

人工智能导论（第一讲）

张玉志

一、课程的目的和目标

- 在人工智能很红火的年代，将来你们很可能从事人工智能。即使不专门从事AI，作为学软件工程出身的，也要看上去很懂。
- 比较全面、准确、清晰的理解人工智能的概念非常重要。
- 我们花比较大的篇幅讲一讲人工智能的概念
- 了解人工智能的历史和发展
- 了解人工智能的主要领域、主要热点
- 了解新旧人工智能的主要方法
- 学一点大家比较欠缺的博弈论

- 我们比较关注基本概念、思想、思路、研究的脉络
- 不注重一些细节知识的讲授
 - 发展变化很快，时间有限，学了用的可能性不是很大，可能价值也就不太大
 - 知识都是重要的，也都是深奥的。你知道有这方面的知识，需要用的时候再去仔细学学更好，我们课堂上不可能学的很深
- 不反对并且鼓励你去学任何的你感兴趣的的具体知识

计划的内容安排

- 4讲 概念、历史、概览
 - 人工智能的概念
 - 宏观看人工智能、大数据（人工智能和大数据对社会的影响）
 - 人工智能的历史、发展和未来，应用领域简述
 - 补充内容和开放式讨论课
- 4讲 传统人工智能
 - 传统人工智能介绍
 - 搜索
 - 专家系统
 - 搜索实验

- 5讲 新人工智能
 - 深度学习1
 - 深度学习2
 - 自然语言处理、GPT-3
 - 知识图谱
 - NKI人工智能平台和深度学习实验
- 2讲 博弈论
 - 简单博弈论
 - 博弈论的思考方式
- 2讲 外请讲座或讨论课

说明

- 我第三次上这门课，可能中间会有所调整
- 风格会跟其他老师有所不同
- 有时候有行政事务不得不调整（后面的知识先讲）
- 课堂方式有可能进行一些创新
 - 外请讲座方式、讨论课、学生讲
 - 鼓励你多学一点东西，思考一点东西，尽量避免知识灌输
 - 涉及到的很多知识，需要你自己去网上查、去钻研

考试问题

- 据说学生觉得最挨坑的事情就是讲的和考的不一样
 - 很遗憾，我准备用这种方法
 - 你可以欣慰的是：我尽量不会给你一个很差的分数
 - 但是：课没听，实验没做，什么都没干，我这里也没法通过

教材和参考书

- 王万良编著 人工智能导论第四版，高教出版社
 - 中规中矩的工科教材
- 李德毅主编 人工智能导论，中国科学技术出版社
 - 面向非计算机专业的，热点都讨论了
- 集智俱乐部编著 科学的极致—漫谈人工智能 人民邮电出版社
 - 非典型教材，比较思辨
- 赵军主编 知识图谱 高教出版社
 - 另外一个侧面
- 第六次国家信息领域技术预测 信息领域国内外技术竞争评价研究报告

- 希望你们从网上找一些知识
 - 比如：台湾李宏毅的深度学习
 - 或者直接听吴恩达的深度学习

- 根据大家已有的知识结构和掌握程度，以后会再补充一些书
- 怎么读这些书？
- 不一定要买。
- 我制作成了电子书
- 翻一下，不一定都读。更不要都仔细学

- 学习比较吃力的，精读第二本就够了
 - 特点是比较权威
- 爱思考的、爱辩论的，一定要读第三本
 - 特点是比较开放，充满了思辨
- 比较专注技术的，仔细学学第一本
 - 特点是比较正统，读下来会掌握一些具体知识
- 第四本可以简单看，有兴趣的仔细读
 - 知识图谱的书
- 第五本，可以挑着看看，国内专家对未来的看法和预测，基本代表了中国的技术发展水平
 - 建议你们翻一下，哪里有兴趣，去查资料

重点

- 掌握思维的脉络最最重要
- 具体的细节，用时再去查
- 大体有个印象，知道这方面有人研究过就不要自己摸索了就行
- 我是非常反对读死书、死读书，不希望你们读书死

二、到底什么叫人工智能，为什么复杂，为什么重要？

- **Artificial Intelligence**
 - 翻译成了人工智能
 - 其实应该翻译成“人造智能”
- 这个词用起来很模糊，内涵和外延不是很明确
 - 就像“漂亮”，什么叫长得漂亮？
 - 有公认的，大家都觉得很漂亮
 - 也有很大差异，有时你觉得不漂亮，别人觉得很漂亮。反之亦然。
- 80-90年代的几乎所有自由讨论最后都归结到：到底什么是人工智能，这个是不是人工智能

我想到的一些描述：从几个角度来说

- 手的延长是机器，脑的延长就是人工智能
 - 太过宽泛了一点
- 应该做个程序，程序不知道该怎么编，实在想不清楚，就叫人工智能
 - 有点不严格，到底谁不会编？
- 最优解求不出来（复杂性太高，往往针对NP问题），寻求满意解的过程
 - 有点太过局限，但这样的问题属于AI，很少有争议
- 寻求解决“传统的计算机解决不了但人能解决得很好”的途径
 - 有点接近实际情况，但是动态的，什么时候不再算是AI不好定

- 让你无法分辨是机器还是人在做（图灵实验）
 - 这是比较经典的，图灵的说法
- 达特茅斯会议：制造一台机器，该机器可以模拟学习或者智能的所有方面，只要这些方面可以精确描述
- Minsky：人工智能是一门学科，是使机器做那些人需要通过智能来做的事情
- 尼尔森：人工智能是关于知识的科学（知识的表示、知识的获取和知识的运用）

有人整理的：人工智能的定义

- **定义1 智能机器**

能够在各类环境中自主地或交互地执行各种拟人任务(anthropomorphic tasks)的机器。

- **人工智能 (Artificial Intelligence - - AI) :**

AI是关于知识的科学 - 怎样表示知识以及怎样获得知识并使用知识的科学。(Nilsson)

AI就是研究如何使计算机做过去只有人才能做的智能工作。(Winston)

人工智能的定义

其他几种关于人工智能的定义。

- **定义2 人工智能(学科)**

人工智能(学科)是计算机科学中涉及研究、设计和应用智能机器的一个分支。它的近期主要目标在于研究用机器来模仿和执行人脑的某些智力功能，并开发相关理论和技术。

- **定义3 人工智能(能力)**

人工智能(能力)是智能机器所执行的通常与人类智能有关的智能行为，如判断、推理、证明、识别、感知、理解、通信、设计、思考、规划、学习和问题求解等思维活动。

人工智能的定义

- **定义4** 人工智能是一种使计算机能够思维，使机器具有智力的激动人心的新尝试（Haugeland,1985）。
- **定义5** 人工智能是那些与人的思维、决策、问题求解和学习等有关活动的自动化（Bellman,1978）。
- **定义6** 人工智能是用计算模型研究智力行为（Charniak和McDermott,1985）。
- **定义7** 人工智能是研究那些使理解、推理和行为成为可能的计算（Winston,1992）。

人工智能的定义

- **定义8 人工智能**是一种能够执行需要人的智能的创造性机器的技术 (Kurzweil,1990) 。
- **定义9 人工智能**研究如何使计算机做事让人过得更好 (Rick和Knight,1991) 。
- **定义10 人工智能**是一门通过计算过程力图理解和模仿智能行为的学科 (Schalkoff,1990) 。
- **定义11 人工智能**是计算机科学中与智能行为的自动化有关的一个分支 (Luger和Stubblefield,1993) 。

其中，定义4和定义5涉及拟人思维；定义6和定义7与理性思维有关；定义8和定义9涉及拟人行为；定义10和定义11与拟人理性行为有关。

人工智能的定义

AI研究如何用计算机来表示和执行人类的智能活动，以模拟人脑所从事的推理、学习、思考和规划等思维活动，并解决需要人类的智力才能处理的复杂问题等。

AI还涉及到脑科学、神经生理学、心理学、语言学、逻辑学、认知科学等许多学科领域。是一门综合性的交叉科学和边缘学科。

- 明确不属于人工智能的一些行当
 - 自动控制类的，如飞机自动驾驶
 - 科学计算类的：石油勘探、天气预报、核计算、分析密码、动画渲染、比特币挖矿、办公室自动化等。
 - 估计目前世界上的90%以上的计算都不属于AI范畴
- 凡是固定了算法的，基本上都不算AI

- 人工智能的范围是不断变化的
- 过去只有人（主要靠智力）才能做的事情，尝试用机器去做，在尝试的这个阶段，属于人工智能的研究范畴
- 如果一个问题已经解决（有了一个明确的算法）得不错了（这个算法大家觉得已经比较满意了），基本上这个问题就不再属于AI了

人工智能与软件工程

- 随着计算机的发展，软件也在发展
- 在K这个量级的时候，汇编/机器语言
- 在M这个量级的时候，高级语言
- 在G这个量级的时候，？一直不明确
 - 深度学习是一个很好的方向

- 计算机的环境是一层层构造出来的
- 如何构造出用于人工智能的运行环境，就是软件工程的事情了
- 与新一代人工智能最密切的学科，就是软件工程了

人工智能和人类智能比较

- 人的智能 ($>$, $=$, $<$) 机器 (可能) 的智能?
- 人能不能造一种机器, 比自己的智能更强大?
- 上帝是万能的, 能不能创造一个石头, 自己搬不动?
- 就进入到悖论了。就非常复杂了

强人工智能VS弱人工智能

- 某个方面的人工智能，弱人工智能
 - 弱人工智能很强
- 综合的多方面的人工智能，强人工智能
 - 强人工智能现在很弱

人工智能的潜力和界限

- 要比较“人的智能”“机器的智能”，就把两者各自搞清楚
- 机器的智能：不管什么计算机，都等价于图灵机
 - 图灵机的概念一定要掌握
- 人的智能：还远远没有认识清楚

- 图灵机的定义，一定要掌握
 - 是一个数学概念
 - 一个七元组：纸带、符号、读写头、规则、状态、起始、结束。（存储、符号、读写、程序、数据、开始、结束）
- 图灵机停机问题
- 可计算性（不可计算的占了大部分）

- 哥德尔定理（哥德尔不完备性定理）
- 成为人工智能界限的基础问题
- 任何一个公理系统，都存在着一些命题，是不可能被证明的（不管真还是假）
- 几句话很难讲清楚，你们要看看书
- 康托尔对角线，哥德尔定理，图灵机不能停机，人工智能的极限，存在着密切的联系，看上去很有趣，也很深奥

- 几百年来，在不断的给“人”降级
 - 神的产物，宇宙的中心
 - 原来不在中心，是进化来的
 - 哥白尼的日心说。达尔文的进化论。
 - 佛洛伊德的潜意识理论，人连自己也控制不了
 - 人工智能说，会思考的不一定只有人，还有计算机
 - 人工智能的发展悲观者认为，从科学的角度，人会被淘汰
 - DNA的发现证明，人的遗传信息也没有太多
 - 有人写书：“人是机器”。有人认为：人就是“一堆肉”
- 人是世界的主宰、人定胜天。
 - 变成了人与自然和谐相处、敬畏自然
 - 逐步觉得其实人没什么了不起
- 我是谁，从哪里来，到哪里去？是永恒的话题

- 人这一头说不清楚，所以很难比较
- 人工智能之所以复杂，就在于还没有定型
- 谁最厉害，哪个方法是主流，还没有“固化”
- 你有一个新理论、新方法，很有效，你就很容易成为大师
- 所以，你自信的话，将来从事人工智能是一个不错的选择
- 涉及到的学科知识很多，涉及到科学、技术、工程

- 科学和技术
 - 科学是从现实总结规律，抽象化到理论
 - 技术是把科学规律、理论用于现实
- 一个是上山，一个是下山
- 技术和工程
 - 技术是能力，是方法
 - 工程是实践，是行动
- 中国缺少科学的基因
- 最近还在争论四大发明的问题
 - 四大发明是技术，不是科学
 - 中国的工程能力，现在是世界第一没问题

- 几个电子器件一搭起来， $高+高=低$ ， $高+低=高$ ， $低+低=低$ ，再把进位考虑进去， $高+高=(进位)高$ ，就是一个二进制的加法器。
- 减法可以变成补码的加法
- 移位就是乘法
- 乘法、除法可以变成移位加上加减法。
- 计算机的原理就这么诞生了。
- 在这上面一点点构造除了现在复杂的计算机世界
- 这就是科学的作用、工程、技术的作用

几种可能性

- 人工智能 < 人类智能
 - 人不用担心，但界限在哪里？
- 人工智能 = 人类智能
 - 人的效率低于人工智能，人类最后会非常被动。但当下没有什么太可怕的地方
- 人工智能 > 人类智能
 - 人现在就已经非常危险了，七八十年前就非常危机了

- 人工智能的概念之所以复杂，在与它实际到多个方面
- 科学的方面，技术的方面；计算机技术的方面，知识工程；数学，逻辑，认知科学等等等等

人工智能为什么重要？

- 人类社会发展的目标是什么？
 - 按照马克思的观点：生产力是社会发展的基础，人类社会发展的推动力是生产力的发展。当物质极大丰富的时候，就进入了共产主义社会了
 - 按照佛教的观点：地狱空了，都进入了佛的状态
 - 按照一些哲学家的观点：人类到哪里去，是一个永恒的基本问题
- 一般我们经常听到的比较具体的是：消除贫困（物质）、消灭战争（人人关系）、人与自然和谐美好（人与自然关系）、人们过上更加美好的生活

- 在人解决了能源之后，最大的追求（之一）就是不断扩展人类的智能了
 - 石油储量的增加、风电光电水电的利用、页岩气的开采、可燃冰、地热、核能、核聚变等发展，一个比一个丰富
- AI成为人类竞争、发展的最重要的手段和目标（之一）
- 物质世界以后都可以交给机器去处理，人只需要处理精神的世界，人类的生存形态将发生根本的改变
- 我们现在就是处在向这样的方向过渡的这样一个历史阶段

- 从人类绝对的生存质量到相对的竞争态势，AI都非常重要
- “这是最后的斗争”，也许人类要灭亡了的“最后”，也许是以后不用斗争了
- 人工智能在竞争中处于改变力量对比的位置
- 不管这么说，AI都有着极端的重要性

三、人工智能的三大流派

- 符号主义
 - 把现实的物，映射到代表它的符号，在符号上完成所有的推理、计算等等
 - 图灵测试就是这个思路
 - 非常有名的图灵测试到底是什么？
 - 屋子里有人、计算机，外边的人提问，交互用打印机、电话等方式，如果提问者无法分辨是人回答的还是计算机回答的，就是通过了图灵测试
 - 受到了数学的深刻的影响
 - 是最经典、接受程度最高的人工智能

人工智能的标准

图灵测试



- 连接主义

- 人脑就是神经元和之间的一些连接
 - 模拟人脑的做法，就可以模拟出人脑的功能
 - 于是提出了神经网络模型
-
- 新人工智能最红火的就是这一块
 - 深度学习的功能真的很强大
 - 但问题也不少，至少现在看是这样

- 行为主义
 - 智能来自更低级的感知和行动，表现得好就行
 - 波士顿动力公司的机器狗
 - 无人机
 - 群体智能

- 三大主义最近都有惊人的出色表现
- 有时三者在互相融合

四、人工智能简史和几个阶段

- 计算机发展非常平顺
- 摩尔定律：每过18-24个月，价格不变功能增强一倍或者价格降低一倍
- 人工智能是起起伏伏

- 早期人工智能
 - 从创立56年到80年代早期
 - 遍地开花，虽然很弱，但各方面都有所涉猎
 - 概念提出、机器定理证明、问题求解、博弈、机器翻译、机器人、模式识别等等
 - 从80年代到2010年代
 - 以专家系统为代表的，以知识为核心
 - 形成一个热潮
 - 九十年代衰退很厉害
 - 从近几年开始的新人工智能
 - 以深度学习为代表的，全新的人工智能
 - 深度学习非常厉害

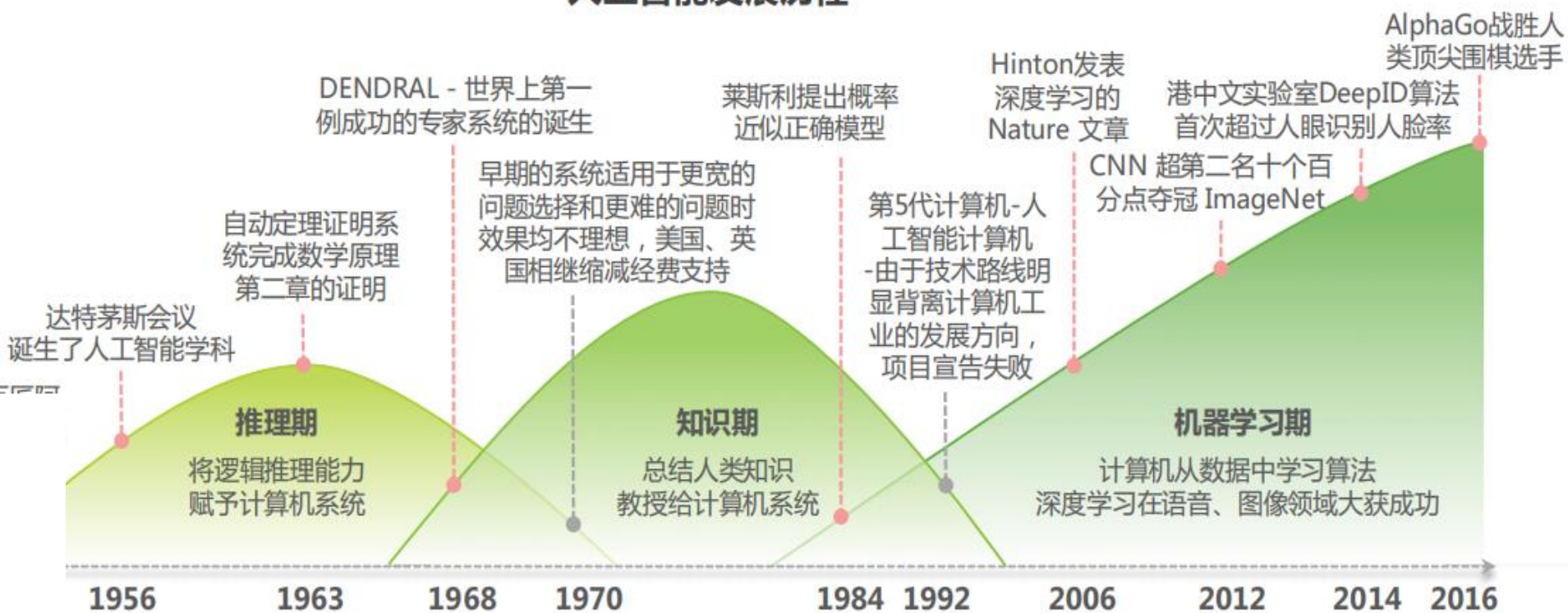
一、人工智能发展的三次浪潮

◆ 人工智能的诞生

- 1956年，斯坦福大学 McCarthy 教授、麻省理工学院 Minsky教授、卡内基梅隆大学的 Simon 和 Newell 教授以及贝尔实验室的 Shannon 、IBM 公司的 Rochester 等学者在美国达特茅斯 (Dartmouth) 学院首次确立了“人工智能”的概念。
- 当时的含义是“利用计算机模拟人类逻辑思维所体现的智能”，即让机器能像人那样理解、思考和学习。该定义由于过于强调“逻辑”，阻碍了神经网络等基于统计的人工智能理论的发展。
- 目前广为接受的定义是：人工智能泛指“一切人造系统的智能”。

一、人工智能发展的三次浪潮

人工智能发展历程



一、人工智能发展的三次浪潮

人工智能发展的三次浪潮

第一次浪潮

20世纪50年代末至70年代初

人工智能学科诞生，基于知识的方法和具有初步智能的机器出现。

由于数学模型存在缺陷，以及计算能力无法完成复杂度呈指数级增长的计算任务，使人工智能发展陷入低谷。

第二次浪潮

20世纪80年代初至90年代初

数学模型实现了重大突破，诞生了多层神经网络、BP反向传播算法和高度智能机器等。

由于人工智能计算机的成本与维护难度都较高，限制了其大规模商业应用和普及，使人工智能再次步入低谷。

第三次浪潮

21世纪初至今

大数据、算法模型和计算能力的多重突破共同驱动了新一代人工智能的快速发展，尤其是2006年深度学习神经网络的提出，使人工智能的性能获得突破性进展，标志着人工智能迎来第三次高速成长期。

一、人工智能发展的三次浪潮

◆ 新一代人工智能基本内容

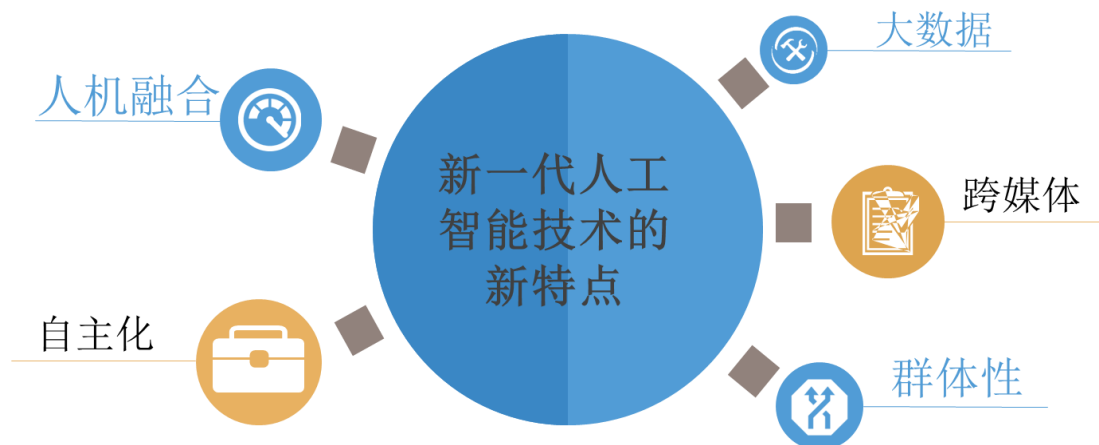
- “大数据驱动知识学习、跨媒体协同处理、人机协同增强智能、群体集成智能、自主智能系统成为人工智能的发展重点，受脑科学研究成果启发的类脑智能蓄势待发，芯片化硬件化平台化趋势更加明显，人工智能发展进入新阶段。
- 当前，新一代人工智能相关学科发展、理论建模、技术创新、软硬件升级等整体推进，正在引发链式突破，推动经济社会各领域从数字化、网络化向智能化加速跃升。”

——《新一代人工智能发展规划》

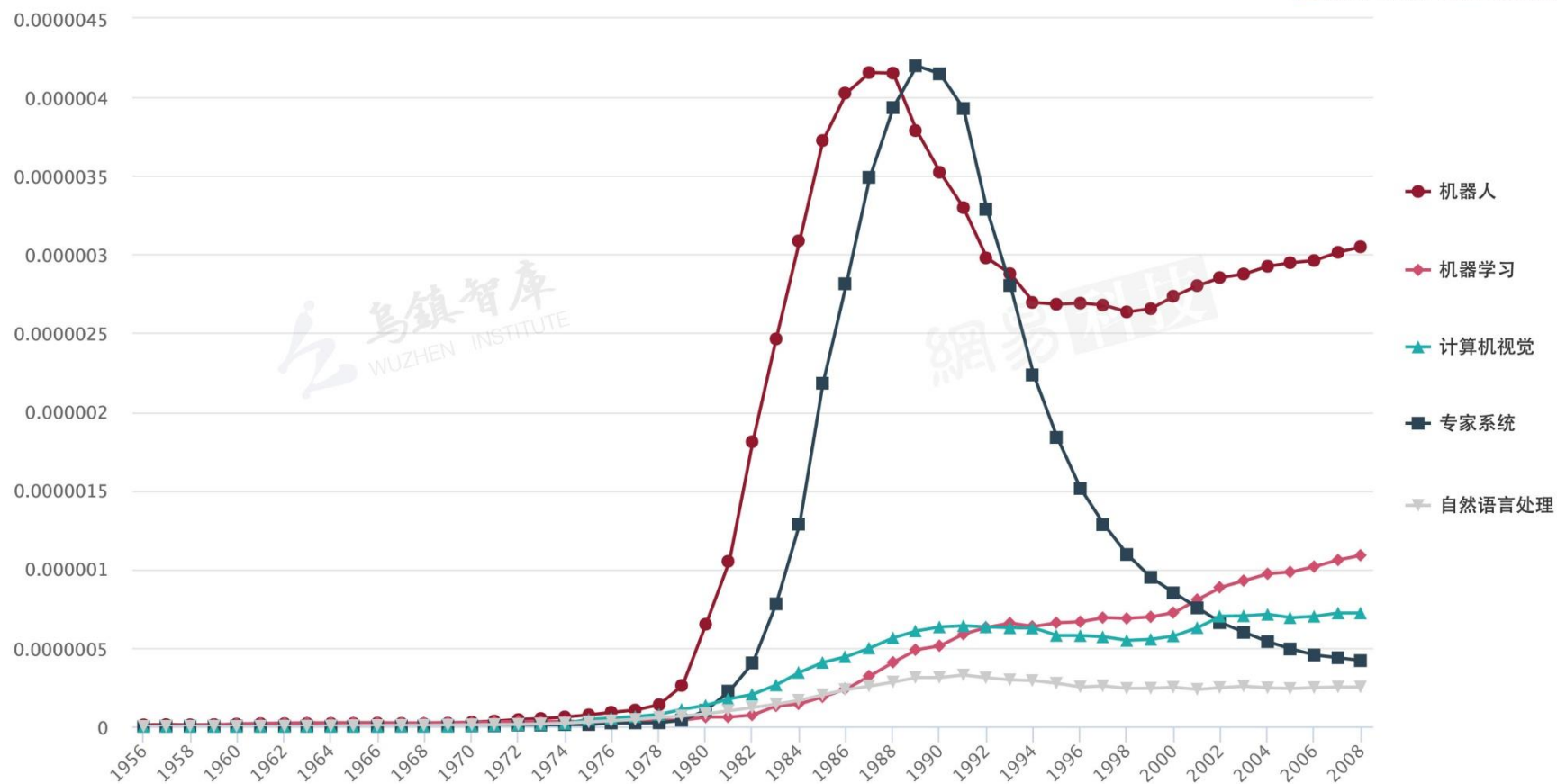
一、人工智能发展的三次浪潮

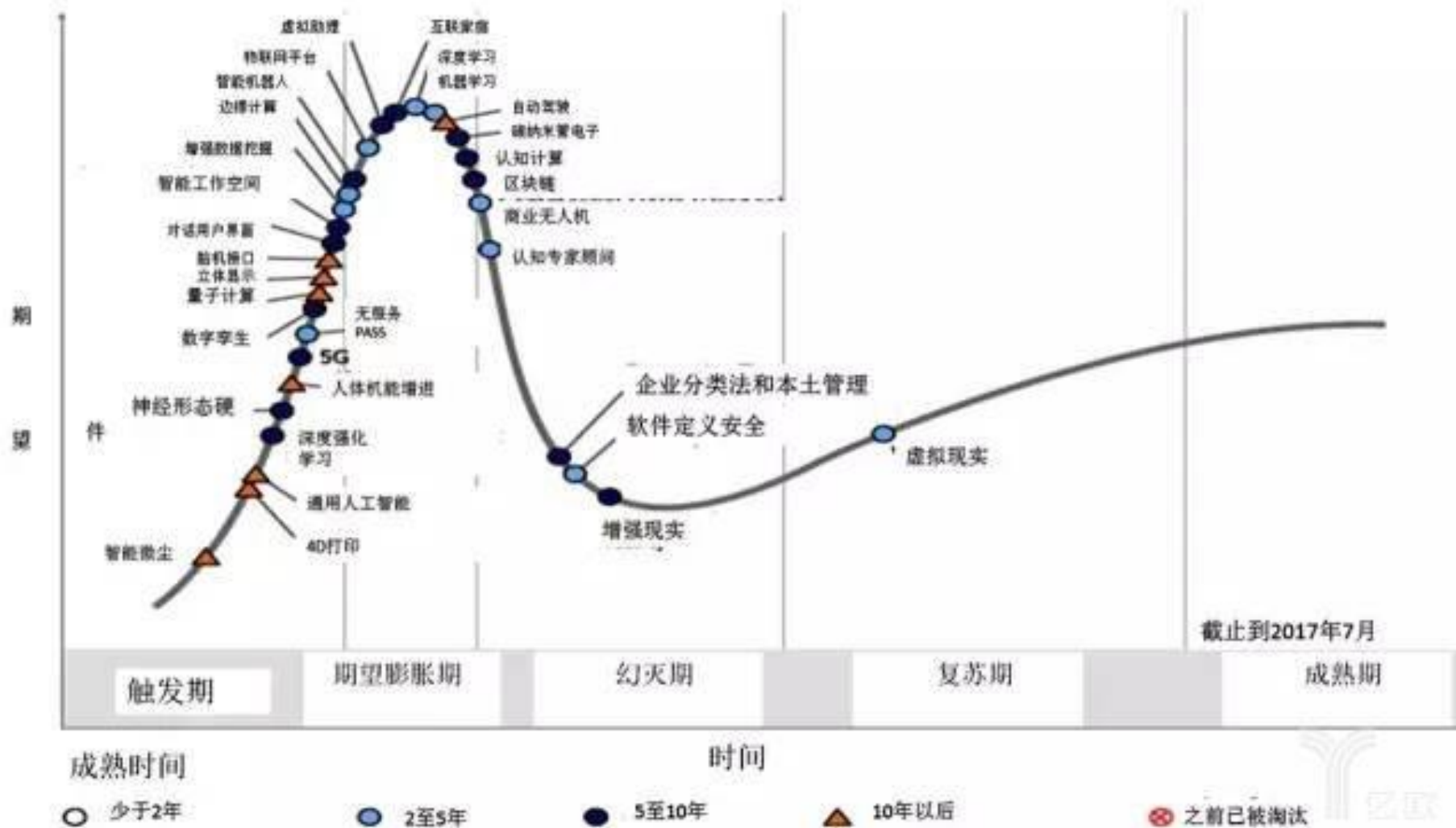
◆ 新一代人工智能与前两轮浪潮的本质区别

- 新一代人工智能技术已经可以大规模应用于经济社会实践，给企业带来盈利，给社会管理带来实质性改善。
- 新一代人工智能发展从学术驱动、研究驱动转变为需求驱动和产业驱动，企业的研发激励巨大、引领作用显著；前两次浪潮中主要依赖政府资助。
- 企业投资及研发主导、政府规划引领，使新一代人工智能发展正进入良性循环。



人工智能发展趋势





- 早期的人工智能是科技贵族的专属品
- 中期的人工智能更加偏向控制
- 在自动化、电子 --- 计算机 ---- 软件 是介于自动化和计算机之间的，自动控制的人出来研究的比较多
- 新人工智能是计算机和软件之间的，软件出来的人较多
- 南开的人工智能学院还是老人工智能为基础

自动化

电子

计算机

软件工程

老人工智能

新人工智能



五、人工智能的领域

- 常见的应用领域
 - 机器定理证明：机器推理过程
 - 早期非常红火，证明了四色定理
 - 中国的王浩教授
 - 中国的吴文俊
 - 早期取得的几个显著的成绩，成为人工智能的助推器

- 常见的应用领域
 - 博弈：人工智能的经典问题
 - 西洋跳棋（应该是穷举了）
 - 国际象棋（IBM的深蓝，有点诡异）
 - 中国象棋（中国大陆、台湾有人研究）
 - 围棋（最著名的是Alpha GO）
 - 成为新人工智能的标志事件
 - 从Alpha Go到Alpha Go Zero到Alpha Zero

- 常见的应用领域
 - 模式识别：声图文
 - 视频图像（目标跟踪、视频理解、智能检索等等）
 - 静态图像（人脸识别、图像理解、智能检索等等）
 - 声音识别（内容识别、声纹识别、语音矫正、语音合成、检索等等）
 - 文字识别（手写、联机手写、印刷体、多语种等等）

- 常见的应用领域
 - 自然语言处理（很多方向）
 - 翻译、理解、问答、摘要、作文作诗
 - 翻译是最复杂的，几乎跟世界一样复杂
 - 当成一个数学转换问题开始
 - 规则学派
 - 统计学派
 - 现在依然没有达到完全实用
 - 科大讯飞的翻译，到底作没作假？
 - （人工智能可以说有造假的传统）

- 常见的应用领域
 - 数据挖掘和知识发现
 - 这几年发展的还不错，各种推荐系统
 - 当年曾经号称发现了开普洛行星运行定律

- 常见的应用领域
 - 专家系统
 - 曾经非常的诱人
 - 把人类专家的知识，通过AI展示出来
 - 一度被费肯鲍姆吹得很神
 - 知识工程
 - 知识总结出来很难
 - 规则很多怎么办？
 - 模糊了怎么办？

- 自动程序设计（程序员的理想）
- 机器人（与硬件结合）
- NP-NPC问题（满意解是人工智能的精髓）
- 群体智能
- 智能+
 - 教育、网络、CAD、绘画、作曲、决策、模拟.....

- 从研究方向的分类
 - 知识表示（IF-then，框架，三元组，统计模型，HMM隐马尔可夫链）
 - 推理方法（归结原理，Lisp语言，Prolog语言）
 - 搜索技术（我们要讲一节，做一个实验）
 - 机器学习（怎计规则，调整参数）
 - 群智算法（简单智能能否配合）
 - 规划（机器人）
 - 神经网络
 - 图计算

- 也许将来南开的人工智能平台会成为一个重要的研究平台
- 希望你们不断的关注
- 也会不断做一些介绍工作