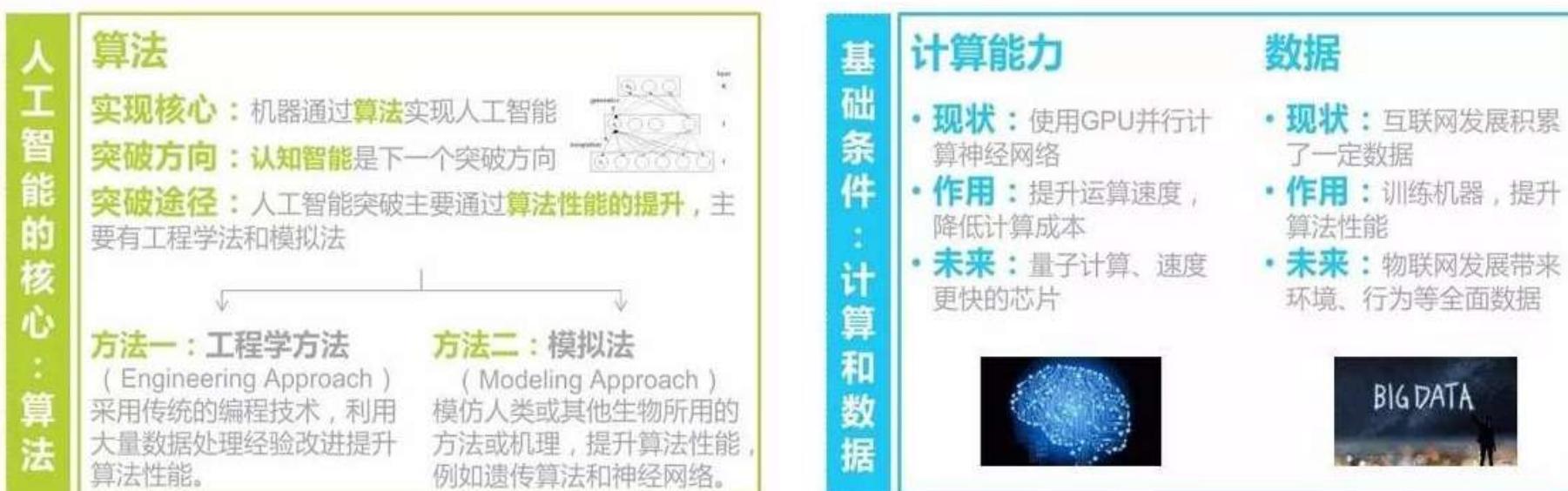
- 目前AI的要素：算法是核心 数据是基础

算法：实现人工智能的核心方法是算法

数据：训练模型，提升算法表现的基础

工程学方法和模拟法是人工智能提升的两个途径

算法发展的现状：目前认知层算法尚未突破



# 1.7 人工智能展望

- 伯克利视角：AI系统的九大挑战 《A Berkeley View of Systems Challenges for AI》（2017-11-07）

四大趋势：

1. 关键性任务的AI (**Mission-critical AI**)

突破生命极限，适应危险复杂环境的实时、稳定、安全AI系统

2. 个性化AI (**Personalized AI**)

个性化AI秘书来临！

3. 跨多组织机构的AI (**AI across organizations**)

让AI兼顾隐私和共享

4. 后摩尔定律时期的AI (**AI demands outpacing the Moore's Law**)

超越摩尔定律的AI数据量

# 1.7 人工智能展望

- 九大研究方向：

**R1:** 连续学习

**R4:** 安全的飞地

**R7:** 领域专用硬件

**R2:** 鲁棒决策

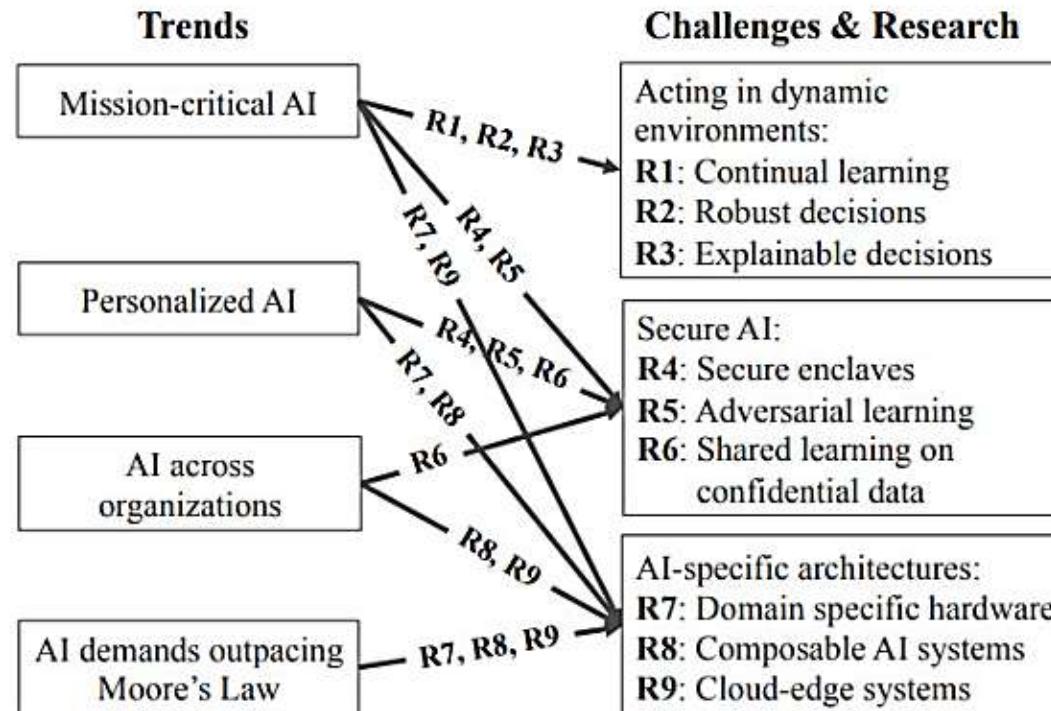
**R5:** 对抗学习

**R8:** 可组合的AI系统

**R3:** 可解释的决策

**R6:** 机要数据的共享学习

**R9:** 云边缘系统

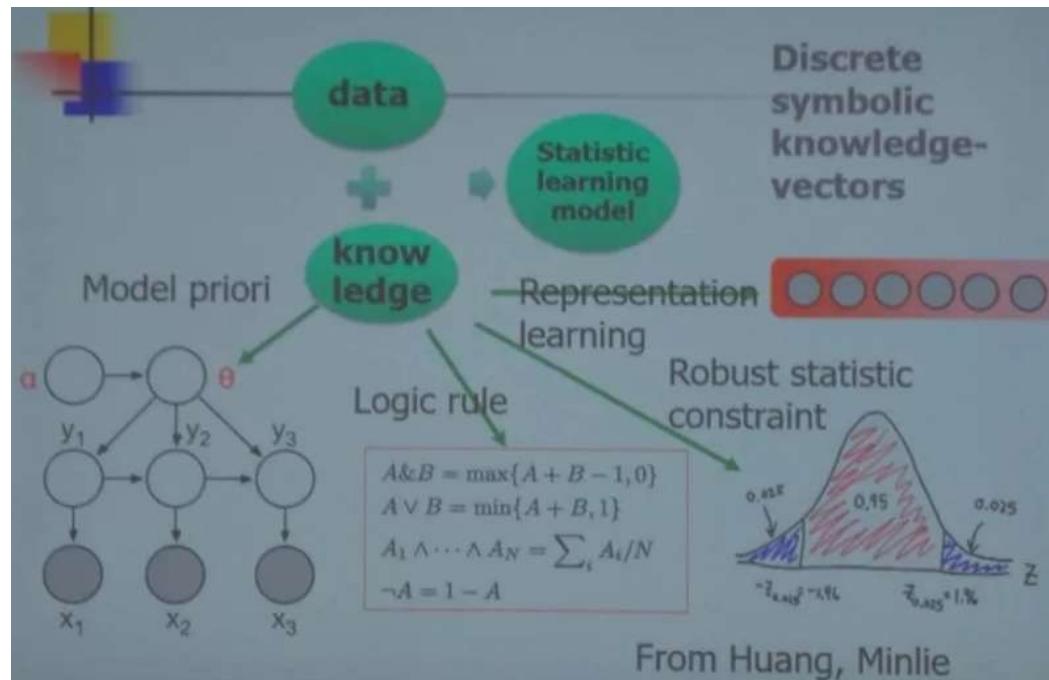


四大趋势同九大研究方向的对应关系

# 1.7 人工智能展望

- 构建同时基于知识+数据的 AI 系统

人类擅长处理知识，计算机擅长处理数据。二者结合：  
数据 + 知识 = 统计学习模型，构建出更智能的系统。



其中知识包括先验模型、逻辑规则、表示学习、强健的统计约束等。

# 1.7 人工智能展望

- 构造真正的 **Intelligence** 系统

AI **without** Understanding → AI **with** Understanding

- 在缺乏知识和数据的情况下依然能够完成任务；
- 在信息不完善（甚至缺乏信息）的情况下依然能够完成任务；
- 能够处理非确定性的任务；
- 能够处理动态任务；
- 能够处理多领域和多任务。

Re: 张钹院士，《AI 和神经科学》，清华大学脑与智能实验室学术研讨会  
2017-12-27

# 1.7 人工智能展望

## • AI的道德准则

目前，AI领域的研究蓬勃发展，但随着科技的进步，许多社会问题也接踵而至。仅去年一年，便发生了许多轰动的新闻：

- Facebook剑桥分析公司丑闻
- 超50位AI顶级学者宣布与制造AI武器的韩国大学断交
- 3100名Google员工抗议五角大楼军事项目

随着AI技术的发展和应用，随之引发的伦理道德问题引人深思，这也是我们科研工作者不得不思考的问题。