

机器人与人工智能

--人工智能

刘之涛

浙江大学智能系统与控制研究所

刘之涛

- 办公室：智能系统与控制研究所（教十八）306室
- Email: ztliu@zju.edu.cn
- 电话：87952233-8239



研究方向：

- 智能无线充电与电池管理
- 车联网与自动驾驶
- 工业大数据与图像处理

- **Russell & Norvig. 人工智能：一种现代方法（第3版），清华大学出版社，2013**
- **王万良，人工智能导论，高等教育出版社,2017年**
- **周志华，机器学习，清华大学出版社，2016.**
- **Ian Goodfellow等，深度学习，人民邮电出版社，2017.**

- 什么是人工智能
- 人工智能的基础
- 人工智能的历史
- 最新发展水平

智能

是人脑活动所表现出来的能力。

智能的基础是知识

智能的关键是思维

智能 = 知识 + 思维

人工智能概念的一般描述

顾名思义,人工智能就是人造智能,其英文表示是“Artificial Intelligence”,简称AI。“人工智能”一词目前是指用计算机模拟或实现的智能,研究如何在机器上实现人类智能。即用机器来模仿人的智能。因此人工智能又称机器智能。当然,这只是对人工智能的字面解释或一般解释。关于人工智能的科学定义,学术界目前还没有统一的认识。下面是部分学者对人工智能概念的描述,可以看做是他们各自对人工智能所下的定义。

像人一样思考的系统

——人工智能是那些与人的思维相关的活动，诸如决策、问题求解和学习等的自动化(Bellman, 1978年)。

——人工智能是一种计算机能够思维，使机器具有智力的激动人心的新尝试(Haugeland, 1985年)。

像人一样行动的系统

——人工智能是一种技艺，创造机器来执行人需要智能才能完成的功能 (Kurzweil, 1990)

——人工智能是研究如何让计算机做现阶段只有人才能做得好的事情(Rich Knight, 1991年)。

理性地思考的系统

——人工智能是那些使知觉、推理和行为成为可能的计算的研究 (Winston, 1992年)。

——人工智能是通过计算模型的使用来进行心智能力的研究 (Winston, 1992年)。

理性地行动的系统

——计算智能是对设计智能化智能体的研究(Poole等人, 1998年)。

——广义地讲, 人工智能是关于人造物的智能行为, 而智能行为包括知觉、推理、学习、交流和在复杂环境中的行为(Nilsson, 1998年)。

什么是人工智能

——Stuart Russell和Peter Norvig则把已有的一些人工智能定义分为4类：像人一样思考的系统、 像人一样行动的系统、 理性地思考的系统、 理性地行动的系统(2003年)。

	模拟人类功能的逼真度 度量	理想的智能度量
思维过程和推理	像人一样思考的系统	理性地思考的系统
强调行为	像人一样行动的系统	理性地行动的系统

可以看出，这些定义虽然都指出了人工智能的一些特征，但用它们却难以界定一台计算机是否具有智能。因为要界定机器是否具有智能，必然要涉及到什么是智能的问题,但这却是一个难以准确回答的问题。所以,尽管人们给出了关于人工智能的不少说法，但都没有**完全或严格地用智能的内涵或外延来定义人工智能**。

■ **人工智能**是研究人类智能活动的规律，构造具有一定智能的**人工系统**，研究如何让计算机去完成以往需要人的智力才能胜任的工作，也就是研究如何应用计算机的软硬件来模拟人类某些智能行为的基本理论、方法和技术。

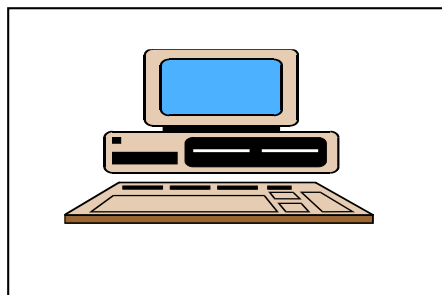
■ **智能系统**

■ **智能机器**

图灵测试方法

Turing (1950) “computing machinery and intelligence” :

- ◆ “机器可以思考吗？” → “机器可以产生智能行为吗？”
- ◆ 测试智能行为的操作型试验
- ◆ 包含AI的主要内容：知识表示，自动推理，自然语言理解，机器学习，机器视觉，机器人技术



智者



询问者

计算机需要具有的能力：

- **自然语言处理**：使之能成功的用英文交流
- **知识表示**：以存储它知道的或听到的信息
- **自动推理**：以运动存储的信息来回答问题并推出新结论
- **机器学习**：以适应新情况并检测和预测模式
- **计算机视觉**：以感知物体
- **机器人学**：以操纵和移动对象

认知模型方法：

人是怎样思考的？

两种途径测定人是怎样思考的

内省：捕捉人类自身的思维过程

心理测试：从输入输出来理解人类的思考

需要对人脑内部行为有科学的、理论上的了解

➤ 抽象进行到什么程度？

➤ 怎样验证？需要

1. 估计和测试人类主体的行为，或者

2. 直接利用神经学的数据进行辨识

Allen Newell and Herbert Simon (1961) 开发的**GPS** (General Problem Solver) 就是以模仿人类思维为目标，并宣称“一个物理符号系统具有必然和足够的方式产生一般智能行为”

理性思维和行为：

亚里士多德：正确的思维过程是怎样的？

三段论：提供了一种在已知前提正确时总能推出正确的结论一些希腊学派提出不同形式的**逻辑**：用来描述思维的**符号和规则**；

逻辑方法的障碍：

- 难以获得形式化的知识并得到逻辑符号表示所需的形式化表达
- 原则上解决问题和实际上解决问题有巨大的差别

理性行为：做正确的事情

正确的事情：给定可得到的信息，预期可最大化目标的那些事情

不一定非要经过思考，例如瞬目反射—但是思考对理性行为有所帮助

理性智能体方法：

- 智能体是某种能够行动的东西，区别于简单“程序”
 - 自主控制的操作
 - 感知环境
 - 持续能力
 - 适应变化
 - 有能力承担其它智能体的目标
 - 通过自己的行动获得最佳结果
- 做出正确的推论是理性智能体的部分功能，但不是理性的全部内容。

- 比“思维法则”方法更为通用
 - 正确的推论只是实现理性的几种可能机制之一
- 比建立在人类行为或思维基础上的方法更经得起科学发展的检验
 - 理性的标准有着清楚而普遍的定义
 - 人类的行为可以很好地适应特定的环境，而且部分地取决于未知的进化过程。

智能体 (agent) 是指能够感知和动作的实体

简单说，一个智能体就是从感知序列到动作的一个函数：

$$f : P^* \rightarrow A$$

对于任何给定的环境和任务，我们设计具有最佳性能的智能体（或主体集）

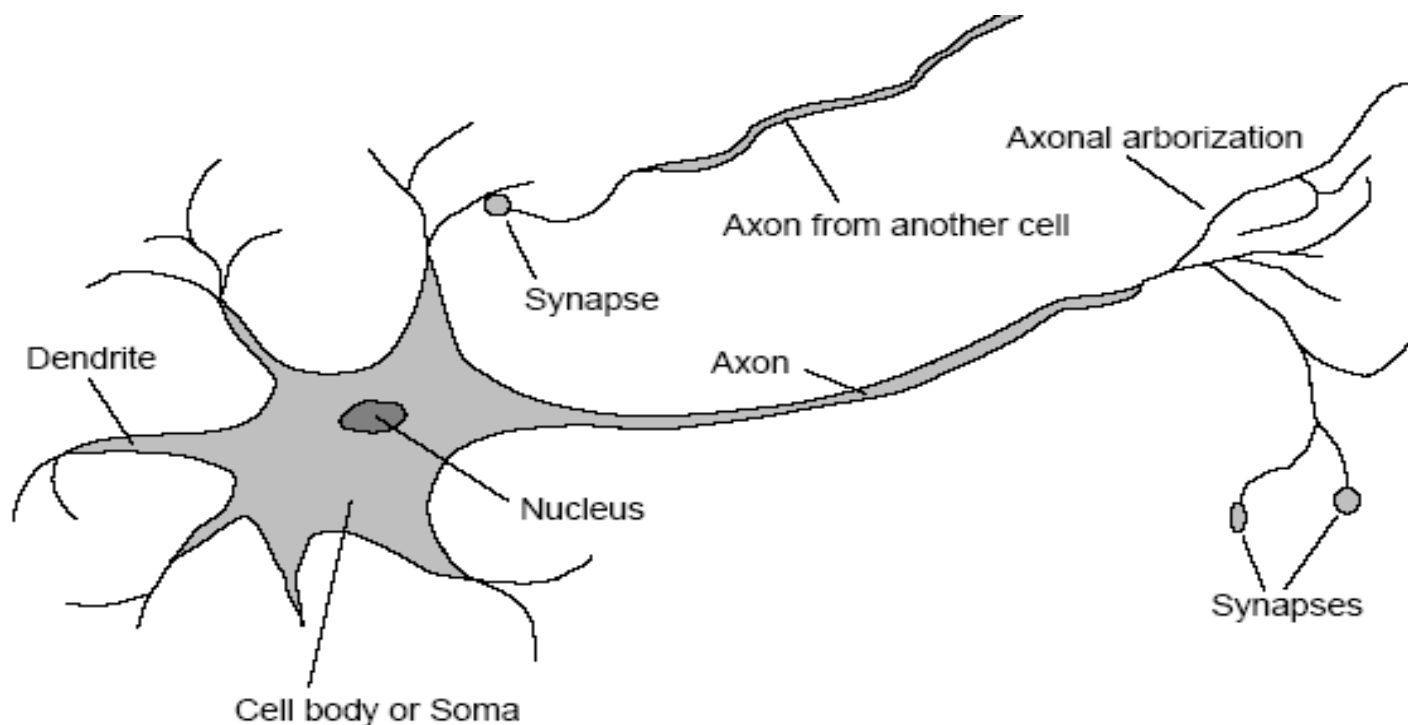
注意：计算能力有限使得最完美的主体不可实现→只能针对给定的机器资源设计最好的程序

- 哲学：标出了AI的大部分重要思想
- 数学：使AI成为一门规范科学
- 经济学：决策理论
- 神经科学：网络，并行处理...
- 心理学：认知理论
- 计算机工程：AI的“载体”
- 控制论：反馈的思想
- 语言学：知识表示、语法
-

数学：

- 什么是抽取合理结论的形式化规则？（形式逻辑）
- 逻辑和计算的极限
 - Godel不完备性定理（1931）：存在不可判断的真值语句。
 - Turing试图精确地刻画哪些函数是能够被计算的
 - NP-完全理论为认识不可操作问题提供一种方法
- 概率
 - 成为处理不确定的测量和不完备的理論的工具
 - Bayes分析形成了大多数AI系统中不确定推理的现代方法的基础

神经科学：大脑是如何处理信息的？



简单的细胞的集合能导致思维、行动和意识

控制论：人工制品怎样才能在自己控制下运转

- 现代控制论和AI的共同点：设计出能随时间变化使目标函数最大化的系统。
- 控制论的主要工具是微积分和线性代数→其主要研究对象是用固定的连续变量集描述的（线性）系统
- AI的部分起因是寻求摆脱控制论的数学方法局限性的途径
 - 不同的工具：逻辑推理和计算
 - 不同的问题：语言、视觉、规划...

计算机工程：如何制造能干的计算机？

- AI需要智能和人工制品，即计算机。
- AI对主流计算机科学的影响
 - 分时技术
 - 交互式翻译器
 - 使用窗口和鼠标的个人计算机
 - 面向对象的编程
 - ...

■ AI的孕育 (1943-1955)

- 1943: Warren McCulloch and Walter Pitts: 人工神经元模型
 - 连接主义计算和学习(Hebbian学习)的第一步
 - Marvin Minsky and Dann Edmonds (1951) 构造了第一台神经网络计算机
- 1945年冯诺依曼(John Von Neumann): 冯诺依曼结构
- 1946年美国人毛奇莱(Mauchly)和艾克特(Eckert):世界上第一台电子计算机ENIAC
- 1948年维纳(N.Wiener) : 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。
- 1950: Alan Turing' s “Computing Machinery and Intelligence”
 - 首次清晰描述AI的全景

AI的诞生 (1956年)

- AI诞生于一次历史性的聚会
- 时间：1956年夏季 (2个月)
- 地点：达特莫斯 (Dartmouth) 大学
- 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(4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 (10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(Carnegie)工科大学

- 会议结果：

由麦卡锡提议正式采用了“Artificial Intelligence”这一术语

- 此后，美国形成了多个人工智能研究组织，如纽厄尔和西蒙的Carnegie RAND协作组，明斯基和麦卡锡的MIT研究组，塞缪尔的IBM工程研究组等。

■巨大的期望 (1952-1969)

- Newell and Simon制作了通用问题求解器

 - 模仿人类问题求解的规程

- Arthur Samuel (1952-)写了西洋跳棋程序达到业余高手水平

- John McCarthy(1958-) :

 - 定义了Lisp语言

 - 建议采纳者 (Advice Taker) , 实现了知识表达和推理的中心原则

■ AI的困境 (1966 - 1973)

- 发展进度比预期慢：不切实际的估计和预言
- 一些系统缺乏可操作性：组合爆炸
 - 程序原则上能够找到解并不意味着程序实际上包含找到解的机制。（组合爆炸问题）
 - 莱特希尔报告
- 技术和表达具有一些根本局限：Minsky and Papert (1969) 感知器

■ AI 的复兴：基于知识的系统 (1969-1979)

- 通用目的 vs. 专业领域
 - DENDRAL程序（质谱分子结构分析）：第一个成功的知识密集系统
- 专家系统
 - MYCIN用于检测血液感染：引入不确定推理
- 对知识表达研究的增加：逻辑语言，框架，语义网络等

■ AI成为工业 (1980 - present)

- 专家系统R1在数字设备公司 (DEC): 为新计算机系统配置订单
- DEC制作了40个专家系统, 杜邦公司有100个专家系统正在使用500个在开发...
- 美国的主流公司几乎都有自己的AI团队, 开发或使用专家系统
- 日本的第五代计算机和美国政府的MCC计划流产

■ 最终导致AI的冬天

■ 连接主义的复苏 (1986 - present)

- 并行分布式处理 (RumelHart and McClelland, 1986); 反向传播

■ AI成为科学 (1987 - present)

■ 严格的定理证明或确凿的实验证据

- 语音识别: 隐马尔可夫模型 (HMM)

- 神经网络

- 不确定推理和专家系统: Bayesian网络

- ...

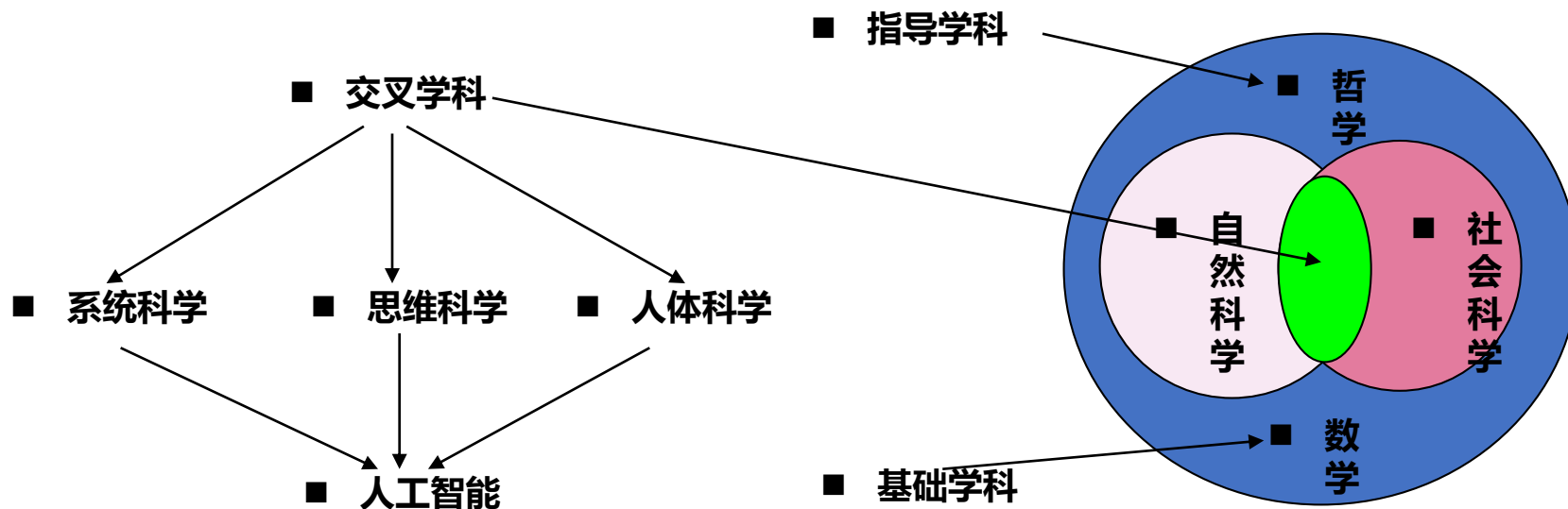
■ 智能化智能体(1995 - present)

■ 完整智能体问题:

- “在真实环境中, 具有连续传感输入的主体是如何行为的”

人工智能的发展

- 是自然科学与社会科学的交叉学科
- 交叉学科：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



与脑科学的交叉研究：

- **脑科学：又称神经科学，其目的是要认识脑、保护脑和创造脑。美国神经科学学会的定义：神经科学是为了了解神经系统内分子水平、细胞水平及细胞间的变化过程，以及这些过程在中枢的功能、控制系统内的整合作用所进行的研究。**
- **现代脑科学的基本问题主要包括：**
 - (1) 揭示神经元之间的连接形式，奠定行为的脑机制的结构基础**
 - (2) 阐明神经活动的基本过程，说明在分子、细胞到行为等不同层次上神经信号的产生、传递、调制等基本过程**
 - (3) 鉴别神经元的特殊细胞生物学特性**
 - (4) 认识实现各种功能的神经回路基础**
 - (5) 解释脑的高级功能机制等**

与**认知科学**的交叉研究

- **认知**：可一般地认为是和情感、动机、意志相对应的理智或认识过程，或者是为了一定的目的，在一定的心理结构中进行的信息加工过程。
- 美国心理学家浩斯顿(Houston)等人把认知归纳为以下5种主要类型：
 - (1) 认知是信息的处理过程；
 - (2) 认知是心理上的符号运算；
 - (3) 认知是问题求解；
 - (4) 认知是思维；
 - (5) 认知是一组相关的活动，如知觉、记忆、思维、判断、推理、问题求解、学习、想象、概念形成及语言使用等。
- **认知科学**：认知科学（亦称思维科学）是研究人类感知和思维信息处理过程的一门学科，其主要研究目的就是要说明和解释人类在完成认知活动时是如何进行信息加工的。

人工智能研究的目标

■ 近期研究目标

- 使现有的计算机更有用（或者说更聪明）

■ 远期研究目标

- 探讨智能的基本机理，研究如何利用各种自动机来模拟人的某些思维过程和智能行为

■ 主要研究方向

- 研究方向一：结构模拟，神经计算
- 研究方向二：功能模拟，符号推演
- 研究方向三：行为模拟，控制进化

人工智能的研究内容

(1) 机器学习 (machine learning)

让计算机能够像人一样自动的获取新知识，并在实践中不断的完善自我和增强能力。

围绕三方面展开：

1) 面向任务的研究：研究和分析改进一组预定任务执行性能的学习系统。

2) 模型研究：研究人类学习过程并进行计算机模拟。

3) 理论分析：从理论上探索各种可能的学习方法和独立于应用领域的算法。

(2) 知识获取 (knowledge acquisition)

从给定的知识库中产生新的知识，根据已有的知识建立动态数据结构，从环境中学习知识并精炼知识。

(3) 智能搜索 (intelligent search)

人工智能中的搜索问题是不确定的，访问搜索空间元素的顺序完全决定于数据集。

(4) 知识表示和推理 (knowledge representation and reasoning)

从一个或多个初始状态推理到达一个预先定义的目标状态。

(5) 规划 (planning)

规划旨在为活动的实体设计按时间顺序的合理活动序列。

(6) 软计算 (soft computing)

软计算是计算的工程方法，它对应于在不确定和不精确的环境下，人脑对于推理和学习的巨大能力。

包括模糊计算、神经计算、进化计算，分别涉及模糊逻辑、人工神经网络和遗传算法三大主要工具。

(7) 智能系统及智能计算机的构造技术

开展对智能计算机模型、系统分析与构造技术、建立工具及语言等的研究。

一、专家系统

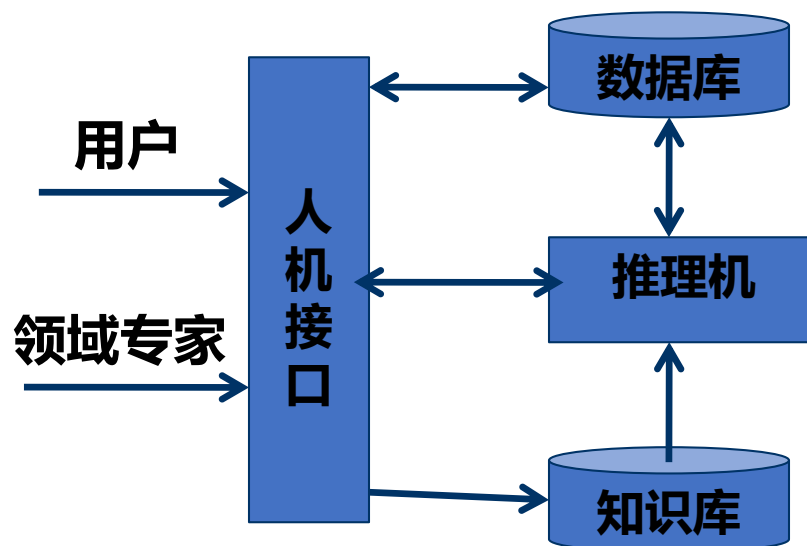
专家系统是一个智能计算机程序系统，其内部具有大量专家水平的某个领域知识与经验，能够利用人类专家的知识 and 解决问题的方法来解决该领域的问题。

著名的专家系统

ExSys: 第一个商用专家系统。

Mycin: 一个诊断系统，其表现出人意料的好，误诊率达到专家级水平，超出一些诊所的医生。

Siri: 一个通过辨识语音作业的专家系统，由苹果公司收购并且推广到自家产品内作为一个人秘书功能。

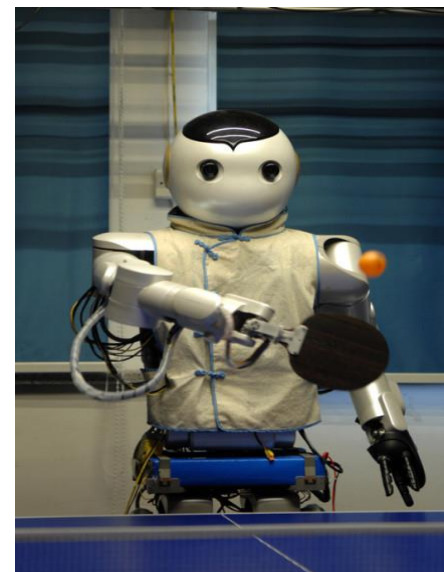
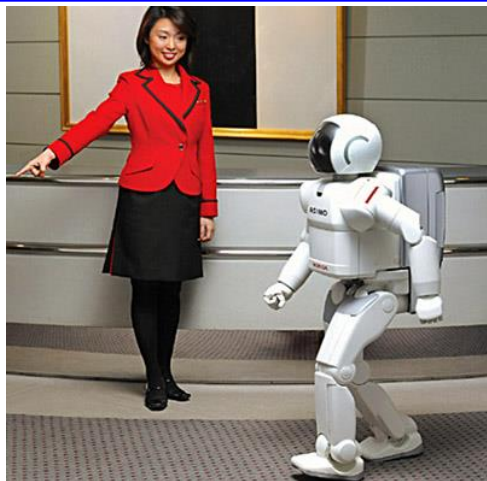


专家系统结构

二、机器人

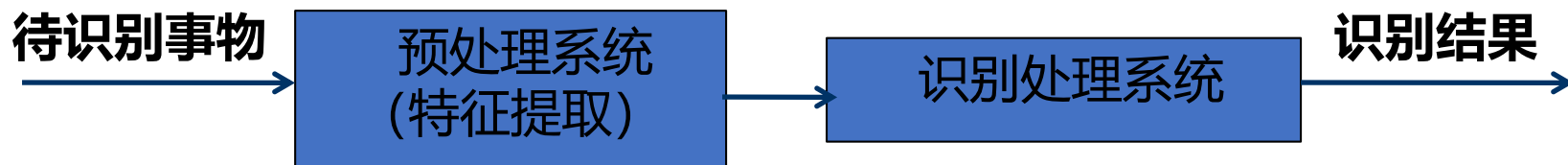
智能机器人的研究和应用体现出广泛的学科交叉，涉及众多的课题，机器人已在各领域获得越来越普遍的应用。诸如工业机器人、太空机器人、水下机器人、家用机器人、军用机器人、服务机器人、医疗机器人、运动机器人、助理机器人、机器人足球赛、机器人象棋赛, 几乎应有尽有。

人工智能的发展



三、模式识别

用计算机代替人类或帮助人类感知模式，是对人类感知外界功能的模拟，也就是使一个计算机系统具有模拟人类通过感官接受外界信息、识别和理解周围环境的感知能力。



模式识别过程

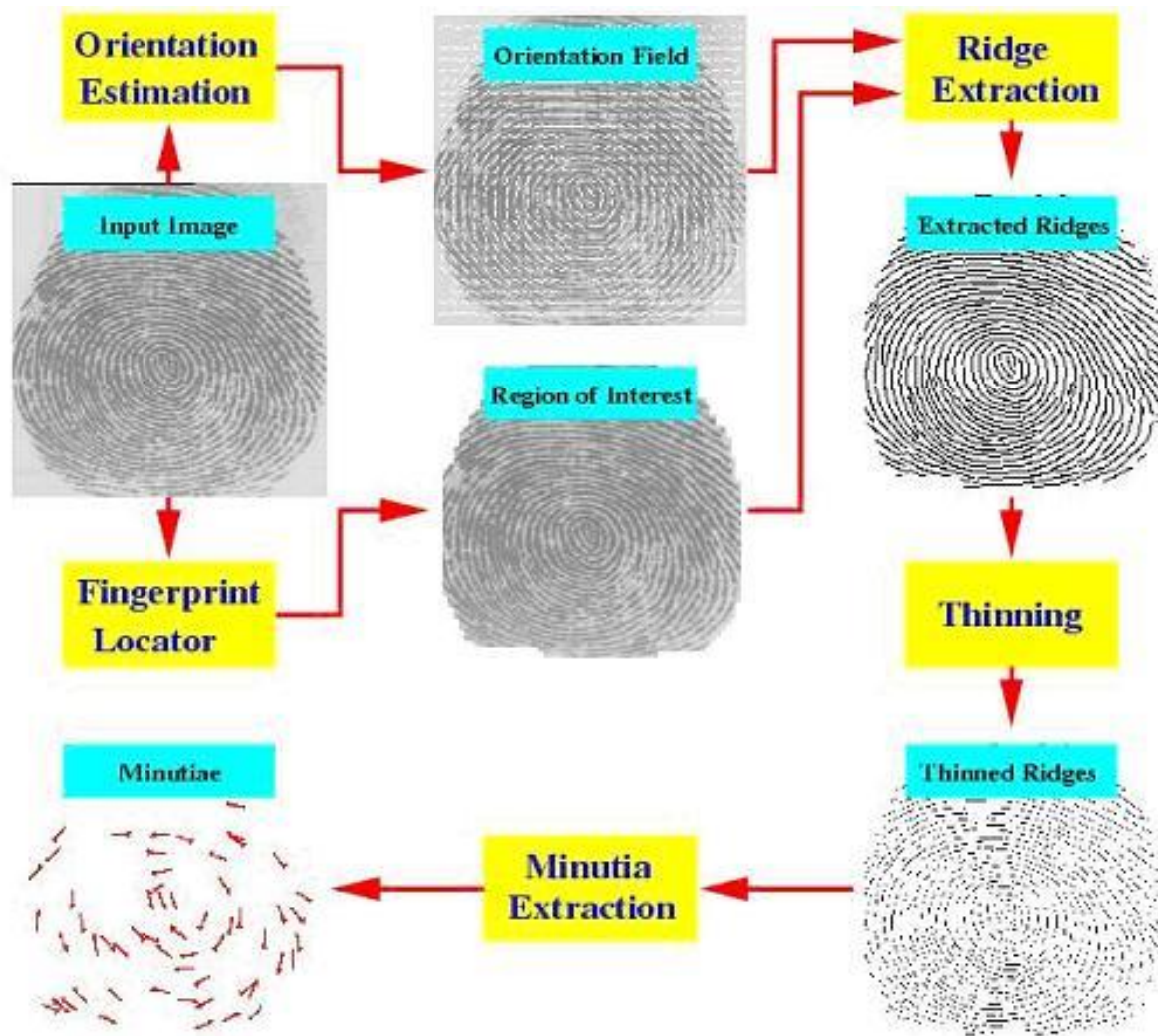


Figure 6: Flowchart of the minutiae extraction algorithm [18]. ©IEEE.

四、计算机视觉

从模式识别的一个研究领域发展为一门独立的学科。研究为完成复杂的环境中运动和在复杂的场景中识别物体所需要的视觉信息，以及如何从图像中获取这些信息的科学领域。

五、人工神经网络

神经网络处理直觉和形象思维信息具有比传统处理方式好得多的效果。人工神经网络用大量简单处理单元经广泛的连接而组成人工网络，模拟大脑神经系统的结构和功能。

六、自然语言理解

语言的生成和理解是一个极为复杂的编码和解码问题。

七、自动程序设计

包括程序综合和程序正确性检验两个方面的内容。

八、博弈

人工智能的许多方法都是从博弈程序中提炼出来的。



卡斯帕罗夫与“深蓝”的比赛



AlphaGO	Lee Se-dol
1202 CPUs, 176 GPUs, 100+ Scientists.	1 Human Brain, 1 Coffee.

需要指出的是, 以上我们仅给出了人工智能应用的部分领域和课题。其他的还包括课本上介绍的**自主规划和调度**、**自主控制**、**规划**等等。其实, 当今的人工智能研究与实际应用的结合越来越紧密, 受应用的驱动越来越明显。现在的人工智能技术已同整个计算机科学技术紧密地结合在一起了, 其应用也与传统的计算机应用越来越相互融合了, 有的则直接面向应用。归纳起来, 形成了以下几条主线:

- 从专家(知识)系统到Agent系统和智能机器人系统。
- 从机器学习到数据挖掘和数据库中的知识发现。
- 从基于图搜索的问题求解到基于各种智能算法的问题求解。
- 从单机环境下的智能程序到以Internet和WWW为平台的分布式智能系统。
- 从智能技术的单一应用到各种各样的智能产品和智能工程(如智能交通、智能建筑)。

- 人工智能是一门**广泛的交叉和前沿科学**，从1956年正式提出人工智能学科算起，已有50多年历史。目前人工智能在发展过程中既有突破但也面临很大的困难
- 除了**计算机科学**以外，人工智能还涉及**信息论**、**控制论**、**自动化**、**仿生学**、**生物学**、**心理学**、**数理逻辑**、**语言学**、**医学**和**哲学**等多门学科。

- 是二十一世纪 (基因工程、纳米科学、人工智能) 三大尖端技术之一
- 人工智能学科研究的主要内容包括：知识表示、自动推理和搜索方法、机器学习和知识获取、知识处理系统、自然语言理解、计算机视觉、智能机器人、自动程序设计等方面。

- 史上首个机器人公民 索菲亚来了: [史上首个机器人公民索菲娅](#)
- AlphaGoZero无师自通: [AlphaGoZero](#)
- 无人驾驶: [无人驾驶汽车](#)

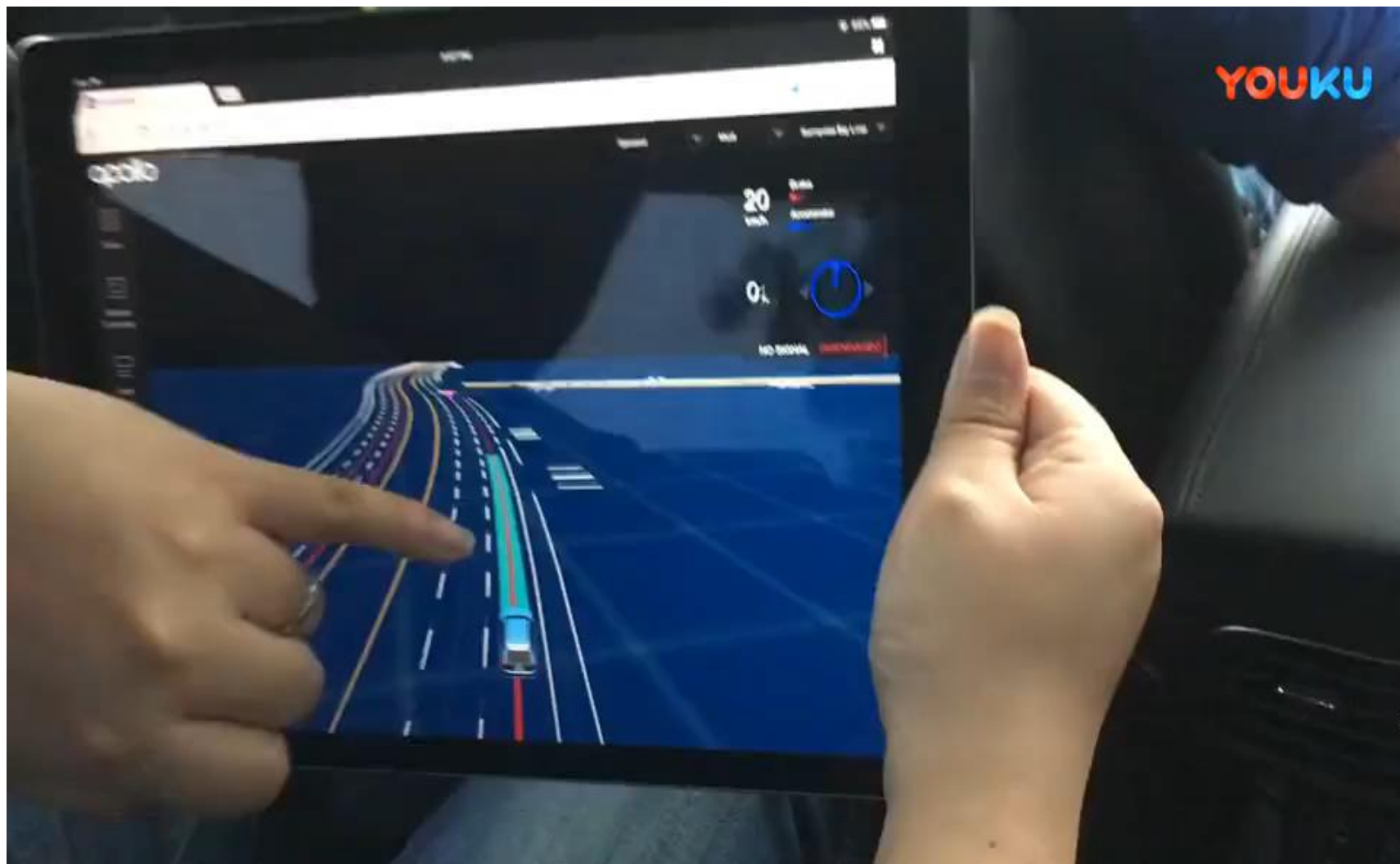
人工智能的发展



人工智能的发展

YOUKU

人工智能的发展



问题?