doi:10.3772/j.issn.2095-915x.2017.01.009

面向机器人竞赛的文献分析与主题词提取

1. 北京师范大学第二附属中学　北京　100088；2. 北京工业大学信息学部　北京　100124
刘涵[1]　刘京旋[2]

摘 要 近年来，机器人教育在我国愈发受重视，针对机器人的竞赛也受到更多中小学、高校的教师与专家的关注。本文通过查阅机器人竞赛相关文献，介绍国内外机器人教育现状，分析国内的机器人教育文献，针对机器人竞赛的文献采用基于统计的方法，自动抽取领域主题词，获取其术语分布。

关键词：知识图谱，表示学习，规则挖掘，推理
中图分类号：G35

The Literature Analysis and Thematic Words Extraction of Robot Competition

1. The Second High School Attached To Beijing Normal University, Beijing 100088, China;
2. College of Information Science, Beijing University of Technology, Beijing 100124, China
LIU Han[1]　LIU Jingxuan[2]

Abstract In recent years, more and more attention were paid to the robot education in our country, and the robot competition attracted much more attention of the teachers from primary and secondary schools, and universities and experts. Through consulting document retrieval of the robot competition, this paper described the present situation of the robot education around the world, and introduced the robot education in China. In this paper, we also presented the distribution of thematic words on robot competition by using the automatic extraction method based on statistic.

Keywords：Robot competition, robot education, subject words extraction

作者简介：刘涵（1999–），学生，研究方向：机器人竞赛，信息处理，email：liuhanalice@outlook.com；刘京旋（1992–），硕士研究生，研究方向：数据挖掘，自然语言处理。

doi:10.3772/j.issn.2095-915x.2017.01.009

1 引言

近年来，伴随经济、网络技术、数字化、人工智能等多项技术的快速发展，包括中国在内的一些国家纷纷将机器人发展视为科技创新和经济增长的动力。随着计算机、电子、网络、机械等多个学科的综合发展，机器人教育逐渐走入中小学生课堂，机器人竞赛也成为促进科技教育和创新教育的方式之一。

美国是最早开展机器人教育的国家，现在各州都设立了自己的机器人课程教育标准，目前美国机器人教育已经发展到了家庭教育层面；紧随其后，日本也注重发展机器人教育，欧洲国家则以玩具切入机器人教育。众所周知，目前日本的机器人技术发展在国际社会处于领先地位，工业机器人产量、安装量世界第一，在科研上已经可以研制出鲤鱼翻身等模仿肌肉动作的机器人，这不仅归功于他们深厚的工业基础，更和对机器人教育的重视密不可分。

在日本，不仅在中小学教育大纲中有机器人教育课程，每所大学都有高水平的机器人研究和教学内容，并定期举办机器人竞赛促进机器人技术的发展；自 1992 年开始，美国政府有关部门在全国推行“感知和认知移动机器人”计划，高中生可免费领取 70 公斤重的零件，自行组装遥控机器人，并参加相关竞赛。

2 国内外机器人教育现状

目前，在众多计算机教育中心和许多著名专家的推动下，我国多个省份与城市开展了机器人教育。教育机器人逐渐成为中小学生技术课程和综合课程的良好载体，帮助学生在实践与创新上取得优异的成绩。自 2001 年中国科协举办了首届中国青少年电脑机器人竞赛，从此拉开了国内中小学机器人比赛的序幕 [1]。2003 年教育部颁布的《普通高中技术课程标准（实验）》，人工智能教育成为了科技技术科目中的选修模块 [2]，人工智能教育在我国教育层面已经逐步走向普及化。2012 年发布的《基础教育信息技术课程标准》[3] 中在小学和初中部分中都包含了机器人模块，在基础教育中开展机器人教育，为我国科技教育开辟一条新的道路，有利于探索教育改革和人才培养，提高学生的科技创新和自身实践能力。

目前国际及国内机器人竞赛种类繁多，1994 年美国三一学院 Jack Mendelssohn 首先提出用机器人灭火的想法，之后同一批国际著名机器人学家发起创办了机器人灭火竞赛；1995 年由韩国科学技术学院的金钟焕教授提出了机器人足球竞赛，之后在韩国政府的支持下举办了微型机器人世界杯足球赛 (FIRA MIROSOT’96）；1997 年在日本组织了国际性机器人足球比赛（Robot World Cup），中国科技大学蓝鹰机器人足球队以 4:0 的成绩战胜日本 Helios 队，获得第 19 届 RoboCup 机器人世界足球赛仿真 2D 冠军，第 20 届 RoboCup 于 2016 年在中国安徽举行；1998 年国际机器人奥林匹克委员会（IROC）成立，旨在使更多的青少年有机会参与国际科技交流活动，展示沟通机器人技术，激发对机器人技术的热情与鼓励对机器人研究的探索。

3 基于机器人竞赛教育的文献分析

3.1 文献数据统计分析

在中国知网（CNKI）查找全国优秀硕士博士学位论文全文数据库、中国期刊全文数据库，全国及国际会议全文数据库和全国报纸全文数据库，截止到2016年12月，分别通过关键词“机器人竞赛”，“机器人教育”，“中小学机器人教育”，“中小学机器人竞赛”按主题进行检索，结果如表1。

表1　机器人竞赛相关文献数据

	机器人竞赛	机器人教育	中小学机器人教育	中小学机器人竞赛
中国硕博学位全文论文数据库	142	307	51	17
中国期刊全文数据库	601	1388	132	52
中国会议论文全文数据库	28	78	5	0
中国报纸全文数据库	271	303	14	17

根据检索数据结果分析，目前机器人教育获得了越来越多的重视，不过从检索内容上看，“机器人竞赛”相关的文献相对于“机器人教育”的文献偏少，而有关“机器人基础教育与竞赛”的文献相比“机器人教育和竞赛”的文献更少，加大机器人基础教育研究对于机器人技术进一步普及和未来发展具有重要意义。

在中国知网（CNKI）查找全文数据库，从2006年截止到2016年12月，查看近10年不同年限发表的机器人相关文献数量，结果如图1。

根据检索时间结果分析，近十年我国相关机器人的文献逐年增多，我国在机器人方面的研究地位逐步提升，人们对机器人的研究工作也逐步加大，这也意味着我国的科技水准已经达到了一定水平。

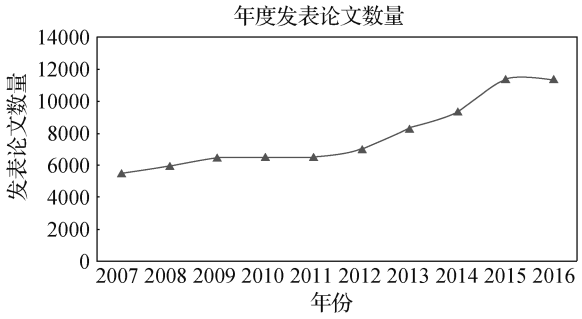

图1　机器人相关文献数量分布

有关机器人竞赛的文献主要分为：机器人竞赛的教育意义、机器人竞赛校本课程综述、机器人竞赛研究现状及问题、机器人竞赛的策略等几个方面，如表2所示。

表2　文献示例

题目	作者	时间	数据库
以机器人竞赛为平台的项目导向教学研究	郭家星	2009	期刊
“微课导学”教学模式构建与实践——以中小学机器人教学为例	王同聚	2015	期刊
中小学机器人竞赛现状以及存在的问题	张清能	2015	期刊
机器人竞赛对学生创新实践能力培养研究	张晓丽；王鹏；崔东辉；范春凤；李彩玲	2015	期刊
浅谈中学机器人竞赛活动培训的有效策略	李玲	2013	期刊

机器人教育并非一蹴而就，有专家介绍[4]，机器人教育要经过三个阶段，建构、编程和控制。一般而言，在幼儿园阶段，主要让幼儿用螺丝将成品零件拼装起来，培养幼儿的动手能力；小学阶段，除了拼装组装，可以进行简单的编程；到了中学阶段，学生们可以更规范的使用机器人，可以自己设计，用高级语言进行简单编程，对机器人进行组装。侯娟[5]分析了教育机器人的特征，以及机器人教育的价值及优势，同时对机器人教育现状进行了总结。

何智在文献[6]中通过阐述中小学生机器人竞赛作为教育研究项目所遵循的理论基础，

doi:10.3772/j.issn.2095-915x.2017.01.009

分析了为什么要开展机器人教育，机器人竞赛教育不光在祖国科技创新上奠定了基础，对于教育方式的变革强有力的推进、学生的综合能力带来了很大提升；机器人作为一种人工智能的高级玩具给青少年带来了乐趣，通过机器人竞赛让青少年将学习乐趣与学习动机结合起来，通过机器人竞赛让他们产生更浓厚的学习兴趣，也通过机器人竞赛锻炼学生们的团队协作能力与宽容为怀的人文品格。

机器人教学能够培养学生创新实践能力和技术素养。常征等[7]从开展机器人兴趣小组教学的实际出发，阐述了机器人教学在提升学生综合素养方面起到的作用。郭家星在文献[8]中指出机器人竞赛教育不光提高了青少年的综合素质与能力，对于师资团队的壮大也极具意义。机器人竞赛平台吸引了各国科学家、科研人员及业内人士，同时也吸引了很多优秀老师；通过带领年轻教师参与竞赛提高他们的专业素质和拓展知识层面，为今后的科研工作打好夯实的基础，为教育教学提供了宝贵的经验。

3.2 文献主题术语抽取

选取主题为“中小学生机器人竞赛”、“大学生机器人竞赛”、“机器人竞赛教育意义”的文献，通过SCP[9]方法和C-value[10,11]方法对文献进行术语和关键词抽取。此方法是一种基于统计和规则相结合的方法进行主题术语抽取，通过对文档结构和论文语言学特点的分析，对SCP和C-value测度加入词频加权因子，并分别来衡量候选术语的单元性和术语性。首先将文本分割成一系列词语；第二步进行停用词过滤，同时也对“敏感区域”进行标注，以便提高这些词作为术语的可能性；第三步对相关信息进行统计，衡量术语的单元性和术语性，以便生成候选术语；第四步通过SCP和C-value方法衡量候选术语的单元性和术语性，从高到低排序候选术语，最终取得前N个术语作为最终术语表。

本文通过SCP（Symmetrical Conditional Probability，对称条件概率）[12]方法来计算候选术语的单元性，对于复杂术语$w_1\cdots w_n$(n>2)公式1如下。

$$SCP(w_1\ldots w_n)=\frac{p(w_1\ldots w_n)^2}{Avp} \quad （公式1）$$

$$Avp=\frac{1}{n-1}\sum_{i=1}^{n-1}p(w_1\ldots w_i)*p(w_{i+1}\ldots w_n) \quad （公式2）$$

其中 $w_1\cdots w_n$ 是一个候选术语，w_i 是组成该候选术语的词语，。是候选术语 $w_1\cdots w_n$ 在语料中出现的概率，可以用词频来计算。

本文通过 C-value 方法来衡量候选术语的术语性，公式 3 如下 [6]。

$$C-value(a)=\begin{cases}\log_2\log_2(|a|+1)*f(a)\\ \log_2\log_2(|a|+1)*f(a)-\dfrac{\sum_{b\in Ta}f(b)}{p(Ta)}\end{cases} \quad （公式3）$$

其中是组成词串 a 的所有词语的个数，是指a在语料库中的词频，Ta是a的所有父串集合，是 a 的所有父串的个数。

抽取的相关术语和关键词分布和对应的C-value、SCP 值，结果如表 3。

由上表可分析，针对中小学生的机器人教育文献的技术词数量较少，针对大学生的文献中技术词相对较多，原因在于中小学生所涉及的科技方面的专业性知识有限，而大学生学习的相关专业知识及运用多。

表3　主题术语词抽取示例

相关主题	术语	C-value	SCP
大学生机器人竞赛	图像处理	3.0	0.0016
	传感检测技术	4.754887502163469	0.00784313725490196
	机械本体结构设计	3.1699250014423126	0.0023872679045092837
	机器人舞蹈	2.0	0.03418803418803419
	计算机编程技术	2.0	0.00810126582278481
	驱动机器	1.0	0.012551229508196721
	机器人搬运	1.0	0.0038461538461538464
	机器人采矿	1.0	0.004
	机器人灭火	6.8	0.0031397174254317113
	机器人控制	2.0	0.11383399209486166
	传感技术	1.0	0.003968253968253968
	通讯传感	2.0	0.0026490066225165563
	机器人舞蹈	2.0	0.0017660910518053376
中小学生机器人竞赛	机器人奥林匹克竞赛	7.924812503605781	1.0
	程序设计	1.0	0.01
	建构机器	1.0	0.007692307692307693
	机器模型	1.0	0.007797270955165692
	硬件设计	2.0	0.0017582417582417582
	应用知识	1.0	0.010245901639344262
	依附性内驱力	1.0	0.014285714285714285
机器人竞赛对教育的意义	工程实践	9.5	0.015503875968992248
	机器人足球赛	7.0	0.009259259259259259
	电子工程	6.0	1.0
	检测技术	5.0	0.09523809523809523
	机械电子	5.0	0.0035929864903707963
	传感检测技术	4.754887502163469	0.0028368794326241137
	图像处理	3.0	0.0011661807580174927
	嵌入式系统	2.0	0.06859410430839002
	自动化类专业	2.0	0.003401360544217687
	仿真软件	1.0	0.012658227848101266
	软件算法	1.0	0.004877626414223043
	模块设计	1.0	0.0033333333333333335
	最优算法	1.0	0.0028735632183908046
	数控维修	1.0	0.0030959752321981426
	软硬件应用		0.0025614754098360654

4 结论

机器人竞赛推动并加速了我国机器人的知识技术的普及，促进我国机器人研究事业的发展和领域人才的培养。在目前条件下，应当尽快建立完善的竞赛体系，降低机器人教育和竞赛的门槛，加强教材建设。术语抽取的结果表明了当前机器人竞赛的内容（比如“机器人灭火”、“机器人采矿”）很实际，贴近人们的日常所需；而机器人技术词抽取的结果表明当前一些“传感技术”、“机械电子”等技术都比较容易掌握，易于让机器人技术融入中小学生的学习课程，提早接触这些技术也为青少年以后的高等教育打下基础。

参考文献

[1] 张国民，张剑平．我国基础教育中机器人教育的现状与对策研究[J]. 现代教育技术, 2008, 18(5): 92-94.

[2] 陶慧萍．高中机器人竞赛校本课程的开发与实践[D]. 南京: 南京师范大学, 2015.

[3] 岳洁．中小学搜救机器人竞赛中的核心问题研究[D]. 南京: 南京师范大学, 2015.

[4] 阳锡叶．竞赛催生的“机器人热”能走多远[J]．辽宁教育, 2015(24): 12-14.

[5] 侯娟. 学校机器人教育现状与思考[J]. 软件导刊: 教育

doi:10.3772/j.issn.2095-915x.2017.01.009

技术, 2016(12).

[6] 何智, 胡又农, 艾伦. 中小学生机器人竞赛的教育价值述评[J]. 中国教育技术装备, 2006(1): 13-15.

[7] 常征, 孙启峰. 机器人教学对学生技术素养提升的探究. 教育现代化, 2016(34).

[8] 郭家星. 以机器人竞赛为平台的项目导向教学研究[J]. 职业教育研究, 2009(12): 133-134.

[9] Silva J F, Lopes G P. A Local Maxima Method and a Fair Dispersion Normalization for Extracting Multiword Units[C]// The 6th Meeting on Mathematics of Language. University of Central Florida, Orlando, FL, USA, July 23–25, 1999.

[10] Maynard D, Ananiadou S. Identifying Terms by Their Family and Friends[C]// Conference on Computational Linguistics. Association for Computational Linguistics, 2000:530-536.

[11] Maynard D, Ananiadou S. TRUCKS: A Model for Automatic Multi-word Term Recognition[J]. Journal of Natural Language Processing, 2000, 8(1): 101-126

[12] 史东娜 . 基于半监督学习的特定领域术语抽取算法的研究 [D]. 北京 : 北京邮电大学 , 2009.