

✦ 人才培养

文章编号: 1672-5913(2020)10-0103-05

中图分类号: G642

人工智能时代的计算机人才创新培养

窦亮, 贺樑, 周爱民, 肖春芸

(华东师范大学计算机科学与技术学院, 上海 200062)

摘要: 针对人工智能时代对计算机人才的更高要求, 介绍专业课程体系建设、人工智能必修课建设、基于 Online Judge 智能推荐的计算思维训练等人才创新培养实践内容, 探讨适合计算机专业本科生的培养之路。

关键词: 人工智能; 计算机专业; 人才培养

DOI:10.16512/j.cnki.jsjy.2020.10.021

0 引言

人工智能是计算机科学研究和应用领域最具活力的发展方向之一, 近年来愈发受到各方的关注。在政府层面, 英美等国家各自发布了人工智能政府报告, 我国政府也予以了极大的重视和支持。2017 年国务院印发了《新一代人工智能发展规划》, 描绘了未来十几年我国人工智能发展的宏伟蓝图。在之后连续三年的政府工作报告中, 人工智能产业建设被反复提起, 对人工智能的描述也从 2017 年的“加快”、2018 年的“加强”变为 2019 年的“深化”^[1]。在企业层面, Google、微软、Facebook、阿里、腾讯、百度等国内外科技企业巨头也纷纷加快在人工智能领域的布局。尽管人工智能产业呈现一片繁荣昌盛, 但目前我国人工智能整体发展水平与发达国家相比仍存在差距, 人工智能尖端人才远远不能满足需求。正如中国人工智能学会理事长、中国工程院院士李德毅所言, “中国要在 2030 年成为人工智能的世界高地, 必须现在就要进行人才培养。”国家和企业之间的竞争, 日益体现为人工智能领域专业人才的竞争。因此, 如何针

对人工智能时代对计算机人才的更高要求, 设计具有针对性的计算机专业本科生培养模式是当前亟需解决的问题。

1 计算机专业本科生培养模式现状分析

目前, 国内计算机专业的本科生培养大多采取“程序设计+算法+应用”的模式, 将数学、物理、基础编程语言作为最核心的基础课组成部分, 在高年级开设人工智能、数据处理等方面的选修课, 提供相关培训。在人工智能的时代需求下, 这种培养模式已经呈现出一些问题。

1) 不利于尽早培养学生的人工智能和数据思维能力。

数据处理和人工智能相关课程属于高年级选修课程, 学生如果不选修这些课程, 就无法深入了解这些领域。传统的计算机专业本科生培养模式, 重点还是理解冯·诺依曼体系结构, 重在打基础; 而人工智能和大数据时代的计算机人才, 除了编程能力和算法能力外, 还需要尽早掌握数据处理和机器学习相关的方法, 即培养以数据为驱动的人工智能思维。

基金项目: 上海高校本科重点教学改革项目 (沪教委高 [2017]71 号); Google 支持的教育部 2018 年第一批产学研合作协同育人项目 (教高司函 [2018]47 号); 国家青年科学基金项目“面向 OJ 训练的适应性学习路径推荐研究” (61907015)。

作者简介: 窦亮, 女, 副教授, 研究方向为软件开发方法与技术, ldou@cs.ecnu.edu.cn; 贺樑 (通信作者), 男, 教授, 研究方向为知识处理与数据智能, lhe@cs.ecnu.edu.cn。

2)不利于培养学生对接新时代的创新能力。

传统培养模式更多的是培养学生编写小型应用程序的技术能力和一定的工程素养,如数据库应用系统开发、Web应用构建等,但这些方法和技术早已被市场框架化和工程化,面向新时代“智能+”产业的创新需求,学生所掌握的技能 and 能力尚不足够。

计算机以强大的生命力飞速发展,但反观计算机专业教学、培养方式和课程体系的发展明显滞后于新技术的发展速度^[2]。

2 计算机人才创新培养模式实践

2.1 优化课程体系,改进培养方案

调整优化计算机专业的本科生课程体系,在培养方案中形成以人工智能为核心的课程群,具体包括以下几方面。

(1)在一年级上学期增加开设专业必修课编程导论,课上使用Python作为新生入门编程语言。利用Python语法简单、关注问题本身、支持库成熟丰富等特点,使学生掌握编程思维和技巧,并通过上机实验为学生储备初步的数据处理与分析的能力。

(2)将人工智能由专业选修课改为专业必修课,开设时间由原来的三年级上学期改为二年级下学期,优化调整原有课程内容,使之成为适合计算机专业本科生二年级学期学习的专业必修课。

(3)在保留传统计算机学科基础课程基础上,以人工智能课为基础主线,发展出三大方向和配套课程,主要开设在高年级,如图1所示。

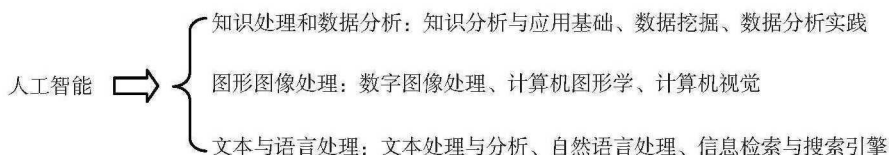


图1 人工智能课程群

整体课程体系遵循“引入—基础—专业—实践—深究”的思路,总体设置为专业基础课、专业必修/选修课、实践环节3类:专业基础课程在低年级开设,旨在打好基础,训练基本的编程能力,依托Online Judge在线判题平台,培养计

算思维;专业必修/选修课在高年级开设,旨在方向钻研,建立完善的理论和应用体系;实践环节则包括课程实践、科创活动、毕业论文等。

2.2 以算法为中心,建立人工智能全局观

当前人工智能的三大发展支柱是摩尔定律描述的算力指数增长,互联网和物联网爆发式增长产生的海量数据,以及智能算法的快速发展^[3]。大数据和算力可以看成是“硬件”,政府和企业都可以做到;建模算法可以看成是“软件”,需要学校来完成^[4]。算法和模型是促进人工智能技术在产业中真正落地的重要保障。因此,我们认为计算机专业的“人工智能”课程教授应该以算法为中心的,要让学生理解算法之理、掌握算法之用和感受算法之美。理解算法之理是要了解什么是算法、算法的运行机理是什么,以及如何实现一个算法;掌握算法之用是要了解如何运用算法解决实际问题,优化改进或者设计一个算法;感受算法之美则是在运用算法解决实际问题的体验中激发学生的学习兴趣。

由于人工智能的主流已从传统的符号系统为主发展到统计机器学习为主,在这样的背景下,调整原人工智能课的教学内容结构,增加遗传算法等启发式搜索、约束满足问题求解、神经网络等内容,适当缩减符号计算内容的讲授比例,形成以搜索和问题求解、机器学习初步、知识表达和推理3个方向比较均衡的讲授内容。人工智能课程作为一门二年级学生的专业必修课,课程目标是要让学生建立基本的人工智能全局观,以算法为中心,要求学生熟练掌握每个大类中的1~2个典型算法,具备基本的算法实现和应用能力,

学会通过人工智能初步解决实际问题。课程的大纲及课时分配见表1。

在课程实验部分,引入Google的开源人工智能学习系统TensorFlow作为综合性实验平台,要求学

生能够复现实验内容并尝试改进效果,使学生掌握使用TensorFlow的基本方法。人工智能课程的实验内容见表2。

2.3 使用智能推荐技术,有效训练计算思维

Online Judge(简称OJ)是业内著名在线

表 1 人工智能课程大纲和课时分配

方向和总课时	序号	章节	主要内容	简介
总论 2h	1	总论	人工智能基本概念、主要研究内容和现状	介绍人工智能的基础知识,包括人工智能基本概念及检验方法,人工智能的目前主要的研究内容,人工智能的发展历程、现状及未来的趋势
	2	一般图搜索	一般图搜索策略、无信息搜索	针对基本搜索问题,介绍图搜索求解方法,主要包括回溯策略,图搜索策略的相关概念及一般算法框架,深度优先和宽度优先等无信息图搜索算法
	3	启发式搜索	启发式搜索 A 算法、A* 算法、AO* 算法	针对一般搜索问题,介绍启发式的图搜索策略,主要包括启发式图搜索的基本概念、A 算法框架及各种衍生算法、A* 算法相关理论分析、与或图及 AO* 算法框架
	4	不确定性搜索	局部搜索、模拟退火算法、遗传算法	针对复杂搜索问题,介绍带有不确定性的图搜索策略,主要包括组合优化及相关概念、局部搜索算法框架及分析、模拟退火算法原理及框架、遗传算法基本概念及框架
	5	博弈搜索	MiniMax 搜索、Alpha-Beta 剪枝、蒙特卡罗树搜索	针对博弈问题,介绍博弈搜索算法,主要包括博弈的基本概念、基本的 MinMax 搜索算法、Alpha-Beta 剪枝算法及分析、蒙特卡罗树搜索基本概念及算法
	6	约束满足问题	约束满足回溯搜索和局部搜索等	针对约束满足问题求解,介绍约束满足问题求解方法,主要包括约束满足问题定义、求解约束满足问题的一般搜索算法、改进搜索算法及启发式算法
搜索和问题求解 10h	7	机器学习基础	机器学习背景、常见机器学习方法	介绍机器学习基础知识,包括机器学习的相关背景、常用的机器学习方法、机器学习目前面临的困难和挑战等
	8	有监督学习	决策树基本概念、决策树常用算法	针对监督分类问题,介绍决策树相关基础知识,主要包括决策树的基本概念、决策树 CLS 及 ID3 等基本算法、监督学习中的基本问题及解决方法
	9	无监督学习和半监督学习	聚类方法概述、划分聚类方法、层次聚类方法	针对无监督聚类等问题,介绍聚类方法的相关基础知识,主要包括聚类方法概述、划分聚类方法、层次聚类方法、聚类方法应用举例
	10	神经网络和深度学习、强化学习	感知器、前馈神经网络、卷积神经网络初步	针对分类回归等问题,介绍人工神经网络相关基础知识,主要包括人工神经网络概述、基本神经元模型、多层网络的反向传播算法、人工神经网络应用举例、深度神经网络基础等
机器学习初步 10h	11	知识表达	演绎系统、产生式系统、框架结构、语义网络、过程知识表示、本体及应用	介绍知识表达的基本概念和基本方法,主要包括知识的定义、知识的要素、知识表示分类及常用的知识表示方法,如逻辑表示、产生式规则表示法、语义网络表示法、框架表示法等
	12	基于谓词逻辑的机器推理	命题逻辑、谓词逻辑基础、归结原理	介绍基于谓词逻辑的机器推理方法,主要包括相关基础介绍、命题逻辑及推理、谓词逻辑归结子句形、归结原理、Herbrand 定理等
	13	不确定性推理	概率论基础、确定性方法、不确定方法	介绍基于概率的不确定性推理方法,主要包括相关基础介绍、主观贝叶斯方法、确定性方法、D-S 证据理论等
	14	贝叶斯网络	概率基础、贝叶斯网络、贝叶斯网络学习与应用	针对一般学习问题,介绍贝叶斯网络相关基础知识,主要包括贝叶斯网络概述、贝叶斯概率基础、贝叶斯网络概述、贝叶斯网络的预测、诊断和训练算法
	15	知识图谱初步	知识图谱概述、知识图谱构建、知识图谱质量控制	介绍知识图谱相关概念及方法,主要包括知识图谱的由来、知识图谱的发展历程、知识图谱应用、实体抽取、关系抽取、百科网站抽取、知识融合、知识图谱精化等
知识表达和推理 10h				
应用 2h	16	人工智能应用	人工智能应用介绍	介绍人工智能的常见应用领域,主要包括自然语言处理、多智能体、机器人、跨媒体智能和物联网等

表2 人工智能课程实验内容

序号	实验名	实验要求
1	路径导航	要求输入节点坐标及连接关系文件,构造距离矩阵,采用最短路径算法、A*算法及局部搜索算法实现路径导航
2	五子棋	实现交互式五子棋程序,采用博弈算法实现,可以采用极大极小搜索方法或者 alpha-beta 剪枝算法实现
3	人脸识别	对人脸样本采用 PCA 降维方法提取特征(前 40 维),采用最近邻分类器、BP 分类器或 SVM 分类器识别人脸,给出分类结果,画出识别率曲线
4	图像分割	给定一幅图像,指定类别数,用 K-means 聚类或者模糊 C 均值聚类算法(FCM)实现图像分割(聚类)
5	关系抽取	使用 Pytorch 建立并训练神经网络模型,预测给定 (e1, e2, description) 三元组对应的实体关系
6	预测 OJ 题目通过率	通过数据清洗、设计特征,并使用 Tensorflow 训练神经网络模型,预测 OJ 中某用户对某题目的通过率
7	垃圾邮件分类	使用 Tensorflow 训练融合了知识与语义的卷积神经网络 KCNN 模型,将文本的语义表示和知识表示融合起来形成新的嵌入表示,在真实的垃圾邮件语料库 trec06c 中进行垃圾邮件分类

编程判题形式,用来检测用户提交的计算机程序源代码的正确性。本专业建设的 OJ 平台称为 EOJ^[5],目前支持 C、C++、Java、Python、Pascal、PHP 等 10 多种语言的编程解题,属于国内领先的、支持语言较多的 OJ 平台之一,最初是为了训练本科生参加程序设计竞赛而设计开发的。学生通过在平台上完整的编程过程,理解了抽象、概括、问题解决等计算思维的本质,能有效锻炼实际的算法设计与分析能力、程序设计与实现能力,于是本专业将大部分专业课程的上机练习和考试建设在了 EOJ 中,既真实考察学生的编程能力,又方便教师的教学管理。目前,本专业依托 EOJ 进行教学辅助的课程包括程序设计基础、面向对象 C++、面向对象 Java、数据结构、算法分析与设计、问题求解与程序设计等。

对多数的计算机专业学生而言,必须在 OJ 上做大量的题目,才能巩固基础,得到充分的计算思维能力训练,培养出必备的专业能力。要想在 ACM-ICPC 等程序设计比赛选拔中脱颖而出,学生更需要花费大量的时间在 OJ 系统上“刷题”。现有的 OJ 系统上存在着几千道题目,有编程基础题目,也有针对不同知识点的题目,还有程序设计竞赛真题,题目类型多样,难度跨越很大,导致学生做题过程中易存在盲目性,出现“认知过载”和“信息迷航”问题,降低了学生的学习效率,甚至影响了学习兴趣。为解决这个问题,我们在 EOJ 上实现了知识驱动的智能推荐,可以

根据学生已提交解题题目的情况,按题目难度或知识点情况智能化地给学生推荐题目,以达到有效训练和提升学生的编程水平的目的。

在 EOJ 上为计算机学科的知识点构建了树型模型,表达了知识点之间的层次关系。在 EOJ 中使用标签来对题目进行知识点标注,允许受信任的学习者修改相应题目的标签信息,目前在 EOJ 上 2 630 道有正确提交的题目中,已对 942 题进行了知识点标签。依据通过题目的学习者数和尝试数以及他们各自所占比例,计算出题目难度,再通过难度向量计算,可以得到学习者在各个知识点上的表现。目前在 EOJ 上实现了 3 种类型的推荐。

(1)按知识点分类,其中又分为推荐已掌握的知识点和接触较少的知识点,提供了学习者两个学习方向,即巩固较为熟练的知识点和拓展不熟练的知识点。

(2)按难度分类,其中又分为难度适中、难度偏低以及难度偏高的题目,即所谓的刷“水”题和挑战难题。

(3)按题目近期热度分类,其中又分为近期提交较多的“热”题和长时间无人问津的“冷”题。

EOJ 的智能推荐功能上线一段时间以来,受到了学生的广泛欢迎,初步统计,通过智能推荐进入解题页面的占比已达 70%。题目智能推荐构建了以学习者需求为中心的个性化学习支持服务与环境,符合现今教育手段与学习方式深度变革

的时代要求。

3 结 语

华东师范大学计算机科学与技术学院针对人工智能时代新形势下计算机专业的新要求,优化课程体系和建设核心课程,探索计算机人才

创新培养模式,取得了较好的实践效果。通过面向人工智能时代的计算机专业人才创新培养实践,学院在多年的实践中探索出了一套较为适合本校学生实际情况的培养方式。希望本文能对其他高校计算机专业的课程建设和培养模式等起到一定借鉴作用。

参考文献:

- [1] 中国机器人网. 政府工作报告连续三年提人工智能: 人工智能产业将迎来爆发期[EB/OL]. (2019-03-06)[2020-05-01]. <http://www.robot-china.com/news/201903/06/55937.html>.
- [2] 李翠平, 柴云鹏, 杜小勇, 等. 新工科背景下以数据为中心的计算机专业教学改革[J]. 中国大学教学, 2018(7): 22-24.
- [3] 张首晟. 人工智能三大发展支柱[J]. 中国经济信息, 2017(14): 80.
- [4] 郑立刚. 基于应用能力培养的“机器学习”课程建设思考[J]. 现代计算机, 2018(21): 66-69.
- [5] 华东师范大学. ECNU Online Judge[EB/OL]. [2019-08-05]. <http://acm.ecnu.edu.cn>.

(编辑: 史志伟)

征稿启事

本刊现向广大读者征集关于提高计算机专业核心课程教学效果的文章, 具体要求如下。

一、论文议题

如何提高计算机专业核心课程(如程序设计基础、计算机组成原理、计算机系统结构、数据结构、操作系统、计算机网络、数据库等)教学效果。

二、论文要求

1. 文章必须是未公开发表过的原创作品, 字数 5000 字左右, 题目自拟。
2. 文体不限, 可以是理论文章, 也可以是优秀的教学案例。
3. 文章应包括题目、摘要(200 字左右)、关键词、正文、参考文献、基金项目和作者简介(第一作者的姓名、性别、职称、研究方向、详细通信地址、电话及 E-mail)。
4. 来稿请统一用 Word 格式, 并在文件名处写明“核心专业课程”。
5. 论文参考格式可上《计算机教育》杂志社的网站下载: www.jsjy.com。

三、投稿方式

投稿邮箱: www.jsjy.com。

电话咨询: (010) 62770175-3402 / 3404 / 3405。

四、论文评审与发表

1. 《计算机教育》杂志社及外审专家组对论文进行评审。
2. 通过的论文将刊登在《计算机教育》杂志相关栏目中。
3. 评审及发表流程与网站投稿论文相同。