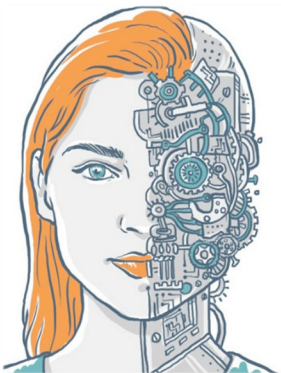


知识图谱

(Knowledge Graph)

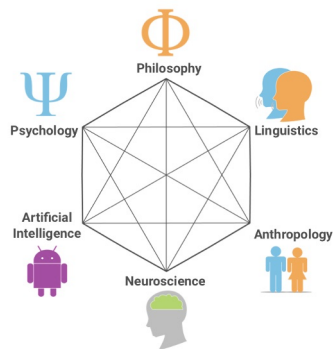


Mind or Machines
Cognitive Science Changing
Artificial Intelligence

汪 鹏

pwang@seu.edu.cn

东南大学 KGCODE实验室



东南大学 计算机科学与工程学院/人工智能学院

课程信息

● 教师

— 汪鹏，计算机楼**217**办公室

— Office time:

周一~周日

(**9:30~17:30**, **19:00~23:00**)

● 助教

— 马鹤文、李一宁

● 同学

— 41人



群聊：《知识图谱》2024年春



该二维码7天内(2月25日前)有效，重新进入将更新

课程主页

公开课程

<https://github.com/npubird/KnowledgeGraphCourse>

Unwatch 98

Fork 995

Star 3.2k

最新: 3.7k
访问量已超过60万次

A systematic course about knowledge graph for graduate students, interested researchers and engineers.

东南大学《知识图谱》研究生课程

时间: 2019年春季 (2月下旬~5月中旬)

每周五下午2:00~4:30

地点: 东南大学九龙湖校区, 纪忠楼Y205

答疑/讨论/建议: 请致信 pwang AT seu.edu.cn

课程内容

第1讲 知识图谱概论 (2019-3-1, 2019-3-8)

1.1 知识图谱起源和发展



爱可可-爱生活

8-31 来自Mac客户端

'东南大学《知识图谱》研究生课程(资料)' by
Peng Wang GitHub: [网页链接](#)

课程内容

第1讲 知识图谱概论 (2019-3-1, 2019-3-8)

1.1 知识图谱起源和发展

1.2 知识图谱 VS 深度学习



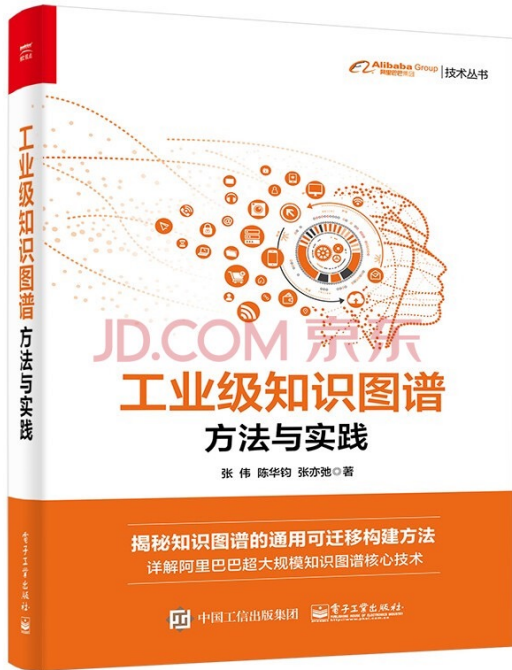
售出>30000册!



国内访问最多的知识图谱课程之一

国内最畅销的知识图谱专著之

参考教材



参考课程

CS 520

2020~2022

Knowledge Graphs

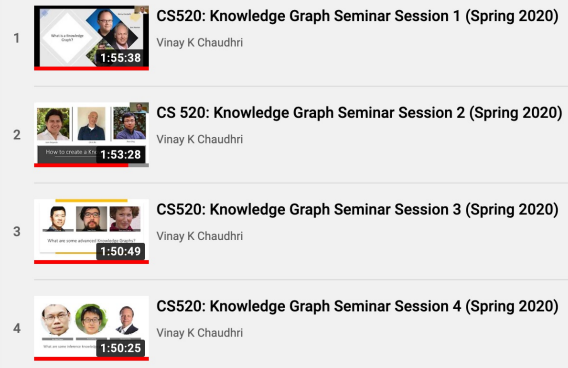
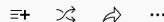
Data Models, Knowledge Acquisition, Inference and Applications

Department of Computer Science, Stanford University, Spring 2021

Tuesdays 4:30-5:50 P.M. PDT and Thursdays 4:30-5:50 P.M. PDT

https://www.youtube.com/playlist?list=PLDhh0lALedc7LC_5wpi5gDnPRnu1GSyRG

https://www.youtube.com/playlist?list=PLDhh0lALedc5paY4N3NRZ3j_ui9foL7Qc

CS520: Knowledge Graphs
Seminar (Spring 2020)10 个视频 · 19,558次观看 · 最后更新时间: 2020
年6月2日

内容安排

- 知识图谱基础：概念、技术和应用
- 知识表示
- 知识建模
- 知识抽取基础
- 知识抽取：实体识别（前沿进展）
- 知识抽取：关系抽取（前沿进展）
- 知识抽取：事件抽取（前沿进展）
- 知识融合（前沿进展）
- 知识表示学习和知识推理（前沿进展）
- 知识图谱和大模型（前沿进展）

实践示例

- 实践示例代码、数据和虚拟机下载：

链接: https://pan.baidu.com/s/1MT_MnnnBX5_s6nRycile3w 提取码: qfp9

- 参照《VirtualBox 安装部署》文件说明，安装VirtualBox虚拟机软件，建议给虚拟机分配至少8G内存。

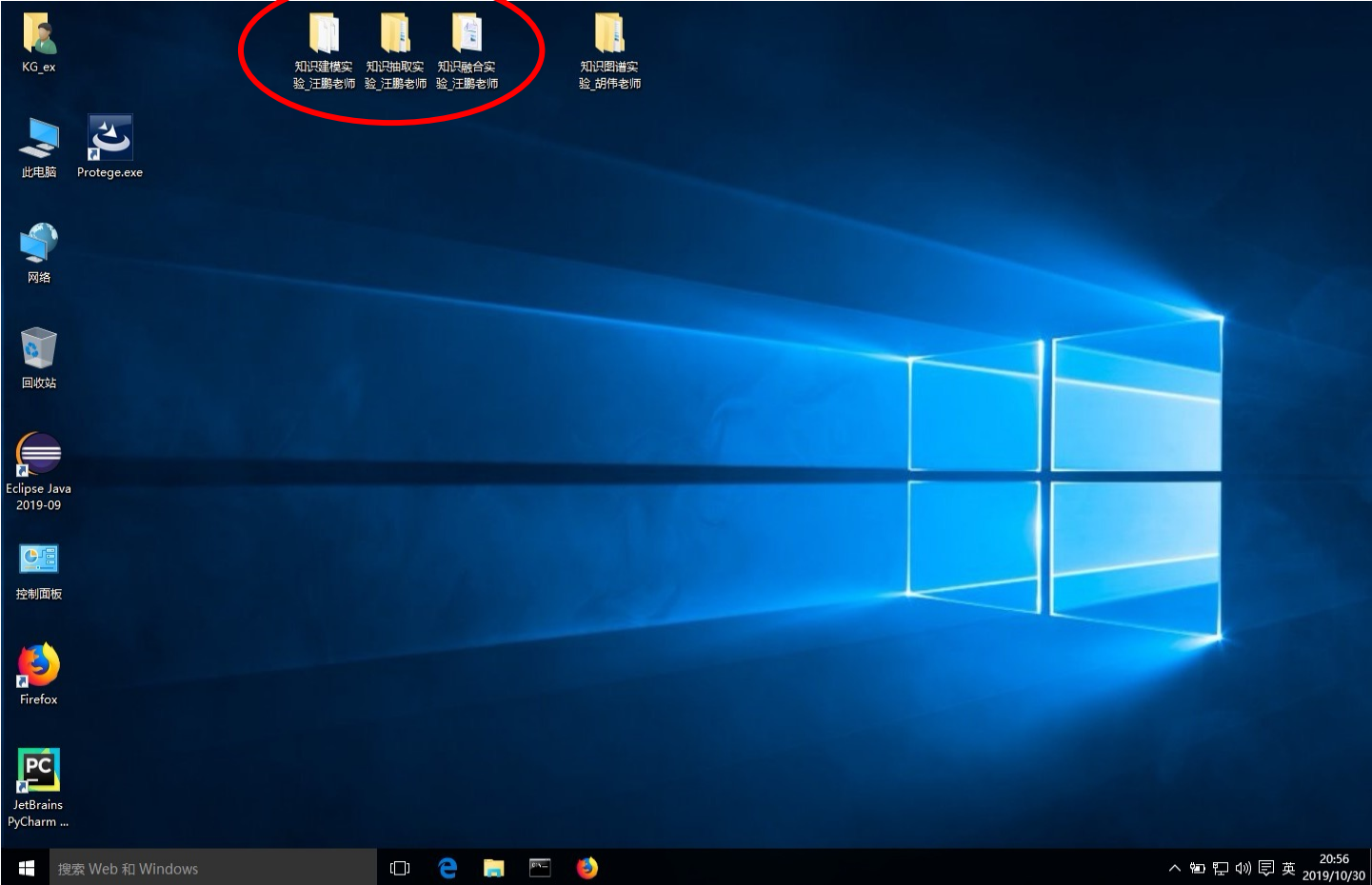
提供了Windows10和Mac OS下的虚拟机软件安装说明。

- 导入已做好的虚拟机文件KG_ex.ova后，运行虚拟机



感谢徐忠楷等同学准备的实践训练

实践示例



课程报告要求

● 报告要求

- 以课程实践内容为基础
- 鼓励多次迭代修改，形成可发表论文
- 报告模板（见统一提供）

课程考核

- 成绩综合考虑
 - 课程报告
 - 评测排名/获奖
 - 课堂报告分享

课程报告

二选一

● 1. 知识图谱相关项目实践

- 独立完成一个特定领域的知识图谱系统，包括：知识表示和建模（本体层）、知识图谱构建（数据爬取和清洗、实例层，实体识别，关系抽取）、知识融合、知识存储、智能应用（搜索、问答、分类、预测、推理等）

● 2. 知识图谱任务评测比赛

- 参加1项知识图谱相关评测比赛，团队成员不超过4人；
- 指导老师包含任课老师。
- 推荐评测比赛（不限于以下）
 - ✓ CCKS2023~2024 NLPCC2023~2024 CCL等学术评测
 - ✓ kaggle、biendata、天池等平台的知识图谱相关任务评测
 - ✓ TAC(<https://tac.nist.gov/>)、SemEval等NLP领域经典评测任务
 - ✓ KDD-CUP, WSDM, ISWC, OAEI等经典会议知识图谱任务

课程报告

● 报告提交

- 报告模板会发在课程群，按照模板填写完整相关信息，根据实践撰写报告内容。
- 对参加评测的实践，给出组队情况和评测结果排名。
- 课程报告由李一宁同学统一收齐。
- 报告最迟提交时间为本学期结束时间。

事件检测——2022年赛题回顾

高鲁棒性要求下的领域事件检测

数据清洗

任务抽象

模型训练

集成推理

训练集id 71

句子文本

.... 空客A330客机向跑道滑行,在经过停在一旁的土耳其航空公司一架A321客机时发生碰撞,A330的机翼将A321的垂直尾翼切断。事件未造成人员伤亡...

事件类型 试验 (Experiment)

触发词

碰撞

论元信息

事故后果: A330的机翼将A321的垂直尾翼切断。事件未造成人员伤亡



训练集id 71

句子文本

.... 空客A330客机向跑道滑行,在经过停在一旁的土耳其航空公司一架A321客机时发生碰撞,A330的机翼将A321的垂直尾翼切断。事件未造成人员伤亡...

事件类型 意外事故 (Accident)

触发词

碰撞

论元信息

事故后果: A330的机翼将A321的垂直尾翼切断。事件未造成人员伤亡



事件检测——2023年赛题思考

多要素交织情况下的领域事件抽取

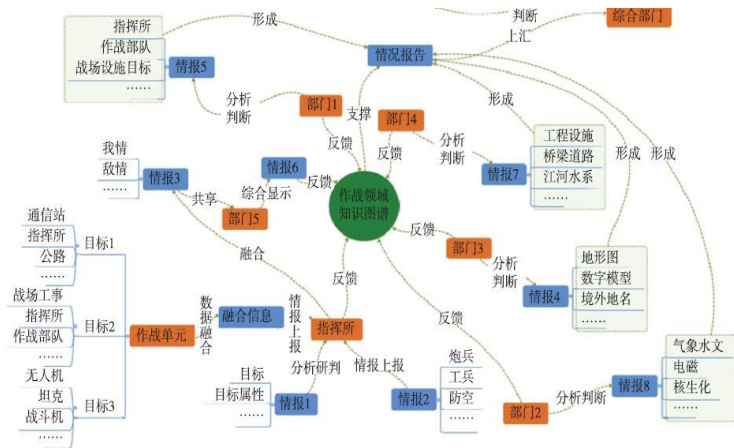
总体背景

- 军事ZZ数据和QB数据中蕴含了大量事件，且一篇文档或一句话中会包含多个事件，呈现出事件交织的情况
- 事件交织下的共享要素，往往反映出事件的顺承、因果等关系，事件关系能够为事件推理提供依据

军事事件文本存在要素交织的情况



事件关系能对动态的客观世界进行推理



事件检测——2023年赛题思考

多要素交织情况下的领域事件抽取

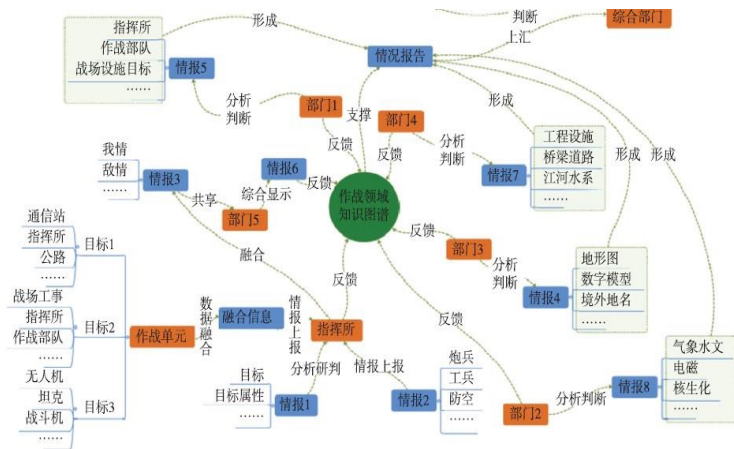
总体背景

- 军事ZZ数据和QB数据中蕴含了大量事件，且一篇文档或一句话中会包含多个事件，呈现出事件交织的情况
- 事件交织下的共享要素，往往反映出事件的顺承、因果等关系，事件关系能够为事件推理提供依据

军事事件文本存在要素交织的情况



事件关系能对动态的客观世界进行推理



事件检测——2023年赛题思考

多要素交织情况下的领域事件抽取

设计思路

- ◆ 一个句子中只包含一个事件
- ◆ 只抽取事件的触发词
- ◆ 评价指标是准确率和Macro-F1

- ① 要素抽取
- ② 评价指标



- ◆ 一个句子中可能会包含多个事件，事件间存在论元重叠
- ◆ 检测出多个事件，同时抽取触发词和论元
- ◆ 引入冗余度、关联度、准确度、完备性等更丰富的评估指标，为后续建立事件图谱和后续的事件推理奠定基础

示例

输入

4月7日,美国海军与印度海军在孟加拉湾举行联合军演。该演习为期10天。4月14日及16日,美印海军联合舰队举行了新型海军装备的展示活动。美军卡尔-文森号航母.....



输出

事件1 触发词: 联合军演

论元: 4月7日 (时间), 美国海军和印度海军 (举办方)
类型: 演习 (Manoeuvre)

事件2 触发词: 展示

论元: 4月14日及16日 (时间), 美印海军舰队 (举办方)
类型: 展示 (Exhibition)

持续学习场景下的军事领域知识抽取

总体背景

- 军事ZZ数据和QB数据具有**实时更新性**，并且存在**实体层次化特性**，在一个阶段中新的数据中往往会出现**新实体类型**或**父子实体类型**，呈现出**实体类型增长**的情况；
- 持续知识抽取**需要模型在不重复访问原始旧数据的情况下，具有快速学习**新知识类型**同时可以维持旧实体类型不被遗忘的能力，能够长期保证军事领域数据的综合处理及利用能力

军事数据实体/关系类型存在**类增量**的情况

国家 / 轰炸机 / 坦克

- ◆ 4月29日，两架**B-1B轰炸机**曾从**美国**本土远程奔赴**中国**南海参加演习。
- ◆ **俄罗斯**媒体“记者报”的报道称，**T-14坦克**在**叙利亚**拉塔基亚省的**普罗维登斯**参加了与反对派武装的战斗。

国家 / 导弹 / 防空系统 / 无人机

- ◆ “**超 530**”型**导弹**是**法国**玛特拉公司研制的，1977 年投入生产。
- ◆ “**铠甲**” S1是一种**俄罗斯**研发的弹炮合一防空武器系统，主要用于中低空的拦截防御作战，按理来说对付**土耳其**察打一体**无人机**应当是得心应手。

军事数据实体存在**实体层次化**的情况

武器装备 / 飞机器 / 轰炸机

- ◆ 4月29日，两架**B-1B轰炸机**曾从美国本土远程奔赴中国南海参加演习。

武器装备 / 陆军装备 / 坦克

- ◆ 俄罗斯媒体“记者报”的报道称，**T-14坦克**在叙利亚拉塔基亚省的**普罗维登斯**参加了与反对派武装的战斗。

持续学习场景下的军事领域知识抽取

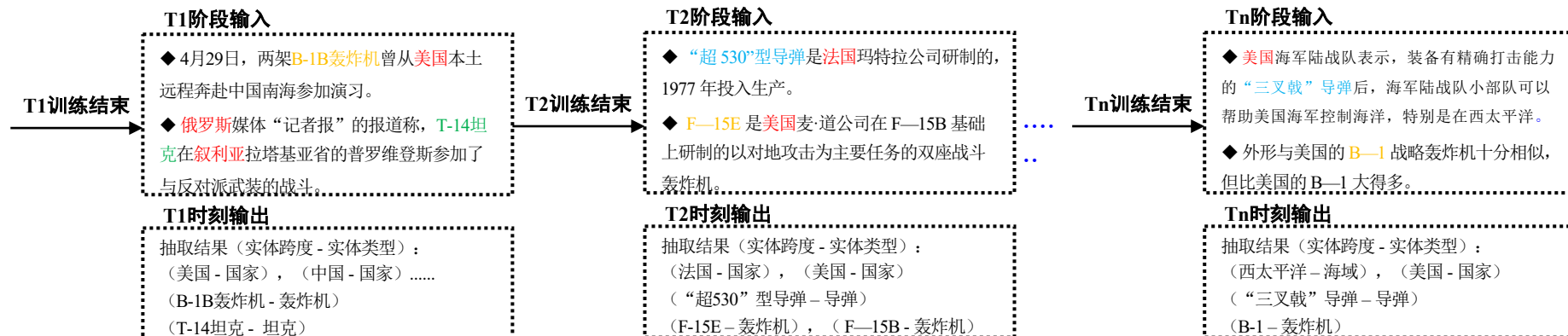
设计思路

- ◆ 模型仅完成一个训练任务
- ◆ 训练任务包含需要学习的所有实体类型
- ◆ 评价指标是Precision, Recall和F1

- ① 任务序列
- ② 评价指标

- ◆ 模型按照顺序完成一个任务序列，序列包含多个不同任务
- ◆ 每个任务包含部分的实体类型，任务间实体类型不重叠
- ◆ 引入遗忘率，前向迁移和后向迁移等更丰富的评估指标，更加全面和准确的评价模型的持续学习能力

示例



धन्यवाद
Hindi

多謝
Traditional Chinese

ขอบคุณ
Thai

Спасибо
Russian

Thank You
English

Gracias
Spanish

شكراً
Arabic

Grazie
Italian

Danke
German

Obrigado
Brazilian Portuguese

Merci
French

多谢
Simplified Chinese

நன்றி
Tamil

ありがとうございました
Japanese

감사합니다
Korean