

《人工智能导论》课程知识点梳理

说明：

本文档基于 20 级钱班《人工智能导论》课程 PPT 整理，仅供课程面向的学生学习交流。文档旨在梳理知识点，以供参考者快速对课程内容留下印象，但**不建议当作背诵提纲使用**。复习最终请**以课程 PPT 内容为准**，相比记忆，更应注重概念理解和各种方法如推理计算方法的应用（相关提示以灰底字的形式给出）。

对应课程 PPT 一到四、五到七和八到十的部分分别由智电钱 001 王艺斐、智能钱 001 桂家彬和智能钱 001 黄智恒编写。王艺斐进行总的整理。

感谢部分 20 级人工智能试验班同学提供的宝贵意见和建议。感谢 20 级钱学组为此文档提供发布的平台。编写仓促，难免有大量不足之处，欢迎批评指正！

智电钱 001 王艺斐

2022 年 6 月 23 日

【一】绪论

- ◆ **关于智能的理论：**多元智能理论、智能三元理论（分析智能+创造智慧+使用智能）、智能发展理论（结晶的智能+流体智力）等
- ◆ **人工智能的奠基：**1956 年，Dartmouth 会议
- ◆ **人工智能三大学派：**
 - ①符号主义——认知即计算（源于数理逻辑，自上而下；研究抽象思维）；
 - ②连接主义——认知即网络（源于对人脑模拟，自下而上；研究形象思维）；
 - ③行为主义——认知即反应（源于控制论，系统工程思想；研究感知思维）。
- ◆ **人工智能发展阶段：**推理期→知识期→学习期→.....
- ◆ **人工智能分类：**符号智能（基于知识）和计算智能（基于数据）
- ◆ **人工智能研究领域：**深度学习、机器学习、自然语言处理、机器人、模糊逻辑系统、专家系统、无人驾驶汽车.....
- ◆ **机器学习的分类：**监督学习、无监督学习和强化学习
- ◆ **人工智能类型：**弱/窄人工智能，强/广泛人工智能，自我意识人工智能
- ◆ **莫拉维克悖论——**容易教会机器高级的逻辑计算推理，却很难完成感觉运动技能
- ◆ **弱人工智能与人类智能差异：**无主观能动性、环境适应性差、无决策能力和社会性
- ◆ **人类智能的神经基础：**大脑的神经网络
- ◆ **神经网络的发展历程：**
 - ①首个人工神经元模型，单层人工神经网络（感知器）——无法解决线性不可分问题；
 - ②神经认知机，Hopfield 网络，玻尔兹曼网络，反向传播算法，卷积神经网络；
 - ③深度学习，深度信念网络，生成对抗神经网络等。
- ◆ **中国新一代人工智能发展方向：**大数据智能、互联网群体智能、跨媒体智能、混合增强

【二】知识表示

- ◆ **知识的类型：**结构性知识、过程性知识、陈述性知识、启发式知识、元知识
- ◆ **数据→信息→知识→智慧**具有金字塔式的层次结构
- ◆ **知识的特性：**相对正确性、不确定性、可表示性及可利用性
- ◆ **基于符号的几种知识表示：**逻辑表示、语义网络表示、产生规则（产生式）、框架表示
- ◆ **知识的表示方法：**
 - ①陈述（叙述）式表示，描述事实性知识“做什么”，不能直接执行；
 - ②过程式表示，描述规则和控制结构知识“怎么做”，可以程序描述。
- ◆ **逻辑的分类：**形式逻辑（包括归纳逻辑与演绎逻辑），数理逻辑（即符号逻辑，包括命题逻辑和谓词逻辑）
- ◆ **命题：**一个能确定真假意义的陈述句（应会判断语句是否为命题）
- ◆ **命题分为：**原子命题和复合命题（原子命题+连接词）
（应会判断，掌握常见连接词、语法、推理规则等）
- ◆ **谓词逻辑表示：**将原子命题分解为个体、谓词、量词（全称量词、存在量词）
（应会理解、使用简单的谓词逻辑表示）
- ◆ **产生式表示：**表示事实、规则以及其不确定性度量
（应会理解、使用简单的产生式表示）
- ◆ **产生式系统：**包括规则库、推理机（控制系统）和综合数据库，是专家系统核心部分
- ◆ **语义网络表示：**由节点和节点之间的弧组成的一个有向图表示事物间关系
- ◆ **语义基元的结构：**三元组(节点 1，弧，节点 2)，弧方向不能随意改换
（应会理解、使用简单的语义网络表示，其中应会用常见的基本语义关系）
- ◆ **框架表示：**包括槽（可由侧面组成）和槽值（以及侧面值）
- ◆ **状态空间表示：**利用状态变量和操作符号表示系统或问题，一般可用有向图描述
- ◆ **状态空间四元组：**
 (S, O, S_0, G) 即(状态集合，操作算子集合，初始状态集合，目的状态集合)、
（应会用状态空间法求解简单的问题）

【三】知识图谱

- ◆ **知识的分类：**
 - ①可分为陈述性知识和过程性知识；
 - ②从知识工程角度可分为概念知识、事实知识、规则知识。
- ◆ **（电子工程中）本体的构成：**实体、类/概念、关系、属性、公理等
- ◆ **知识图谱的基本单位（核心）：**

关系三元组（“实体 1-关系-实体 2”）和属性三元组（“实体-属性-属性值”）

- ◆ **知识图谱的存储方式：**基于 RDF（资源描述框架）的存储和基于图数据库的存储
- ◆ **数据的类型：**结构化数据，半结构化数据和非结构化数据
- ◆ **知识图谱的架构：**
 - ①逻辑架构——模式层，数据层；
 - ②技术架构——信息提取，知识表示，知识融合，知识加工，知识更新。
- ◆ **技术架构各步中的关键技术——**
 - ①信息提取：实体抽取、关系抽取、属性抽取
 - ②知识表示：传统表示方法与深度学习为代表的表示学习基础
 - ③知识融合：实体链接（实体消歧和共指消解）与知识合并
 - ④知识加工：本体构建（人工/自动构建），知识推理（基于逻辑/图）和质量评估
 - ⑤知识更新：全面更新与增量更新

【四】搜索求解

- ◆ **搜索技术的关键：**搜索目标和搜索空间
- ◆ **搜索算法的分类：**
 - ①按问题表示方法，可分为状态空间搜索（状态空间法）和与/或树搜索（问题归纳法）；
 - ②按启发式信息的使用，可分为盲目搜索和启发式搜索；
 - 1.盲目搜索——深度优先搜索、深度限制搜索、广度优先搜索等；
 - 2.启发式搜索——贪婪搜索、A 搜索、A*搜索等。
- ◆ **搜索算法的评价指标：**完备性，最优性，时间复杂度和空间复杂度
复杂度相关的参数包括分支因子（每个结点最大分支数） b ，目标结点最浅深度 d ，最大路径长度 m ，状态空间大小 n 等
- ◆ **深度优先搜索（DFS）：**优先扩展深度最深的结点，限制最大深度成为深度受限搜索，通过直接修改状态生成可撤销的后继则称为回溯搜索
(应掌握用 DFS 及其衍生搜索方法求解简单问题)
- ◆ **广度优先搜索（BFS）：**优先扩展同级直接相连的结点
(应掌握用 BFS 求解简单问题)
- ◆ **DFS 与 BFS 的比较——**
 - ①完备性： b 有限的 BFS 完备，DFS 不一定完备；
 - ②最优性：若路径代价是结点深度非降函数 BFS 完备，其余情况 BFS 与 DFS 均不完备；
 - ③时间和空间复杂度：BFS 都为 $O(b^d)$ ，DFS 为 $O(b^m)$ 和 $O(bm)$ （PPT 上有误）。
- ◆ **一致代价搜索：**优先扩展代价最小的结点，是 BFS 基础上的扩展
(应掌握用一致代价搜索方法求解简单问题)
- ◆ **爬山搜索：**移动到邻近更优解，只能找到局部最优解，很依赖初始点位置

- ◆ **启发函数 $h(n)$** : 是结点的函数, $h(n) \geq 0$, 越小离目标越近, 等于 0 代表到达目标
- ◆ **评价函数 $f(n) = g(n) + h(n)$** : $g(n)$ 评估已付出代价, $h(n)$ 评估离目标距离;
 $f(n) = g(n)$ 为等代价搜索 (如 BFS), 有完备性; $f(n) = h(n)$ 没有完备性。
- ◆ **贪婪搜索: $f(n) = h(n)$** , 不最优也不完备, 最大复杂度 $O(b^m)$
(应掌握用贪婪搜索求解简单问题)
- ◆ **A 算法**: $g(n)$ 为实际路径代价, 启发函数 $h(n)$ 为评估的剩余代价 (不唯一)
- ◆ **A*算法**: $g^*(n)$ 为起始点到结点 n 的最低路径代价, $h^*(n)$ 为到目标的最低代价
 $g(n) > 0$ 是 $g^*(n)$ 的估计, $h^*(n) \geq h(n)$
- ◆ **A*算法一致性和完备性的条件**:
 - ①采纳性, 即 $h(n)$ 不会过高估计到达目标的代价;
 - ②一致性, 即通过结点 n 到达目标的代价不高于从结点 n 经 n' 到达目标的代价。
 (应掌握用 A 和 A*搜索方法求解简单问题)

【五】博弈搜索

- ◆ **博弈论**: 二人在平等的对局中各自利用对方的策略变换自己的对抗策略, 达到取胜目的
- ◆ **博弈的基本要素**: 博弈参加者, 博弈信息, 策略, 博弈次序, 博弈收益
- ◆ **博弈分类**:
 - ①按参加者的顺序, 分为静态博弈, 动态博弈 (先后;
 - ②按得益分类, 有零和博弈, 常和博弈, 变和博弈;
 - ③按策略分类, 有限博弈, 无限博弈;
 - ④按博弈信息 (如策略空间、收益) 了解情况分类, 有完全信息博弈, 不完全信息博弈;
 - ⑤按博弈过程的信息了解情况分类, 有完美信息博弈和不完美信息博弈;
 - ⑥按参与人是否合作, 分为合作博弈, 非合作博弈。
- ◆ **复杂求解问题表示**: 与/或树表示, 求解代价计算 (应掌握画法与代价计算方式)
- ◆ **零和博弈过程的表示**: 博弈树 (特殊的与或树), Max 方为或, Min 方为与
- ◆ **极小极大搜索**: 有限深度中使用带回溯的 DFS, 选择出对自己最有利的走法
(应掌握用极小极大搜索解决问题, 参考 PPT 理解)
- ◆ **$\alpha - \beta$ 剪枝**: 同时计算节点估计值和倒推值, 剪去无用分支
(应掌握用 $\alpha - \beta$ 剪枝搜索解决问题, 参考 PPT 理解)
- ◆ **非合作博弈的分类及对应的均衡理念**:

博弈特点	静态	动态
完全信息	完全信息静态博弈 纳什均衡	完全信息动态博弈 子博弈精炼纳什均衡
不完全信息	不完全信息静态博弈 贝叶斯纳什均衡	不完全信息动态博弈 精炼贝叶斯纳什均衡

- ◆ **非平衡博弈**：纳什均衡，每个参加者选择对自己最有利的策略；可以没有纳什均衡，也可以有多个纳什均衡

（应掌握其分析计算）

【六】确定性推理

- ◆ **推理的两个基本要素**：已知事实（证据）和知识（规则）
- ◆ **推理的内容**：前提和结论的命题内容
- ◆ **推理的形式**：是前提和结论的命题形式之间的联结方式
- ◆ **推理逻辑形式的有效性**：凡是从真前提必然得出真结论的推理，就是有效的
- ◆ **推理形式的解释**：将一个推理形式中的逻辑变项替换成具体内容的词项或命题
- ◆ **推理逻辑形式的可靠性**：
推理的前提和结论的联系是必然的，当且仅当推理是有效的，并且其所有前提都是真的
- ◆ **命题逻辑**：命题概念，命题联结词，复合命题符号化等（具体略）
（应会判断命题，进行复合命题符号化操作等，主要参考 PPT 例子，下同）
- ◆ **范式**：命题公式标准形式，其中析取范式为有限个简单合取式的析取（合取范式类似）
- ◆ **谓词逻辑**：谓词表示法，量词，谓词公式，逻辑等价式，推理规则等（具体略）
（应会理解各种谓词逻辑表示及其应用）
- ◆ **推理的控制策略**：指如何使用领域知识使推理过程尽快到达目标的策略
- ◆ **主要控制策略**：推理方向，冲突消解，求解策略，搜索策略，限制策略等
- ◆ **推理的方向**：正向推理，逆向推理，混合推理和双向推理
- ◆ **冲突消解策略**：解决冲突时规则的所用顺序，基本思想是对匹配的知识或规则进行排序
- ◆ **推理的分类**：
 - ①按知识的确定性分为确定性（精确）推理和不确定性（不精确）推理；
 - 1.精确推理有自然演绎推理、归纳演绎推理、与/或形演绎推理；
 - 2.不精确推理有似然推理，近似推理（模糊推理）；
 - ②按逻辑基础分为演绎推理（一般到个别，核心是三段论，不产生新知识），归纳推理（个别到一般，可以产生新知识），缺省推理。
- ◆ **演绎推理**：三段论法，假言演绎推理（充分/必要/充分必要假言推理）（具体略）
- ◆ **归纳推理**：
完全归纳推理（有必然性），不完全归纳推理（有或然性，包括简单枚举和科学归纳法）
（应会进行相关的演绎推理或归纳推理，判断推理形式有效性）
- ◆ **默认推理（缺省推理）**：知识不完全，没有足够反对证据就认为结论正确
- ◆ **自然演绎推理**：假言推理，拒取式推理，假言三段论等（具体略）
（应会进行对应的推理或识别推理中的错误）
- ◆ **归结推理的基本思想**：将证明 $P \rightarrow Q$ 转化为证明与 $\neg(P \rightarrow Q)$ 等价的 $P \wedge \neg Q$ 不成立
- ◆ **前束形范式**：所有全称量词和存在量词都非否定地出现在公式最前段，且辖域延伸到公

式末端，且公式中不存在蕴含“ \rightarrow ”也不存在双向蕴含“ \leftrightarrow ”

- ◆ **斯克伦范式**：前束形范式消去存在量词，后面跟合取范式
(应会运用相关规则将谓词逻辑表示化为前束形范式或斯科伦范式)
- ◆ **文字与子句**：文字是原子谓词公式及其否定，子句是文字或文字之析取（还有空子句）
(应会将谓词公式化为子句集)
- ◆ **子句集的形式**：原谓词公式不可满足（永假）和子句集不可满足等价
- ◆ **海伯伦域（H 域）**：判定此域上的一切解释，即可判断子句集不可满足性（应会构造）
- ◆ **鲁滨逊归结与归结反演**：命题公式与谓词公式的归结，不断归结直到出现空子句 NIL
(应会用归结反演方法求解或证明逻辑问题)

【七】多智能体系统

- ◆ **智能体**：具有自治性，社会性，反应性和预动性的基本特性的实体（软件、人、机器人）
- ◆ **多智能体系统**：多个相互作用的智能体构成
- ◆ **智能体三个主要工作构件**：传感器、执行器、效应器
- ◆ **PEAS——智能体的任务环境**：（应会指出智能体实例的 PEAS）
 - ①Performance Measure: 性能度量； ②Environment: 环境；
 - ③Actuators: 执行器； ④Sensors: 传感器。
- ◆ **理性智能体**：能够依赖感知序列能够最大限度地改进其性能度量
- ◆ **全知智能体**：明确知道行为实际结果，并且做出相应动作的智能体，现实不存在
- ◆ **智能体的任务环境分类**：
 - ①完全可观察（国际象棋）vs 部分可观察环境（驾驶）；
 - ②确定性环境（国际象棋）vs 随机性环境（驾驶）；
 - ③竞争环境（国际象棋）vs 合作环境（驾驶中互相避让）；
 - ④单智能体 vs 多智能体环境； ⑤动态环境 vs 静态环境；
 - ⑥离散（国际象棋）vs 连续环境（驾驶）；
 - ⑦片段式（问答机器人）vs 延续式环境（打网球）。
- ◆ **智能体的结构**：架构（机器）、函数（感知到行动的映射）、程序（函数的实现）
- ◆ **智能体的三种形式**：人类智能体、机器人智能体、软件智能体
- ◆ **智能体的特点**：情境性、自主性、适应性、社交性
- ◆ **智能体的类型**：简单反射智能体、基于模型的反射智能体、基于目标的智能体、基于效用的智能体（以及学习智能体，任何智能体都可以通过学习改善性能）
- ◆ **狭义智能体的特性**：自主性、交互性、协作性、可通信性、长寿性
- ◆ **智能体的强性质**：移动性，诚实性，无私性，理性

【八】自然语言处理

- ◆ **自然语言**：指英语、汉语等人们日常使用的，随着人类社会发展演变而自然约定俗成的语言，区别于程序设计语言等人造的语言
- ◆ **自然语言的常见形式**：口头语言（语言）和书面语言（文字）
- ◆ **自然语言处理（NLP）**：用计算机处理加工自然语言，实现人机间的信息交流
- ◆ **自然语言的主要分类（基于结构形态）**：屈折型语言、黏着型语言、孤立型语言
- ◆ **语言与语法**：语言由语句构成，以词为基本单位；语句由单词构成，词由词素构成；语法制约词素如何构成词，词如何构成词组和句子；词法包括构词法和构形法
- ◆ **自然语言处理的主要方法**：理想主义方法、经验主义方法、连接主义（神经网络）方法
- ◆ **自然语言处理两大核心任务**：自然语言理解、自然语言生成
- ◆ **自然语言处理的层次化**：语音分析→词法分析→句法分析→语义分析→语用分析
- ◆ **自然语言处理的具体方法**：
 - ①基于传统机器学习的方法：语料预处理→特征工程→选择分类器
 - ②深度学习的 NLP 流程：语料预处理→设计模型→训练模型
- ◆ **自然语言语料预处理的步骤**：
 - ①英文，分词→词干提取→词形还原→词性标注→命名实体识别→分块；
 - ②中文，中文分词→词性标注→命名实体识别→去除停用词。
- ◆ **典型的分词方法**：基于字典匹配，基于统计，基于深度学习
- ◆ **中英文分词的典型区别**：中文更难分词，英文单词形态较多，中文分词要考虑粒度问题
- ◆ **中文分词难点**：无统一标准，歧义词区分，新词识别
- ◆ **自然语言模型**：通过语料计算某个句子的出现概率，如 N-Gram 语言模型
- ◆ **自然语言理解（NLU）难点**：语言的多样性、歧义性、鲁棒性、知识依赖和上下文
- ◆ **自然语言理解（NLU）的实现方式**：基于规则→基于统计→基于深度学习
- ◆ **自然语言生成（NLG）的方式**：文本到语言（text-to-text）、数据到语言（data-to-text）
- ◆ **自然语言生成的三个等级**：简单的数据合并、模板化的 NLG、高级 NLG
- ◆ **NLG 的主要步骤**：
 - 内容确定→文本结构→句子聚合→语法化→参考表达式生成→语言实现
- ◆ **自然语言处理应用**：情感分析，机器翻译，聊天机器人，语音识别等

【九】机器学习

- ◆ **机器学习**：指使计算机无需明确的变成就具有学习能力的研究领域
- ◆ **机器学习的关键因素**：数据量、计算能力和存储能力
- ◆ **机器学习的分类**：监督学习、无监督学习以及强化学习
- ◆ **机器学习的阶段**：训练阶段与测试阶段，机器学习框架即训练与测试
- ◆ **机器学习的本质**：从数据中进行预测或决策
- ◆ **机器学习的基本架构**：

环境→学习模块→知识库→执行模块，执行模块又作用于学习模块

◆ **机器学习中的数据拆分：**训练数据、验证数据、测试数据

◆ **数据的属性：**体积、多样性、速度、价值、准确性

◆ **机器学习的任务与方法：**

特点	监督学习	无监督学习
离散	分类、归类	聚类
连续	回归	降维

◆ **分类的结果与性能指标：**混淆矩阵、准确率；第一类错误与第二类错误；精确率、召回率以及特异性（后三者是准确率的细化补充）

◆ **二分类的性能指标：**ROC 曲线与 AUX（ROC 曲线下面积，越接近 1 越好）

（简单了解以上指标的计算方法以及相互之间的关联）

◆ **特征选择方法：**过滤法、包装法、嵌入式方法

◆ **监督学习：**

①模型，决策函数或条件概率分布；

②策略，包含损失函数，风险函数，经验损失，极大似然估计等（具体略）；

③算法，分为生成方法（先学习出联合概率密度和特征概率密度再计算条件概率密度）

与判别方法（直接学习决策函数或条件概率密度）。

◆ **具体的主要的监督学习算法：**

①分类算法，有朴素贝叶斯、决策树、SVM（支持向量机）、逻辑回归等；

②回归算法，有线性回归，回归树等；

③分类+回归算法，有 K 最近邻，Adaboosting，神经网络等。

◆ **朴素贝叶斯分类：**假设特征属性相互独立且权重相同，用贝叶斯定理计算各类别概率

（应掌握相关的计算）

◆ **基于特征空间距离的分类：**以特征空间中两样本点间的距离（如闵可夫斯基距离，包括曼哈顿距离和欧氏距离等）作为相似性的反映

◆ **K 最近邻算法：**基本思路是看样本最近邻的 K 各样本中的大多数属于哪一类别

◆ **决策树的步骤：**特征选择→决策树生成→决策树剪枝

◆ **决策树基础：**

基本量有信息量 $I(x_k) = -\log_2 p(x_k)$ ，信息熵 $H(X) = \sum I(x_k)p(x_k)$ ，经验分割信息熵等

①信息增益 $g(D, A) = H(D) - H(D|A)$ ，用在 ID3 算法（信息增益愈大愈好）

②信息增益率 $g_R(D, A) = H(D)/H_A(D)$ ，用在 C4.5 算法（先选信息增益大于阈值的，然后信息增益率愈大愈好）

③基尼指数 $Gini = \sum p_k(1 - p_k) = 1 - \sum p_k^2$ ，用在 CART 算法（基尼指数越小越好）

（应结合课件掌握相关的计算）

◆ **支持向量机（SVM）：**用支持向量进行超平面划分，完成分类或回归问题

◆ **无监督学习：**

聚类（层次聚类、K 均值聚类）和降维（主成分分析、奇异值分解、独立成分分析）

◆ **强化学习：**算法与环境 and Agent 交互有关，基于“奖励假设”

【十】深度学习

◆ **神经元的典型结构与工作机制：**抑制与兴奋对应“0-1”律

◆ **人工神经网络（ANN）基本构成：**输入层、隐藏层、输出层

◆ **人工神经网络的参数与功能：**权重、阈值、求和函数、激活函数

◆ **最简单的前馈神经网络——感知器：**

基本思想、组成（见 PPT）和局限（不能解决线性不可分问题）

◆ **人工神经网络发展：**感知器（单层神经网络）、多层神经网络、循环神经网络、随机神经网络、卷积神经网络等（绪论部分也有相关内容）

◆ **深度学习：**是基于对数据进行表征学习的算法，具有多层结构的网络

◆ **深度学习的分类：**

①起源于感知器的是有监督学习（深度神经网络→卷积神经网络）；

②起源于受限玻尔兹曼机的是无监督学习（深度信念网络，深度玻尔兹曼机）；

◆ **卷积神经网络（CNN）：**具有局部连接、权值共享等特点的深层前馈神经网络

◆ **CNN 的结构：**输入层、卷积层、池化层、全连接层、输出层（结构及作用见 PPT）

◆ **卷积神经网络的核心出发点：**局部感受野、多卷积核、权值共享、池化汇聚

◆ **当前深度学习瓶颈：**泛化能力问题、强泛化能力不足

【十一】人工智能哲学观与伦理

※考虑到内容特点与上课讲授情况，此处不作梳理，有兴趣自行阅读 PPT 即可。