

## 第一讲 知识工程概述 (1)

- 知识工程基本概念 (1)
  - 什么是知识 (1)
  - 什么是知识库 (1)
  - 知识工程 (2)
- 知识工程的认知基础 (2)
  - 认知语言学 (2)
    - 意向图式 (3)
    - 范畴 (4)
  - 认知神经科学 (6)
    - 学习知识的四个支柱 (6)
    - 深度学习 (7)
- 知识工程的发展历程 (8)
- 大数据时代的知识工程：知识图谱 (9)
  - 简介 (9)
  - 知识的类型 (9)
- 大数据时代的知识工程：预训练语言模型 (12)
  - 符号表示和数值表示对比 (13)
- 知识工程生命周期 (14)
  - 符号化工程生命周期 (14)
  - 参数化知识工程的生命周期 (17)

## 第二讲 符号化知识表示理论和方法 (20)

- 概述 (20)
  - 知识表示的五个主要角色 (20)
  - 知识类型 (21)
    - 语言知识 (21)
    - 世界知识 (22)
    - 常识知识 (23)
  - 知识表示 (23)
- 产生式规则 (24)
  - 正向推理 (24)
  - 反向推理 (25)
- 语义网络 (25)
- 框架 (26)
- 脚本 (27)
- 一阶谓词逻辑 (28)
- 描述逻辑 (30)
- 语义网 (31)
  - XML (32)

- RDF (33)
- 本体, OWL (34)
- 知识图谱 (36)
- 代表性的符号化知识工程 (38)

## 第三讲 数值化知识表示理论和方法 (45)

- 语言的分布表示 (45)
  - 传统独热 (46)
  - 矩阵表示 (46)
  - 神经网络 (47)
- 知识的分布表示 (48)
  - 位移模型 TransE (49)
- 预训练语言模型 (51)
  - 形式语言模型 (符号) (51)
  - 统计语言模型 (符号) (51)
  - 神经语言模型 (数值) (52)
  - 预训练语言模型 (数值) (53)
- 预训练语言模型可以作为世界模型吗? (60)
  - 什么是世界模型? (60)
  - 如何构建世界模型? (61)
  - 语言模型中的时空表示 (61)
  - 语言模型可以作为世界模型 (62)
- 预训练语言模型可以作为知识库吗? (63)
  - 大模型如何存取知识 (64)
  - 如何更新LLM里的知识? (65)

## 第四讲 知识获取 (上) (66)

- 信息抽取概述 (66)
- 命名实体识别 (67)
  - 定义, 任务, 特点 (67)
  - 典型方法 (68)
    - 基于词典的方法 (68)
    - 基于统计的命名实体识别 (68)
    - 基于阅读理解命名实体识别 (69)
    - 基于模板生成的命名实体识别 (70)
    - 基于大模型的命名实体识别 (71)
    - 多模态命名实体识别 (71)
- 关系知识抽取 (73)
  - 面向非结构化文本的关系抽取 (73)
    - 限定域关系抽取 (73)
      - 关系分类 (74)
      - 实体关系联合抽取 (75)
    - 开放式关系抽取 (77)
      - 传统方法 (77)

## 第五讲 知识获取（下） (78)

1. 事件知识抽取 (78)
  1. 事件元素抽取 (79)
  2. 事件关系抽取 (80)
  3. 多米太事件知识抽取 (81)
2. 脚本知识抽取 (83)
3. 多粒度知识联合抽取 (84)

## 第六讲 大模型知识分析、萃取与增强 (85)

1. 大模型的知识分析 (85)
  1. 知识探测 (86)
    1. 基于提示的知识探测是否严谨? (87)
  2. 知识定位 (87)
    1. 语言模型如何存取知识? (87)
    2. 基于梯度归因的知识定位 (88)
    3. 基于因果分析的知识定位 (89)
    4. 多种语言、多种架构大模型的知识定位 (89)
    5. 语言无关/简并 知识神经元 (89)
  3. 知识分析 (90)
    1. LLM难以学习长尾知识 (90)
    2. 共现频率影响事实知识 (91)
    3. 逆转诅咒 (91)
2. 大模型的知识萃取 (92)
  1. COMET (92)
  2. ATOMIC10X (92)
  3. CN-AutoMIC (93)
  4. 大模型中萃取脚本知识 (94)
3. 大模型的知识增强 (94)
  1. 大模型中的幻觉现象 (94)
  2. 知识增强的背景 (95)
  3. 知识增强的分类 (95)
  4. 检索增强 (95)
    1. 面向黑盒大模型的检索增强: REPLUG (95)
    2. 模仿人类搜索过程的大模型: WebGPT (96)
    3. 利用大模型生成外部知识进行检索增强 (97)
    4. 知识拉锯战: 检索增强中的知识冲突问题 (97)
4. 大模型的工具增强 (97)
  1. 工具增强的背景 (97)
  2. Toolformer (98)

## 第七讲 知识图谱构建实战 (99)

1. 知识图谱实践的系统工程观念 (99)
  1. 工程观 (99)
  2. 系统观 (99)
2. 领域知识图谱框架构建 (100)
  1. 基本思路 (100)
3. 半结构化文本中的知识抽取 (101)
4. 非结构化文本中的知识抽取 (103)
5. 知识图谱众包构建 (106)
6. 知识图谱质量控制 (107)
  1. 质量评估维度 (108)
  2. 质量评估方法 (108)
7. 实践案例 (109)
8. 实践经验 (113)
  1. 基本原则 (113)
9. 大模型时代知识图谱构建实战 (114)
  1. 知识图谱框架构建 (114)
  2. 非结构化文本中的知识抽取 (115)

## 第八讲 知识图谱数据管理：存储与检索 (117)

1. 知识图谱数据管理概述 (117)
2. 符号化知识图谱数据管理 (117)
  1. 知识图谱数据模型 (117)
    1. RDF图模型 (118)
    2. 属性图模型 (118)
  2. 知识图谱数据的存储 (119)
    1. 基于表结构的存储 (119)
      1. 关系数据库 (120)
    2. 基于图结构的存储 (122)
      1. 常见系统  
Neo4j, OrientDB, HyperGraphDB, InfiniteGraph
3. 知识图谱数据的检索（查询） (123)
  1. RDF图 (123)
    1. SPARQL (123)
  2. 属性图 (126)
    1. Cypher Gremlin PGQL (126)
4. 代表性工具简介
  1. Neo4j CQL (127)
3. 参数化知识图谱（大模型）数据管理 (130)
  1. 大模型知识存储 (131)
  2. 大模型的知识编辑（检索与更新） (132)

- 1. 全量微调+正则化 (132)
- 2. 超网络(Hyper-network)方法 (元学习思想) (132)
- 3. 定向更新方法 (133)
- 3. 代表性工具简介 (134)

## 第九讲 知识建模与知识融合 (135)

- 1. 知识建模 (135)
  - 1. 知识体系概述 (135)
  - 2. 常用的知识组织形式 (136)
    - 1. Ontology (本体) (136)
    - 2. Taxonomy (分类体系) (138)
    - 3. Folksonomy/Metadata (开放标签) (138)
  - 3. 典型知识体系 (138)
  - 4. 知识体系手工建模方法 (139)
    - 1. 知识体系构建的目标 (139)
  - 5. 知识体系自动建模方法 (140)
    - 1. 基于结构化、半结构化数据的知识体系构建 (140)
    - 2. 基于非结构化数据 (纯文本) 的知识体系构建 (141)
      - 1. 术语、概念抽取 (141)
      - 2. 同义词挖掘 (143)
      - 3. 关系挖掘 (143)
        - 1. 上下位关系 (143)
        - 2. 属性 (144)
- 2. 知识融合 (144)
  - 1. 知识融合概述 (145)
  - 2. 知识体系融合方法 (146)
    - 1. 基于文本相似度的映射方法 (146)
      - 1. 基于字符串匹配的映射方法 (146)
      - 2. 基于语言处理的映射方法 (146)
      - 3. 基于语义匹配的映射方法 (147)
    - 2. 基于结构相似度的映射方法 (147)
      - 1. 基于内部结构的映射方法 (147)
      - 2. 基于外部结构的映射方法 (147)
      - 3. 基于网络表示学习的映射方法 (147)
  - 3. 知识实例融合方法
    - 1. 实体消歧定义 (149)
    - 2. 实体消歧分类 (149)
      - 1. 基于无监督聚类的实体消歧 (149)
      - 2. 基于实体链接的实体消歧 (150)
        - 1. 输入输出 (150)
        - 2. 主要步骤 (150)
          - 候选实体的发现(150)

- 候选实体的链接(151)
- 无链接实体 (NIL) 的聚类
- 3. 大模型中的知识融合 (153)
  - 1. 大模型对齐技术概述 (154)
  - 2. 大模型对齐方法 (154)

## 第十讲 知识推理 (155)

- 1. 概述 (155)
  - 1. 什么是推理 (156)
  - 2. 推理方式及分类 (157)
  - 3. 知识推理典型应用 (159)
    - 知识图谱补全, 知识问答, 搜索与推荐, 行业应用
  - 4. ChatGPT推理能力分析 (159)
- 2. 演绎推理 (161)
  - 1. 经典逻辑推理 (161)
  - 2. 基于产生式规则的推理 (162)
  - 3. 基于概率逻辑学习的推理 (163)
  - 4. 自然语言演绎推理 (165)
- 3. 归纳推理 (167)
  - 1. 归纳推理概述 (167)
  - 2. 归纳逻辑程序设计 (167)
    - 1. FOIL算法 (168)
  - 3. 路径排序算法(PRA算法) (168)
  - 4. 关联规则挖掘算法(AMIE算法) (169)
- 4. 基于深度学习的知识推理方法 (170)
  - 1. 基于表示学习的知识图谱推理 (170)
  - 2. 基于强化学习的知识图谱推理 (171)
- 5. 大语言模型下的推理方法 (171)
  - 1. 大模型推理的背景 (172)
  - 2. 大模型推理增强方法 (172)
    - 1. 引入外部工具集 (173)
    - 2. 引入外部知识源 (173)
    - 3. 引入多模型决策 (174)
    - 4. 优化推理路径 (175)
  - 3. 大模型推理的未来发展 (176)

## 第十一讲 问答系统 (177)

- 1. 问答系统概述 (178)
  - 1. 历史 (178)
  - 2. 定义 (178)
  - 3. 分类-根据答案资源分类, 根据问答形式分类 (178)
- 2. 检索式问答系统 (179)

1. 已有方法 (179)
  2. IBM Watson (180)
  3. 阅读理解-分类:选择式, 填空式, 抽取式, 生成式 (181-182)
  4. 小结 (182)
3. 社区问答系统--主要任务是相似问题检索 (183)
1. 传统基于符号匹配的相似问句匹配方法--VSM, BM25等 (184)
  2. 基于深度学习的相似问句匹配方法 (184)
  3. 小结 (185)
4. 知识图谱问答概述 (185)
1. 基于语义匹配的知识图谱问答 (186)
    1. 基于语义图的知识图谱问答 (187)
    2. 基于端到端模型的知识图谱问答 (189)
    3. 小结 (191)
  2. 基于语义解析的知识图谱问答 (191)
    1. 语义表示(Lambda演算、DCS Tree、...)(191)
    2. 基于规则的语义解析方法: CCG (192)
    3. 基于统计模型的语义解析方法: PCCG (193)
    4. 基于神经网络的语义解析方法: Seq2Seq (194)
  3. 面向表格数据的问答系统 (196)
    1. 任务描述、典型数据集 (196)
    2. 面向单张表格的NL2SQL (197)
    3. 面向多张表格的NL2SQL (198)