

2022 东南大学计算机科学与工程学院

第二届国际暑期学校项目总结报告

Final Report of SEU CSE International Summer School Program

项目主题 Theme

感受多元 AI 前沿技术

Experiencing Frontier Technologies of Artificial Intelligence

一、 项目介绍

本项目是由东南大学计算机科学与工程学院组织，通过开设人工智能领域计算机视觉、知识图谱、人机交互等方向的研讨课程及基础的专业阅读与写作课程，旨在培养学生人工智能方向学习兴趣及基本的科学研究能力。每门课程均包括理论部分、经典案例实践，由院内有长期海外学习和工作经历的老师为主授课，邀请海外相关领域知名专家联合授课。本项目共开设 4 门课程，每门课程 2 学分，项目实施时间为连续 4 周。项目拟向考核优秀的同学颁发荣誉证书。

二、 项目执行情况

感受多元 AI 前沿技术国际暑期学校于 2022 年 8 月 8 日到 9 月 4 日之间举办。该项目是东南大学 2022 年（国际）暑期学校项目之一，由东南大学教务处和东南大学计算机科学与人工智能学院共同举办。项目为期四周，上课模式采取线上线下相结合模式：院内老师课程主要采用线下教学，海外专家课程主要采用线上会议。根据每一门课程的内容及特点，我们的院内老师与海外专家一起做了不同的课程设计。如感知与人机交互，考虑到课程人机交互的特点，本次课程在实践环节设计上增加了应用性和趣味性，根据课程需求采购了实验设备，使学生在课堂中“玩中练”，在“练中学”。机器视觉与应用、知识图谱及应用两门课的专家们则是结合这两个领域的前沿技术研究及在人工智能领域的应用，通过经典应用实践让学生更好的理解技术，同时通过前沿技术领域的论文阅读及研讨，使同学们对人工智能研究有了更高层面的认识，激发了同学们的研究热情。专业阅读与写作则是一门为学术科研设计的基础课程，从选课人数看，一如既往的受到了同学们的广泛欢迎。本次国际暑期课程的选课情况如下，感知与人机交互 37 人，机器视觉与应用 32 人，知识图谱及应用 42 人，专业阅读与写作 121 人。

1) 感知与人机交互 (全英文、研讨)

本课程旨在使学生深入了解人机交互的内涵,人机交互的研究内容及发展趋势,熟悉人机交互领域的一些典型技术,如虚拟现实、动作捕捉、机器人等技术,掌握人机交互界面构造的一般性方法和评估方式。本课程理论部分由计算机科学与工程学院丁玎副教授主讲,同时邀请香港理工大学钟俊培助理教授、荷兰代尔夫特理工大学助理教授 Ujwal Gadiraju 带来人机交互相关应用实例讲座。通过课堂学习、研讨交流与分组项目实践,学生从真实的科研案例和动手项目开发实践中学习人机交互前沿知识,从各场景典型应用中增强对人机交互的综合认识与理解,为未来从事人机交互、交互设计、人工智能等相关交叉领域的科研工作打下坚实基础。

丁玎老师围绕人机交互这一主题,从人机交互的定义与发展历史、代表性应用与存在问题、人机交互涉及的认知心理学、人因工程理论知识、需求发现方法及交互设计模式、人机交互评估方法及技术等多个维度,讲授了人机交互涉及的核心知识、关键技术和典型应用,通过展示大量技术实例和前沿项目趣味视频,增强学生对课程的兴趣度;通过课堂上以项目为导向的小组讨论、分享与研学,增强学生课程的参与感;并通过课后作业、小组项目合作及最后的小组汇报演讲促使学生将课堂所学知识实践应用,培养其协同合作、团队沟通及解决问题的能力。教学内容由浅入深,理论与应用结合,授课方式轻松活泼,课堂氛围良好。

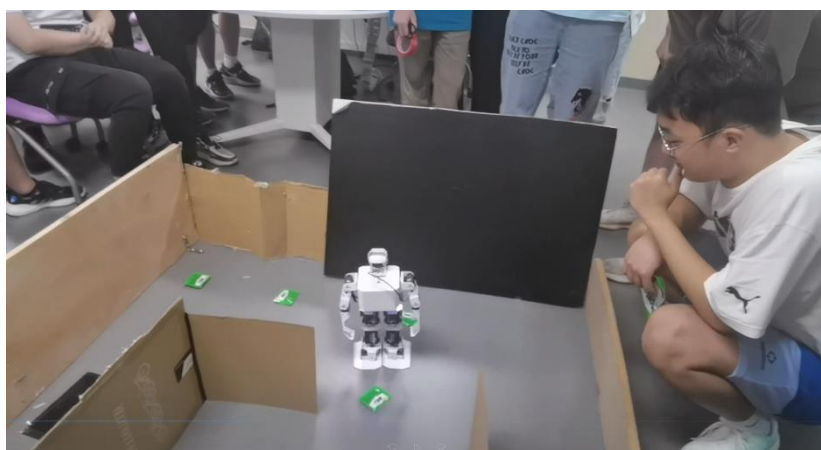
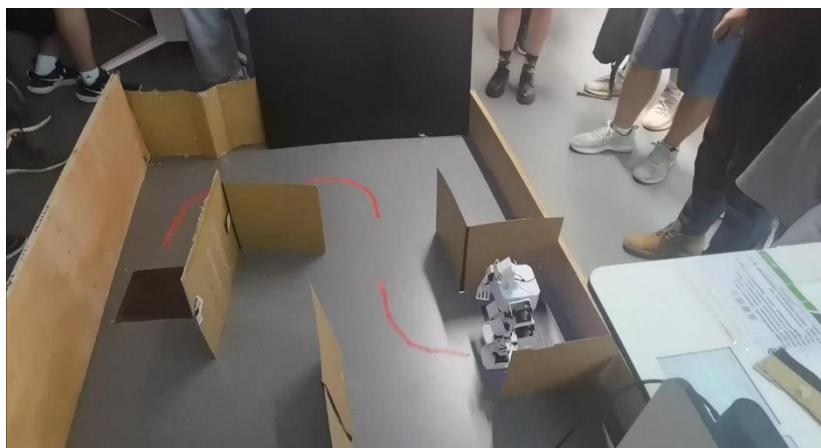
针对去年暑期课程实践环节相对简单且偏理论设计这一问题,今年特意加强了实践环节,精心设计了两大课程实践项目: 1. 基于虚拟现实和动作捕捉技术的考试恐惧体验系统: 疫情的冲击导致传统的线下课堂教育、考试转为线上进行,而随着疫情好转,学校逐步开放线下考试,然而由于长时间缺乏线下考试的经历,部分学生容易出现“考试恐惧症”,不能在考试中发挥日常所学,影响了考试的成绩。本项目基于虚拟现实和动作捕捉技术,希望深度还原考试场景,帮助学生迅速适应正常线下考试,具有很强的实际应用背景。2. 基于计算机视觉的机器人迷宫导航: 机器人寻路可以应用于众多实际场景中,如灾害现场搜救以及逃离、野外勘探、复杂管道作业等。本项目基于计算机视觉技术,通过迷宫模拟复杂地形,实现机器人的准确、快速导航寻路,对同学们的算法理论、实践能力都有较

高要求。4-5 位同学组成一组，在实际开发中运用课堂学到的以及研讨环节掌握的知识，并最终安排虚拟现实样片展示和机器人实机比拼环节，围绕“考试恐惧症暴露疗法”和“迷宫导航”两大趣味主题，让同学“玩中练”，在“练中学”。

其中参与虚拟现实和动作捕捉项目的小组有的着重场景搭建，有的关注复现真实完整的考试流程。在考试的开始和结束时刻都设置铃声作为提醒，监考老师时不时走到考生身边，对于考生的手部动作捕捉也力争精细，综合这些特点以便给使用者带来身临其境的沉浸感，发挥了虚拟现实技术的优势。参与机器人项目的小组有的另辟蹊径，不同于其他小组使用小球、有色胶带等标志物，他们使用了外包装色彩鲜明的零食作为机器人的识别对象。在实机比拼环节，他们的“贪吃机器人”追逐着提前放置好的零食，顺利走出了迷宫，取得了良好的成绩。

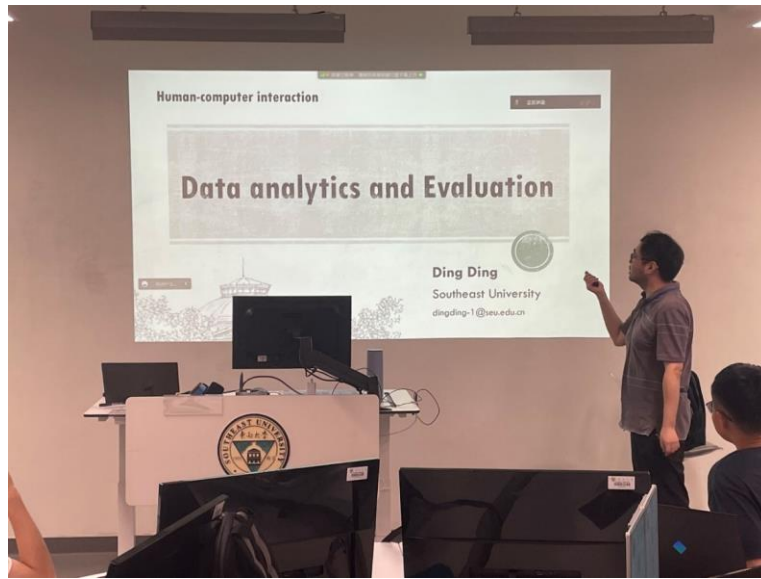


虚拟现实考试恐惧体验系统开发





机器人迷宫导航实机比拼



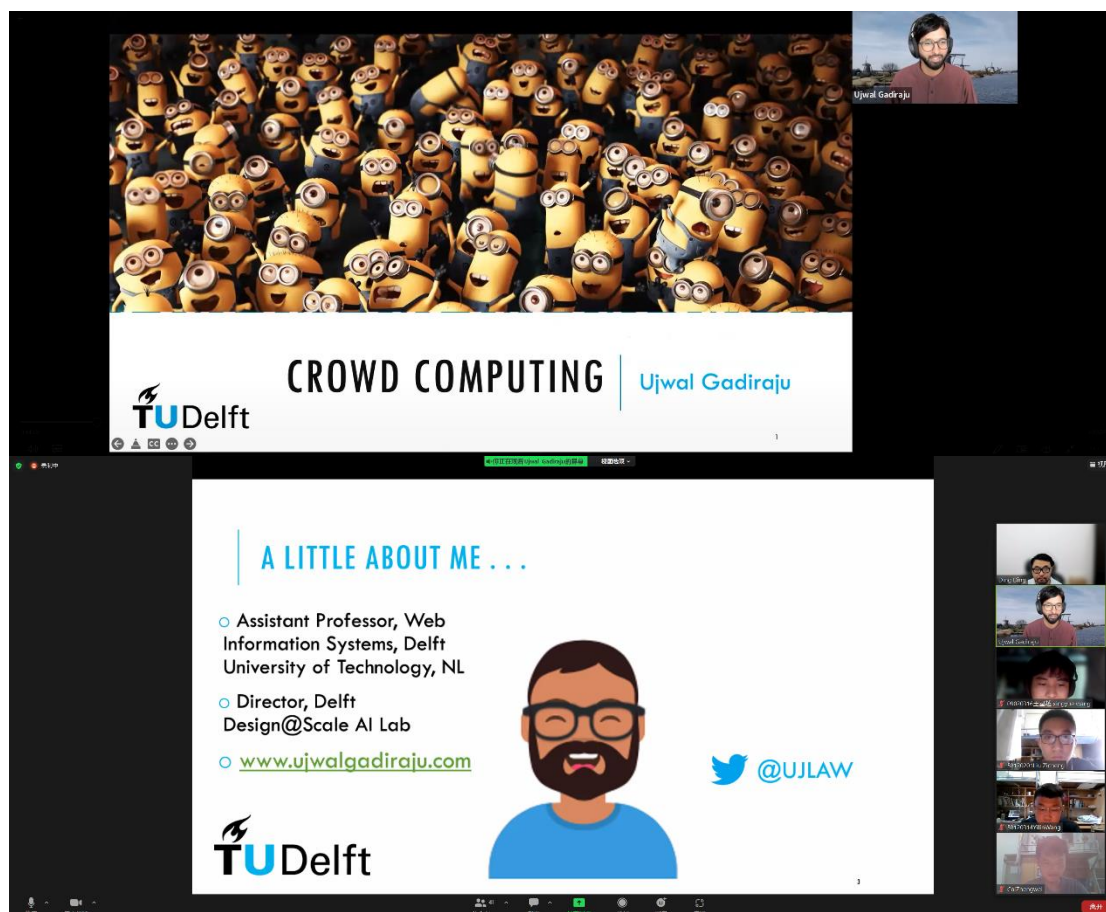
丁玘老师课堂授课实况

香港理工大学钟俊培助理教授以机器人这一典型的人机交互应用技术为切入点，介绍了人与机器人交互的理论和关键技术，重点讲授了其中的对话系统、人-机器人交互的不确定性问题，通过分享众多精彩的视频示例，让同学们对该领域的技术及研究热点有了清晰生动的认识。



钟俊培老师在线授课实况

荷兰代尔夫特理工大学助理教授 Ujwal Gadiraju 老师分别讲授了人机交互领域的关键技术--对话界面以及最新研究热点--群智计算，通过结合有趣的案例和图例，自身的最新研究成果，深入浅出的介绍了对话界面和群智计算的概念、理论、技术和研究点。

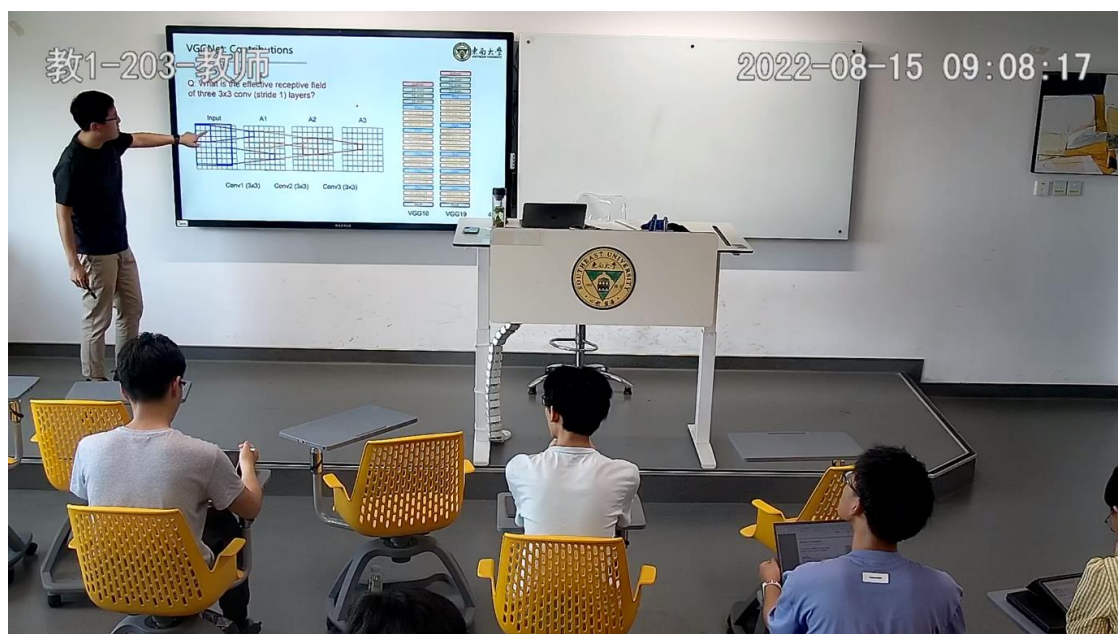


Ujwal Gadiraju 老师在线授课实况

2) 机器视觉与应用(研讨)

本课程围绕前沿的计算机视觉和机器学习领域问题，结合学术界研究与工业界应用进行技术讲解，使学生掌握机器视觉的基本概念和理论方法。本课程由计算机科学与工程学院周毅副教授主讲，同时邀请新加坡 A*STAR 高级科学家付华柱博士、英国牛津大学沈钰明博士分别从医学图像分析及基于内容的大规模图像检索技术两个应用领域带领学生深入探索并动手实践，培养学术科研能力与工程实践能力。

周毅老师首先为同学们介绍了机器视觉在当前人工智能领域的广泛应用，比如工业场景、自动驾驶场景、医学图像分析场景等，引起了学生们的浓厚兴趣。紧接着，周老师沿着课程的核心主线，讲授了不同的传统视觉特征提取方法、基于深度神经网络的卷积神经网络模型、Transformer 技术、以及目标检测、语义分割、图像生成等视觉主流任务的技术原理。结合着丰富的前沿课题推荐，同学们分为小组开展热烈的研讨，比如讨论了在智能安防中的行人重识别问题、医学图像分析中的断层图像配准问题、智能绘图中的风格迁移技术等，引发了同学们的开拓性思考。同时，结合实例，全面深入地进行剖析，使学生们对机器视觉模型开发有了一定掌握，并给出了一些实践案例，让学生练习从而更好地理解技术理论。





周毅老师课堂组织研讨实况

付华柱博士以医学影像分析与识别领域为引线，将医学图像人工智能基本背景、眼科疾病辅助诊断、胸肺疾病辅助诊断三大主题衔接起来，层层深入，环环相扣，前来学习的老师与学生受益匪浅。最后，付华柱博士为我们介绍了基于CT的新冠肺炎病灶自动分割模型案例，引发了同学们的热烈讨论。

The screenshot shows a Tencent Meeting window with a slide titled "Medical Imaging". The slide text reads: "Develop methods for combining, interpreting, visualizing, storing/accessing data from different imaging modalities, including PET, MRI, CT, ultrasound, optical, molecular imaging, and other techniques to increase their diagnostic power." Below the text is a diagram of a deep learning architecture. It starts with an "Input" of three different medical images. These pass through a "Feature learning" stage consisting of two "Convolution" and "Pooling" blocks. The output then goes through a "Classification" stage with a "Fully connected" layer and a "Softmax" layer. The source is cited as "Source: Esteva 2019, Nature Medicine". On the right side of the meeting window, there is a vertical list of participants: 周毅, Hz_FU(A*STAR), 09020138 王正今, 71120115 唐语蓓, and 71120129 陈代伟. At the bottom, it says "Hz_FU(A*STAR) 的屏幕共享".

The screenshot shows a Tencent Meeting window with a slide titled "Skin Cancer Detection". The slide displays a grid of skin lesion images categorized into "Benign" and "Malignant". It also shows a diagram of a "Deep convolutional neural network (Inception v3)" with layers: Convolution, AvgPool, MaxPool, Concat, Dropout, Fully connected, and Softmax. A central diagram shows "Training classes (757)" and "Inference classes (varies by task)". Training classes include: Acral-lentiginous melanoma, Amelanotic melanoma, Lentigo melanoma, Blue nevus, Halo nevus, Mongolian spot, Basal cell carcinomas, and Melanomas. Inference classes include: Epidermal benign, Epidermal malignant, Melanocytic benign, Melanocytic malignant, and Seborrheic keratosis. The source is cited as "Source: Esteva, Nature 2017". On the right side of the meeting window, there is a vertical list of participants: 周毅, Hz_FU(A*STAR), 09020138 王正今, 71120115 唐语蓓, and 71120129 陈代伟. At the bottom, it says "Hz_FU(A*STAR) 的屏幕共享".

付华柱老师在线授课实况

沈钰明博士以机器视觉中的大规模图像检索问题为中心，对信息检索中的无监督学习、表达学习、度量学习等算法，以及这些算法在图像检索和推荐系统中的应用，深入浅出地进行讲解，前来学习的老师与同学获益良多。最后，沈钰明博士为我们介绍了其团队在图像视频检索领域中的最新工作，引起了同学们的积极关注与讨论。

腾讯会议

录制中

Unsupervised Retrieval

- Pseudo labelling
 - CIMON: Towards High-quality Hash Codes

(a) Generation of Semantic Information: An input image I is processed by a Pre-trained ViT (ViT) to produce Features. These features are used to generate a Local View and a Global View. The Local View is processed by a Hashing Module H to produce Hash Codes H . The Global View is processed by a Confidence Matrix W to produce Semantic Information S .

(b) Consistency Learning: The Hash Codes H and Semantic Information S are used to generate a Hash Code Similarity Matrix M . This matrix is used to learn Consistency, which is split into Cross-Semantic Consistency and Intra-Semantic Consistency. The process also involves Semantic Information W_1 and S_2 , and Semantic Information W_2 and S_1 .

Dr. Yuming Shen 的屏幕共享

腾讯会议

录制中

MATRIX FACTORIZATION

Mean Squared Error

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	1	0	0	1	0
User 2	0	0	0	0	1
User 3	0	0	0	1	0
User 4	0	0	1	0	0

Utility Matrix

0.7	0	0	1.1	0
0	0	0.5	0	0.4
0.5	0	0	0.7	0
0	0	0.6	0	0.5

User Matrix

0	1.3
-0.6	0
0	0.9
-0.8	0

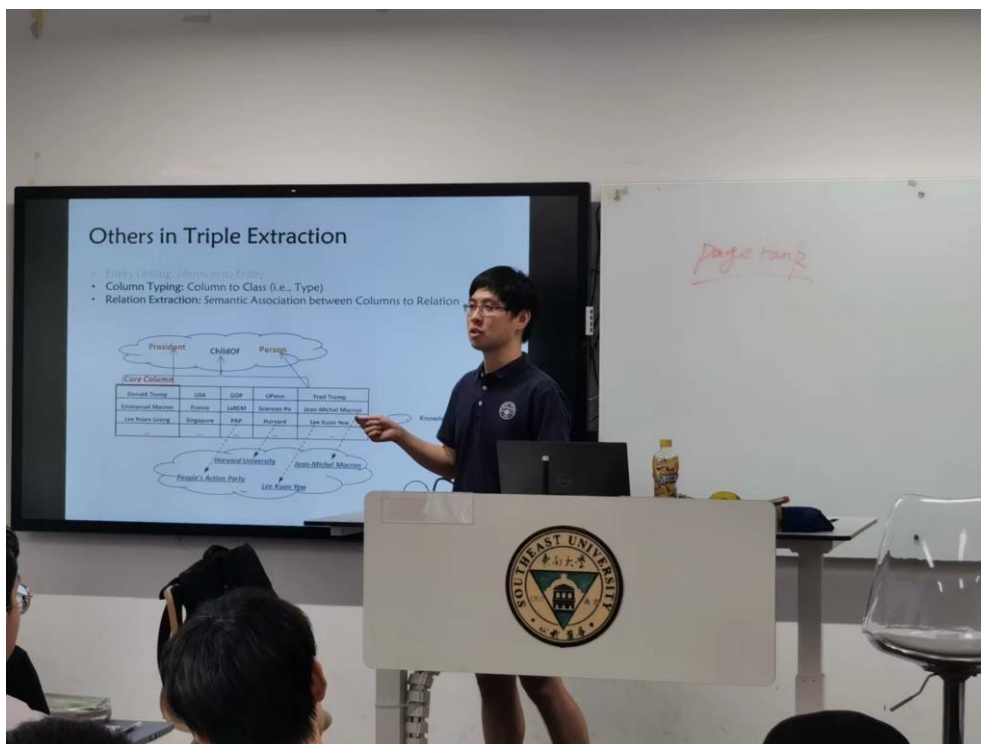
Item Matrix

0	0	-0.8	0	-0.6
0.5	0	0	0.8	0

Multiplication

Dr. Yuming Shen 的屏幕共享

沈钰明老师在线授课实况



吴天星老师线下授课实况

Guohui Xiao 老师以知识查询为授课的核心内容，从图查询语言 SPARQL 的理论基础出发，讲授了不同算子的应用场景，并将面向异构数据库查询的典型应用—基于本体的数据访问技术细节进行详细讲解，使得同学们清楚地了解在当前知识图谱技术发展的背景下知识查询的重要价值，并且理解相关的背景技术，获益良多。

SPARQL

Example VALUES

ex:Book1	ex:title	"SPARQL Tutorial".
ex:Book2	ex:title	"SemWeb".

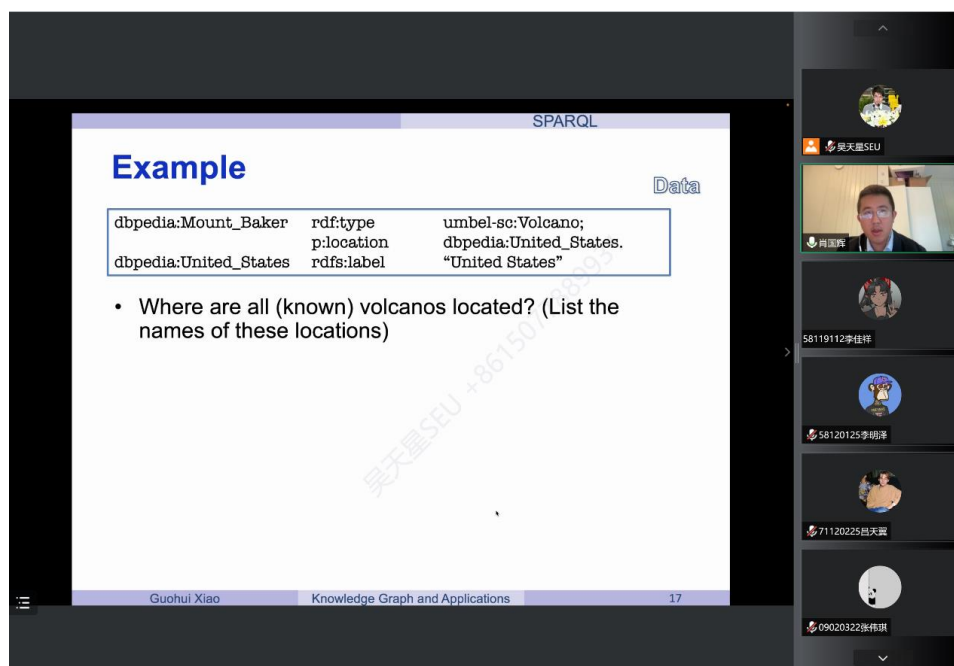
SELECT ?title WHERE {
 ?b ex:title ?title
 VALUES ?b { ex:Book1 }
}
Query

?title → "SPARQL Tutorial"
Result

正在讲话: 肖国辉
肖国辉

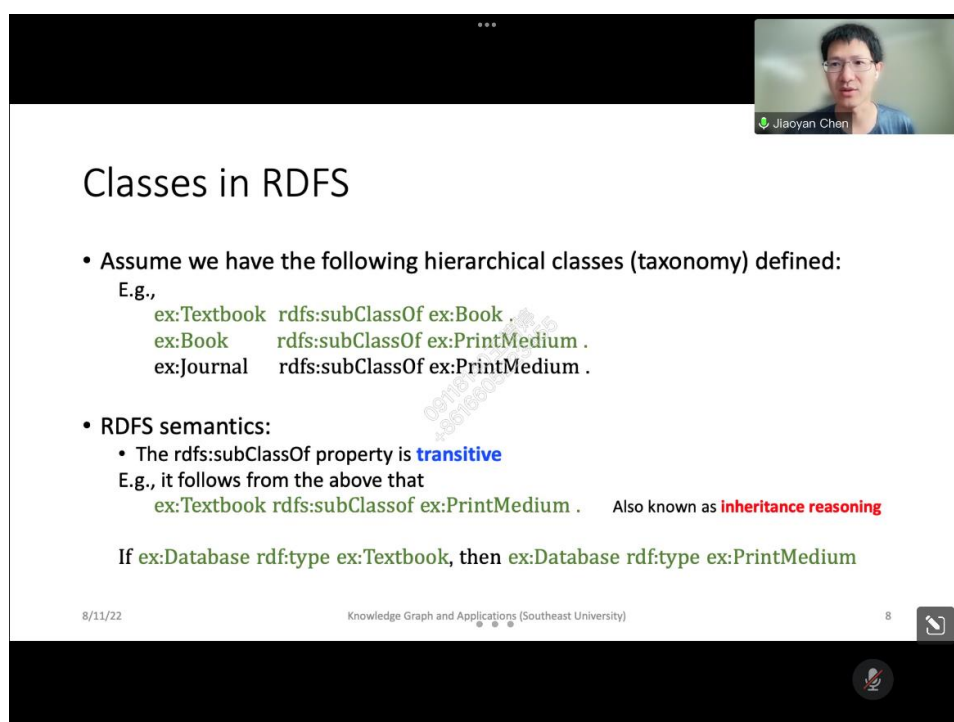
肖国辉的屏幕共享
Guohui Xiao

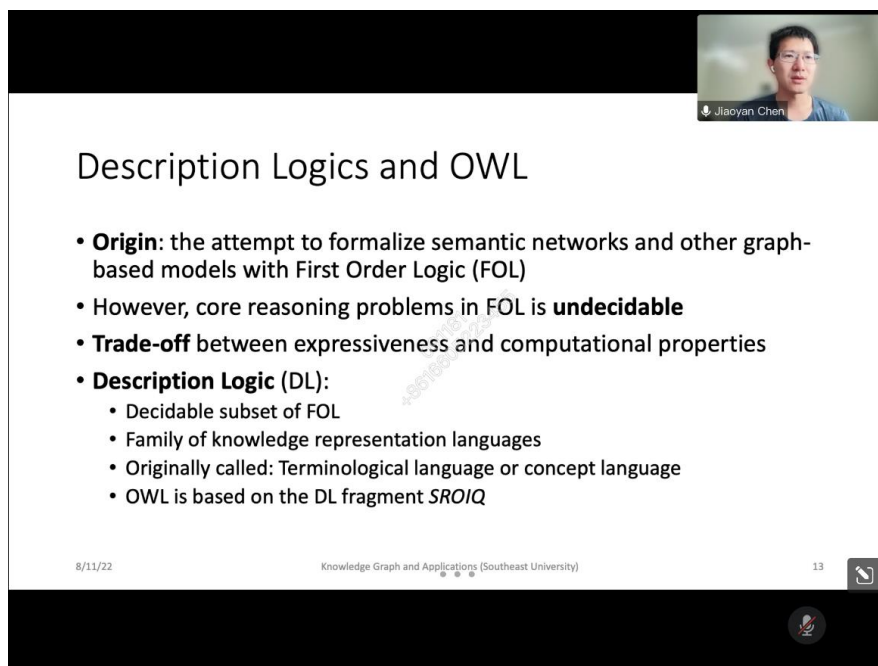
Knowledge Graph and Applications 62



Guohui Xiao 老师线上授课实况

Jiaoyan Chen 老师以知识推理为核心内容，从符号推理与统计推理的角度出发，不仅进行了生动的基础理论介绍，包括基于 RDFS、OWL 的知识推理及基于知识图谱表示学习、图神经网络的知识推理，紧扣当前的国际研究热点，介绍了本体的表示学习与推理技术，并在线上与同学们展开了热烈的讨论，使得同学们受益匪浅。





Jiaoyan Chen

Description Logics and OWL

- **Origin:** the attempt to formalize semantic networks and other graph-based models with First Order Logic (FOL)
- However, core reasoning problems in FOL is **undecidable**
- **Trade-off** between expressiveness and computational properties
- **Description Logic (DL):**
 - Decidable subset of FOL
 - Family of knowledge representation languages
 - Originally called: Terminological language or concept language
 - OWL is based on the DL fragment *SROIQ*

8/11/22

Knowledge Graph and Applications (Southeast University)

13

Jiaoyan Chen 老师线上授课实况

通过整个课程的学习研讨，同学们对于人工智能的研究方向有了全新的认识，很多同学表达了对于知识图谱技术进行深度研究的浓厚兴趣。以班级中的王玟雯同学为例，该同学在整个课程中积极认真，踊跃提问，并在研讨汇报的过程中不仅针对自己分配到的论文进行精彩汇报，并对其他组的论文工作做了充分对比分析，短短两周的课程研讨，完成了对十数篇前沿论文的汇总梳理及报告；此外，在与该同学的交流的过程中，她表达了对于未来从事知识图谱方向研究的志向，并申请了美国麻省理工学院知识图谱相关的硕博项目。课程本身不仅教授理论知识，更重要的是提升本科的同学对于人工智能专业层面更高的认识，激发研究兴趣，真正做到培育人才。

4) 专业阅读与写作（研讨）

本课程由计算机科学与工程学院王帅教授主讲，通过本课程学习，学生将提前接触科研，直观地体验研究生学习，掌握专业文献的检索技巧以及撰写和报告学术报告的能力。它包括：如何检索专业材料，如何评估专业材料，如何阅读专业材料，专业材料的通用结构，如何编写专业材料，如何制作学术风格的 PPT 以报告专业内容。通过本课程学习，学生还将掌握 GIT、Latex 等工具。

王帅教授以计算机领域写作为引线，将计算机子领域的分类研究工作、检索和阅读高质量专业论文以及学术论文的写作与汇报三大主题衔接起来，层层深入，环环相扣，使得前来学习的学生受益匪浅。研讨过程中，学生们积极地对学术报告的要点和细节进行交流与探讨，王帅教授也细致地总结学生们的讨论与理解。最后，王帅教授以典型的计算机领域顶会论文为例，为同学们做了精彩的汇报，引发了热烈的讨论。



王帅老线下课堂实况

本课程同时邀请美国佛罗里达州立大学计算机系王广助理教授，以大数据系统领域写作为引线，将物理信息融合系统、大数据分析和人工智能的研究及应用三大主题衔接起来，条理清晰，由浅入深，得到了前来听课的老师同学们一致好评。最后，王广教授为我们介绍了大数据系统中典型的城市大规模电动车系统写作论文样例，引发了同学们的热烈讨论与积极提问。

Cyber-Physical Systems (CPS)

The diagram illustrates the CPS architecture. At the top, a central cloud labeled 'Cyber' is connected to a 'Physical' domain below it. The 'Cyber' domain includes 'Computation' and 'Control' components. The 'Physical' domain includes 'Sensing' and 'Control' components. The 'Sensing' loop connects the Physical domain to the Cyber domain, and the 'Control' loop connects the Cyber domain to the Physical domain. The Physical domain is represented by a grid of images showing various physical systems: power lines, a yellow robot, a white car, a bicycle, a white scooter, a hand holding a card, and a person's face.

Technologies

The mind map details the technologies used in CPS:

- Sensing Infrastructure** (GPS, Cellphone, Wi-Fi, Beacon, RFID, Bluetooth)
- Real-Time Systems**
- Communication Systems**
- Big Data**
 - Data Analytics (Pandas/Hadoop/Spark)
 - Visualization (Python, Arc GIS, MapBox, Processing)
- AI**
 - Inference**
 - Data Mining
 - Statistics
 - Prediction**
 - ML (XGBoost, Bayesian, SVM)
 - DL (CNN, LSTM, GNN)
 - Decision Making** (Scheduling, Dispatching, Recommendation, Alert!)
- Integer/Combinatorial Optimization**
- Reinforcement Learning**

王广老师线上授课实况

三、 课堂教学反馈情况

本次暑期学校结束后,针对暑期课程的教学情况,组织了学生在线问卷调查,以便学院后续更好的组织下一届的国际暑期学校。学生对老师们的授课整体满意度比较高,也有部分同学对于课程内容提出了一些建议,比如有的同学反应课程内容比较多,课程考核工作量大,全英文授课理解有困难等。

问卷调查课程	2、老师教学态度端正,能够发挥正向价值引领作用	3、课程教学组织形式能有效促进我掌握课程内容	4、老师能有效开展课程管理,教学秩序不杂乱	5、我有困难时,能得到老师的有效指导	6、课后作业适切,我能及时得到老师的反馈	7、请您反馈对课程安排和教师教学方面的意见和建议	8、我对课程的总体评价
机器视觉与应用	10	10	10	10	10	10 无	10
机器视觉与应用	10	10	10	10	10	10 老师讲的由浅入深,条理清晰,思路明确,举例合理	10
机器视觉与应用	10	8	10	10	10	8 无	8
知识图谱及应用	10	10	10	10	10	10 无	10
知识图谱及应用	10	10	10	10	10	10 内容太多有点赶,或许可以减少一点	10
机器视觉与应用	10	8	10	10	8	10 内容把每一点讲清楚	10
机器视觉与应用	10	10	10	10	10	10 无	10
机器视觉与应用	10	7	10	10	10	8 无	9
机器视觉与应用	10	10	10	10	10	10 课程安排非常合理,在教授了基础知识后,寻找优秀的相关论文进行研讨。小组内部也能够妥善学习。唯一的不足就是分组上六人仍然容易出现较大的任务量差距,绝对平均并不一定是小组合作的最佳解。	10
机器视觉与应用	9	9	9	9	9	9 没学过机器学习的软件学生可能读论文会比较吃力	9
机器视觉与应用	10	10	10	10	10	10 无	10
机器视觉与应用	10	3	10	10	10	10 课程安排全英文其实挺不利于学生学习知识。首先老师把ppt从中文自己的思路改成英文,然后传达到学生这里,学生再翻译为自己的思路,内容接收会大打折扣。 希望老师可以适当加入中文方面的补充说明,如果课程要求上课用英文,是否可以提供中文版ppt作为参考补充。 同时研讨课不应该分布在一上午或一下午,一周五天时间完全可以平均分配,这样有消化吸收的时间,上课效率也高	5
感知与人机交互	9	8	8	8	8	9 无	9
感知与人机交互	10	10	10	10	10	10 无	10
感知与人机交互	10	10	10	10	10	10 无	10
机器视觉与应用	10	10	10	10	10	10 老师很棒,教学内容非常合适,教学过程十分耐心,在为期一个月的授课过程中我也学到了很多。	10
专业阅读与写作	10	10	10	10	10	10 无	10
感知与人机交互	9	9	9	9	10	9 无	9
机器视觉与应用	10	10	10	10	7	5 任务少点	8
感知与人机交互	10	10	10	10	10	10 无	10
感知与人机交互	10	8	10	10	10	8 无	9

问卷调查部分原始数据

项目总结

本次感受多元 AI 前沿技术国际暑期学校是我学院第二次承办的国际暑期学校，所开设课程均为计软智学院现有全英文课程，契合国家面向新一代人工智能发展战略，受到了全体学生的一致好评。本次课程根据第一届国际暑期学校反馈的理论部分偏多，案里实践方面安排偏少情况，做了课程设计上的改进，使课程理论与应用结合更紧密，课堂教学趣味性更强。课程还存在以下问题，需要我们去更进一步的思考与改进：

1、四门课程内容设计需要进一步优化：目前由于国际暑期暑期学校时间不长，但是我们的课程内容是按照长学期课程来设计的，因此在排课上存在课时比较多且集中的特点，给学生的学习带来了不小的压力。后续我们需要在课程内容设计上根据国际暑期学校的特点，做进一步改进。

2、学生生源国际化需要进一步推动：目前我们的国际暑期学校课程做到了师资的国际性，但是在生源上基本都是校内学生，后续需要在生源国际化上做更多的推动工作，比如动员院内有海外学习背景的老师给我们在国外高校进校国际暑期学校的宣传，以引入更多的海外高校学生。

我们将不断地改进提高，力争把东南大学国际暑期学校项目举办成我校的名片项目。

计算机科学与工程学院、人工智能学院

2022 年 9 月 11 日