

教育技术与人工智能

贾积有 著

吉林大学出版社

序

教育,是人们传授知识、培养能力和提升素质的一项基本社会活动,旨在不断改善人类认识世界和改造世界的能力,为人类的生存和发展创造越来越好的条件。可以认为,教育,是推动人类社会发展的最为基础性的活动。没有教育的社会,是一种没有开化的社会;一个国家或地区的经济和社会发展水平如何,在很大程度上取决于它的教育发展程度。

作为人类的基础性活动,人是教育成败的首要因素。但是为了提高教育活动的效果和效率,除了应当具备必要的人力资源和教育基础设施之外,还需要借助恰当的教育工具。没有工具的教育活动,是原始的教育活动,很难达到高效率和高水平的程度。一切可以用来帮助改进教育质量、提高教育水平和改善教育效率的工具以及制造这些工具的工艺,都可以称之为教育技术。

教育技术越是先进,教育的效果和效率就可能越好。随着教育事业的发展,随着科学技术的进步,历代先人创造了各种各样的教育工具:从比较简单到比较复杂,从比较低级到比较高级。而在迄今所创造的各种教育技术之中,人工智能教育技术可能是最有前景的一类教育技术。这是因为,教育活动是教育者与被教育者之间复杂的人文交互过程,而且必须实现互相理解,于是要求介入教育活动过程的教育工具也必须具有一定程度的人性化的魅力,具有一定程度的理解能力;而只有人工智能技术才能赋予教育工具以一定的情感能力和理解能力。

当然,人工智能教育技术也不是万能的。特别是在人工智能技术本身还处在发展初级阶段的现今,情形更是如此。然而,与历来的各种教育技术相比,只有人工智能教育技术才有可能实现一定的情感能力和理解能力,因而才有可能比较有效地达到提高教

育活动的效果和效率的目的。

正因为这个缘故,当我看到北京大学教育学院教育技术系贾积有副教授撰写的《教育技术与人工智能》书稿时,就有眼前一亮的感觉,认为这是一部值得推荐的好书。因为,一方面,它准确地抓住了当代教育技术发展的最有希望的方向—人工智能技术;另一方面,它也抓住了人工智能技术发展和应用的最有意义的领域—教育领域。我相信,把人工智能技术与教育的需求结合起来,一定可以创造出最有希望的教育技术。

本书论述了教育、教育技术、人工智能的基本概念及其相互关系,介绍了与教育应用相关的人工智能基本技术,特别值得称道的是,书中介绍了作者自主研究开发的基于网络的智能语言教学系统 **CSIEC**,阐述了它的理论基础、系统架构、相关的人工智能技术和系统应用的情况。我相信,本书的内容,特别是 **CSIEC** 的研制、应用以及改进的设想,将对智能教育技术的发展起到启迪、借鉴和推动的作用。我也相信,只要教育技术和人工智能领域的同仁携起手来,互相学习互相合作,必将使教育技术更好地走向智能化,为教育事业和人工智能事业的发展做出意义重大的贡献。

钟义信

2009 年 10 月于北京

前 言

关于教育、教育技术与人工智能之间的关系,国内外目前尚无同类专门著作进行系统阐述。一些专著和教材章节中对三者之间的关系作了探讨,如美国加涅等编著的教育技术学经典教材《教育技术学基础》等。一些期刊和重要会议论文集中也有一些这方面的专题文章,如 BJET (British Journal of Educational Technology), ReCALL, ET & S (Educational Technology and Science), ITS (Intelligent Tutoring System), AIED (Artificial Intelligence in Education) 等,但都比较分散。所以这本著作的出版,不仅对教育技术学等相关专业的科研工作,而且对教育技术学的专业教学,都将具有填补空白的理论和实践意义。

从理论意义上讲,本著作系统地论述了教育、教育技术与人工智能之间的关系,对人工智能和教育关系的理论探索具有一定的参考作用。从实践意义上讲,本著作融入了科研工作的最新进展和成果,对教育技术和其它相关领域的研究人员和学生应用人工智能技术进行教育领域的系统设计和开发,具有一定的借鉴作用。

本书共九章。第一章为绪论,从历史发展和科学哲学的角度论述教育、教育技术与人工智能的基本概念与相互关系。第二章从系统论的角度分析一般教学系统的构成元素及其关系、动态性、开放性,教育者的功能以及教学法理论基础。第三章介绍计算机辅助教学系统的构成、形式和其中计算机的功能。第四章到第八章分别讲解人工智能的主要领域及其教育应用,分别为知识表达技术(知识工程、谓词表示法、产生式表示法、语义网络和本体论、Prolog、框架表示法、面向对象表示法、脚本表示法等),人工神经网络,机器学习,自然语言处理(词法、语法和语义分析,概率统计方

法,语音识别和合成等),专家系统和 Agent 技术等。第九章详细介绍希赛可(CSIEC)系统,一个作者自主开发研制的具有一定影响的智能网络语言教学系统,包括其理论基础、系统架构、人工智能技术(自然语言处理、Agent、专家系统等)、自由应用和教学整合成果等。

本书是作者依据自己多年的教学实践和科研体会,吸收前辈经验,归纳、提炼、创新而形成的具有一定特色的专著。这本专著跨越教学论、人工智能和教育技术学三个领域,从教育的角度来分析和理解人工智能理论,从教学应用的角度来讲解人工智能技术,而不是单纯介绍人工智能技术细节。参考文献大多为来自国际和国内相关领域著名期刊和会议论文集的最新论文,可以让读者接触到国际前沿最新研究成果。

作为一本新兴、边缘、交叉学科的专著,本书可作为教育类、教育技术类、计算机类、人工智能类、科学哲学类等相关学科的本科生和研究生教学使用,也可作为相关领域的教学者、研究者和管理者等各类人员参考之用。

本书借鉴和参考了国内外很多优秀的论文、专著和教材,由于编排体制限制未能在文中一一注明,在此向各位学者一并表示诚挚的谢意!本书的教学讲义尽管使用过五年多,我也全力以赴地投入了本书的编写,但是错误与不足之处仍然难免,这也正是我今后必须不断努力改进的地方。

本书是在北京大学教育学院教育技术系历经四年的研究生教学讲义的基础提炼而成,其间参与课程的学生提出了不少有益的意见和建议。特此表示衷心感谢!

教育技术系的老师们,都对本书的写作提出了宝贵意见,在此一并感谢!

本书中重点介绍的希赛可智能英语教学系统,先后得到了德国阿登纳基金会、北京大学、教育部、北京市教委、韩国高等教育财团、日本桐山基金会等组织机构的项目资助,特此表示诚挚的谢意!该系统的开发、应用和评估过程,先后有十多位研究生和本科生同学参与,如陈维超、李福攀、黄渊、焦越、徐照银、陈思彪、郝正

元等等,特此表示由衷的谢意!对于参与英语课程整合的英语老师们,如程英、向东方、阮美贤等,以及可爱的同学们,作者更是感激不尽!

作者

2009年8月于北京大学

目 录

第 1 章 绪论	1
1.1 什么是教育?	1
1.2 什么是人类的自然智能?	2
1.3 什么是人工智能?	4
1.4 什么是教育技术?	5
1.5 教育,自然智能,教育技术与人工智能之间的关系 ...	7
第 2 章 一般教学系统	9
2.1 定义	9
2.2 一般教学系统中各个要素之间的关系	10
2.3 一般教学系统的动态性	10
2.4 一般教学系统的开放性	12
2.5 一般教学系统中教育者的功能	12
2.6 一般教学系统中的教学方法和理论基础	14
2.6.1 行为主义:学习依靠强化训练	14
2.6.2 认知主义:学习依靠理解	16
2.6.3 建构主义:学习依靠经历和解释	17
第 3 章 计算机辅助教学系统.....	19
3.1 定义	19
3.2 计算机在一般教学系统中的功能	19
第 4 章 知识表达技术	23
4.1 概述	23

4.1.1	知识的定义	23
4.1.2	知识的分类	23
4.1.3	知识的自然表达方法——自然语言	26
4.1.4	计算机和自然语言	27
4.2	谓词表示法	28
4.2.1	命题逻辑	29
4.2.2	谓词逻辑	30
4.2.3	谓词公式的等价性与永真蕴涵	32
4.2.4	自然演绎推理方法	34
4.2.5	逻辑编程语言 Prolog 简介	35
4.3	产生式表示法	37
4.3.1	定义	37
4.3.2	产生式系统的组成	39
4.3.3	产生式表示法的特点	41
4.4	语义网络表示法	41
4.4.1	定义	41
4.4.2	语义网络表示法的特点	42
4.4.3	语义网络系统: WordNet、HowNet 和 CYC	43
4.5	框架表示法	46
4.5.1	定义	46
4.5.2	框架表示法的特点	47
4.6	面向对象表示法	48
4.7	脚本表示法	49
4.7.1	定义	49
4.7.2	脚本的推理	
4.8	小结	51
4.9	在教育技术上的应用	52
第 5 章	人工神经网络	53
5.1	人工神经网络简介	59
5.2	神经网络的特征	59
5.3	人工神经网络在个性化教学上的应用案例	65

5.3.1	人类性格——大五理论	66
5.3.2	人类智能——多元智力理论	66
5.3.3	通过神经网络识别个人特征	67
5.3.4	学习行为评价	69
5.4	人工神经网络方法在高校教学评估中的应用	70
5.5	人工神经网络方法在高校发展预测中的应用	71
第6章	机器学习	72
6.1	概述	74
6.2	基本模型	74
6.3	机械学习	74
6.4	传授式学习	76
6.5	类比学习	76
6.6	归纳学习	76
6.7	基于解释的学习	77
6.8	在教育技术上的应用	77
第7章	自然语言处理	77
7.1	自然语言理解和产生	80
7.1.1	概述	80
7.1.2	词法分析	80
7.1.3	句法分析	83
7.1.4	语义分析	89
7.2	智能答疑系统	95
7.2.1	聊天机器人:xiaoi 和 CSIEC	98
7.2.2	ELIZA: 关键词匹配技术	98
7.2.3	ALICEBOT 与 AIML	100
7.2.4	模式匹配原理的优点和缺点	107
7.2.5	关键词匹配技术应用:网络搜索	110
7.2.6	汉字智能答疑系统	111
7.3	语音识别和合成技术	114
7.3.1	语音识别技术及其教学应用	115
7.3.2	语音合成技术	116

7.3.3 语音识别和合成软件介绍	124
第8章 其它技术	127
8.1 专家系统	129
8.1.1 发展简史	129
8.1.2 专家系统的定义	129
8.1.3 分类	129
8.1.4 基本结构	129
8.1.5 设计与建造	130
8.1.6 系统评价内容	132
8.1.7 开发工具	132
8.1.8 专家系统与教育	133
8.2 Agent 技术	133
8.2.1 Agent 历史与概念	134
8.2.2 Agent 的特性	134
8.2.3 多 Agent 系统	135
8.2.4 智能和社会性 Agent 技术在教育上的应用—— 以希赛可系统为例	135
第9章 希赛可系统(CSIEC:Computer Simulation in Educational Communication)	136
9.1 动因	137
9.2 指导理论	137
9.3 相关工作	139
9.4 CSIEC 中的英语语法体系	140
9.4.1 句子	143
9.4.2 从句	155
9.4.3 短语	155
9.4.4 单词	159
9.5 希赛可系统结构和技术机理	161
9.5.1 浏览器/服务器接口	162
9.5.2 英语解析器	162
9.5.3 NLML 解析器	163

9.5.4 自然语言数据库 NLDB (Natural Language Database)	165
9.5.5 常识性知识	165
9.5.6 文本蕴涵机理 GTE (Generation of Textual Entailment)	166
9.5.7 交流性反应机理 CR (Communicational Response)	185
9.5.8 多通道用户界面和可以选择的聊天模式	193
9.5.9 适应于用户偏好和话题的自由聊天	197
9.5.10 在给定场景中的引导性对话	198
9.5.11 无限定答案填空练习的自动评分	206
9.5.12 两个机器人的对话演示(“二人转”)	207
9.5.13 听力训练	208
9.5.14 积分机制	208
9.6 系统应用和评估	208
9.6.1 自由用户使用情况分析	210
9.6.2 系统在韩国和日本应用情况分析	216
9.6.3 英语课程整合和评估	223
9.6.4 希赛可系统课程整合研究结论、反思与讨论	249
附录1 插图索引	265
附录2 表格索引	267
参考文献	269

第 1 章 绪论

1.1 什么是教育？

教育作为一个名词，是指一个由人和物所组成的动态系统或者指这个动态系统的变化过程。在这个系统中，必须有一个人是受教育者。所谓的受教育者是指他（她）的（自然）智能在这个动态系统中得到提高的人。换句话说，教育是一个受教育者的智能得到提高的动态系统，或者指受教育者在这个动态系统中的智能得到提高的过程。这个动态系统和过程必须包括受教育者，也可以包括教育者。教育者是指帮助受教育者提高智能的人。

教育作为一个动态系统和过程，它的工作效率可以用这样一个公式来表示：

教育效率

$$= \frac{\text{该系统中所有受教育者的智能的提高程度总和}}{(\text{该系统所耗费的教育者的人力资源总和} + \text{该系统所耗费的自然资源总和} + \text{该系统所耗费的时间})}$$

在动态系统的层面上，我们可以说教育系统，教育制度，教育制，教育组织，教育经济，教育历史，教育发展，教育管理等。这样一个大定义涵盖了各种各样的教育形式，如家庭教育，社会教育，学前教育，学校教育（基础教育，初等教育，中等教育，高等教育），成人教育，企业培训，继续教育，终身教育，自学，远程教育（函授教育，网络教育等），计算机辅助教育（电化教育）等。

在过程的层面上，我们可以说教育业绩，教育评估，教育方法（手段），教育理论，教育技术，等等。如果我们强调受教育者，这个过程就是学（习）；如果我们强调教育者，这个过程就是教。

1.2 什么是人类的自然智能？

人类的自然智能就是人类具有的智力和行为能力。别的生物（动物和植物）也有自己的自然智能。人类的自然智能包括感知能力、记忆能力、思维能力、行为能力和语言能力五个方面（张仰森，2004）。其中的记忆和思维能力是人在大脑中实现的，是内省的。感知、行为和语言能力是人在和周围环境交往中产生的，是外露的。

感知能力指人通过视觉、听觉、触觉、味觉、嗅觉等感觉器官感知外部世界的能力，是人类获取外部信息的基本途径。动物都在不同的方面有不同程度的感知能力，在某些方面动物的感知能力甚至比人还强，比如狗的嗅觉。研究人的感知能力的科学领域有（人体）生理学、生物学、心理学等，研究生物的感知能力的学科有生物学、动物学、植物学等。

行为能力是人们对感知到的外界环境信息的一种反应能力。外界环境可以是异质和同质环境。异质环境就是自然环境，同质环境就是人类社会。对自然环境的行为能力就是适应自然和改造自然的能力。人类对社会环境的行为能力就是适应和改造社会的能力，包括人际交往能力。形形色色的生物对自然环境和种内个体也有不同程度的适应和改造能力。研究人和自然环境关系的学科有环境学、生态学、人类学等，研究人和人的关系的学科有政治学、社会学、经济学、伦理学、心理学等。研究其它生物和自然环境的关系的学科有生物学、动物社会学等。（克劳斯·迈因策尔著 & 曾国屏译，1999）

语言能力其实是一种特殊的行为能力，即人类社会环境中的一种行为能力。人类的语言能力包括语音方式和文字方式。某些动物也具有一定程度的语音方式下的语言交际能力。文字方式的语言能力是人类独有的一种行为能力。研究人类语言能力的学科包括语言学、物理学（声学）、文字学等。

记忆与思维能力是人脑不可分割的最重要的两个功能，也是人类具有智能的根本原因所在，也是至今为止研究的最不清楚的

功能，也可能是永远都搞不清楚的功能。为什么这样说呢？因为我们是在用人脑研究人脑，而且我们不可能打开一个活动的人脑去研究活动的人脑！某些动物也具有记忆能力，但是它们有思维能力吗？他们可以在实验中帮助我们认识人脑，但是终究不能取代人脑！行为能力和语言能力都处在思维能力的控制之下。有意识的感知能力也处于思维能力的控制之下。研究记忆和思维能力的学科有生理学、生物学（神经学）、心理学、脑科学等。最近二十多年来出现的一个交叉科学——认知科学，就是以人类的记忆和思维能力为主要研究对象的一门学科。

记忆指由感觉器官感知到的外部信息和思维产生的结果在大脑中的存储和以后的调用。这些信息和思维结果我们称之为知识。知识有先验的和后验的。先验性的知识是不依赖于人的后天经验而存在的；后验性的知识是外部世界在人脑中的映像。

思维指大脑对记忆的信息进行处理，即利用已有的知识对信息进行分析、计算、比较、判断、推理、联想、决策等，以获取新知识以及运用知识求解问题，并指挥人体进行感知、行为和语言等活动。思维可分为逻辑思维、形象思维以及顿悟思维等。逻辑思维和形象思维是两种基本的思维方式。逻辑思维又称为抽象思维，它是一种根据逻辑规则对信息进行处理的对性思维方式，反映了人们以抽象的、间接的、概括的方式认识世界的过程。归纳与演绎能力是人类进行问题求解的两种主要的推理方式。归纳能力是人们可以通过大量实例，总结出具有一般性规律的知识的能力；而演绎能力则是人类根据已有知识和所感知的事实，推理求解问题的能力。形象思维又称为直感思维，是一种以客观现象为思维对象、以指导创造物化形象的实践为主要目的、以意象为主要思维工具、以感性形象认识为思维材料的思维活动。顿悟思维是在潜意识激发下获得灵感而忽然开窍的思维。（张仰森，2004，p. 5）

学习可以说是一种特殊的思维能力。按照著名人工智能专家、心理学家和经济学家西蒙（Herbert Alexander Simon）的定义，学习是一个系统的有特定目的的知识获取过程，其内在行为

是获取知识、积累经验、发现规律，其外在表现是使系统性能改进、系统实现自我完善、自适应于环境。学习是人类的本能，每个人都在随时随地地学习，既可能是自觉的、有意识的，也可能是不自觉的、无意识的；既可以是由教师指导的，也可以是自己实践的。人们的学习是通过与环境的相互作用而实现的，通过学习可以积累知识，增长才干，充实完善自己，适应环境变化。

1.3 什么是人工智能？

人工智能就是用人工的方法在机器（包括计算机）上实现的智能；或者说就是人们使用机器模拟人类和其它生物的智能，包括感知能力、记忆和思维能力、行为能力、语言能力。由于人工智能是在机器上实现的，因此又称为机器智能。所谓智能机器就是能够在各种环境中自主或交互地执行各种拟人活动的机器，这些活动包括人脑所从事的推理、证明、识别、理解、设计、学习、思考、规划以及问题求解等记忆和思维活动，感知器官从事的感知和识别活动，自然语言理解和产生能力，四肢和身体的活动能力，在社会中的适应能力和人际交往能力，学习能力，等等。比如人工智能的一个研究领域——机器学习，是研究如何使用计算机来模拟人类学习活动的一门学科；更严格地说，就是研究计算机获取新知识和新技能、识别现有知识、不断改善性能、实现自我完善的方法。（蔡自兴，陆汝铃主编，2001）

为什么特别强调计算机作为人工智能的载体？这要从计算机的诞生和历史谈起。计算机（电脑）是为了代替人脑的功能而设计诞生的。这从莱布尼茨（Leibniz）、帕斯卡（Pascal）、Zuse、冯·诺依曼（Von Neumann）、图灵（Turing）等计算机的鼻祖和设计师的文章中就可以看出。计算机科学的发展历史，其实也就是人工智能的发展历史。只不过习惯上，研究者把无法用成熟的数学方法所刻画的人脑功能在计算机上的实现称做人工智能罢了。

从这个定义可以看出，研究人工智能的科学，即人工智能科学是一门跨学科的交叉性的科学。首先，从事该学科的研究人员

要了解包括计算机在内的机器的基本工作原理（硬件或软件），因此人工智能学科一般被划分为计算机科学的一个分支，其次他（她）要从研究自然智能的某个学科中汲取研究成果，并将其在机器上表现出来。人工智能研究的成功之路就依赖于计算机科学与上面提及的学科门类的相互借鉴和融合，包括生物学、动物学、生理学、心理学、人类学、神经学、脑科学、物理学、化学、经济学、社会学、政治学、语言学，和研究所有学科的学科（或者说作为所有学科的基础的学科）——哲学（如其分支认识论、本体论、科学哲学等），以及所有学科的研究工具——数学，等等。

人工智能几千年来的发展和现代人工智能学科五十多年的发展历史已经充分证明了这一点。近现代人工智能学科的奠基人和重要人物（不管他自己承认与否），如 Leibniz, Pascal, Zuse, Hilbert, Goedel, Turing, Von Neumann, Simon, Newell, Minsky, McCarthy, Chomsky, Searle, Schank, Winograd, Rumelhart, Hebb, Hopfield, Habermas, 等等，无不涉足两个或两个以上的学科并颇有建树。

从另外一个角度来讲，如前所述，因为很多有关人类记忆和思维的科学研究（如认知科学）无法获得直接的证据（用大脑来研究大脑），所以很多理论还处于假说阶段。如果在人工智能研究中利用这种假说取得了较好的实验效果，那么这种人工智能的研究成果就可以说为这种假说提供了一种佐证。所以说人工智能研究也是其他相关学科的一个实验场所，比如著名的脑科学中的联结主义便得以在串行计算机上模拟，并取得了令人瞩目的实验结果。

1.4 什么是教育技术？

关于教育技术和教育技术学的定义已经很多了，其中比较著名的是美国 AECT 1994 的定义，即“教学技术是为了促进学习，对有关过程和资源进行设计、开发、利用、管理和评价的理论与实践。”作者的理解是，教育技术是在教育这个动态系统中教育

者使用的各种技术手段，目的是使受教育者的自然智能得到最大限度的提高和教育系统得到最大的工作效率。也就是说教育技术是教育系统或过程的一个组成部分。在这个意义上，教育技术是强调技术手段的教学方法。

为了最大程度地提高教育系统的工作效率，教育者可以使用各种各样的工具，如黑板、粉笔、教鞭、图片、幻灯、投影仪、电影、录像、试验仪器和计算机（硬件和软件）等；并减少教师数量、教学经费和教学时间。究竟如何和何时使用这些技术和手段才能取得最大的工作效率，是教学理论的一个重要研究内容。而教学理论是伴随着教育系统的出现而出现的，深受心理学、哲学（如认识论）等思想的影响。所以说，认知主义、行为主义、建构主义、联结主义等教育技术理论并非是最近几十年来才出现的新理论，而是几百年来心理学、哲学等传统学科发展积累形成的成果。

计算机作为一种特殊的工具被应用在教育中，是 20 世纪 50 年代以来的事情。相比于其他传统教学工具，计算机具有能够和人交互的优点。在其他教学工具面前，受教育者只能是被动的接受者。而在计算机面前，受教育者却可以成为主动的探索者。这种交互性不仅是计算机作为人工智能载体的一个特征，也是它作为人脑的替代工具的功能体现。换句话说，计算机是教育系统中一种具有一定智能的特殊的教学工具。这种特殊性吸引越来越多的各个相关学科的专家学者来致力于它的研究。计算机科学家研究如何将最新的计算机技术应用到教学中；心理学家和教育学家研究计算机在教育上的应用能否提高教育系统的工作效率。不管人们怎么称呼它——CAI（Computer Assited Instruction）、CAL（Computer Assited Learning）、CAT（Computer Assited Teaching）、CBE（Computer Based Education）、CBT（Computer Based Teaching）、WBT（Web Based Teaching）、E-Learning，等等，计算机是一种应用于教育系统的工具和手段的本质并没有被改变。

随着计算机的功能越来越强大，它在教育系统中应用的范围越来越广。那么它究竟有没有可能替代教育系统中的教育者，并