

中国科协新一代信息技术系列丛书

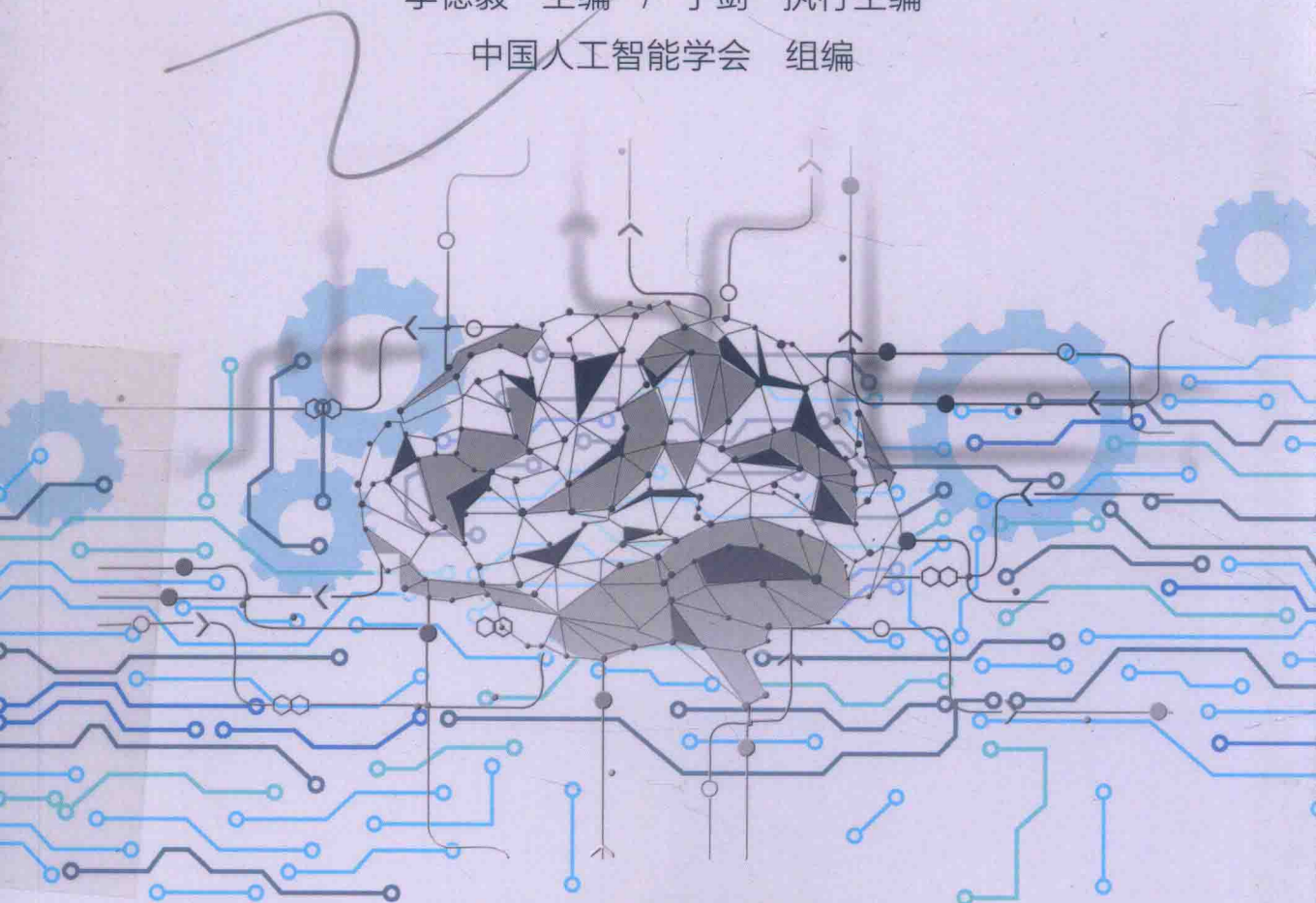
面向非计算机专业

# 人工智能导论

Introduction to  
Artificial Intelligence

李德毅 主编 / 于剑 执行主编

中国人工智能学会 组编



 中国科学技术出版社  
CHINA SCIENCE AND TECHNOLOGY PRESS

中国科协新一代信息技术系列丛书

面向非计算机专业

# 人工智能导论

## Introduction to Artificial Intelligence

主 编 李德毅

执行主编 于 剑

副 主 编 马少平 王万良

参 编 于 剑 山世光 马少平

朱 军 孙富春 李涓

周 明 高 阳 陶建华

中国人工智能学会 组编



中国科学技术出版社

· 北 京 ·

本书是中国科协新一代信息技术系列丛书之一。

本书内容包括知识表示、知识获取、知识应用三部分。其中，知识表示主要介绍概念表示、知识表示、知识图谱；知识获取主要介绍搜索技术、群智能算法、机器学习、人工神经网络与深度学习；知识应用涉及专家系统、计算机视觉、自然语言处理、语音处理、规划、多智能体系统与智能机器人六部分。力求将人工智能的发展脉络、技术理论、产业成果以翔实的形态展现于人前。除了必要的知识点与宽泛的知识图谱，本书还深入浅出地介绍了有关智能搜索技术、机器学习、神经网络、计算机视觉、语言智能、机器人等在内的不同领域的应用实践成果。

本书主要面向大学非计算机类的工科专业的高年级学生与研究生，帮助学生了解人工智能的发展过程与基本知识，熟悉人工智能产业的发展现状与市场需求，培养人工智能应用能力。同时，对于计算机相关专业的学生，本书也可作为人工智能专业课程的先导学习材料。

#### 图书在版编目 ( CIP ) 数据

人工智能导论 / 李德毅主编；于剑执行主编；中国人工智能学会组编. —北京：中国科学技术出版社，2018.8（2018.9重印）

ISBN 978-7-5046-8119-5

I. ①人… II. ①李… ②于… ③中… III. ①人工智能 IV. ①TP18

中国版本图书馆 CIP 数据核字 (2018) 第 182231 号

---

责任编辑	李双北 韩颖
装帧设计	中文天地
责任校对	焦宁
责任印制	李晓霖

---

出版发行	中国科学技术出版社
地 址	北京市海淀区中关村南大街16号
邮 编	100081
发行电话	010-62173865
传 真	010-62179148
网 址	<a href="http://www.cspbooks.com.cn">http://www.cspbooks.com.cn</a>

---

开 本	787mm × 1092mm 1/16
字 数	328千字
印 张	17.25
版 次	2018年8月第1版
印 次	2018年9月第2次印刷
印 刷	北京盛通印刷股份有限公司
书 号	ISBN 978-7-5046-8119-5 / TP·408
定 价	49.00元

---

（凡购买本社图书，如有缺页、倒页、脱页者，本社发行部负责调换）

# 《人工智能导论》编写组

## 顾问

- 陆汝钤 中国科学院院士  
桂卫华 中国工程院院士  
谭铁牛 中国科学院院士  
吴朝晖 中国科学院院士  
戴琼海 中国工程院院士

## 主编

- 李德毅 中国工程院院士

## 执行主编

- 于 剑 北京交通大学教授

## 副主编

- 马少平 清华大学教授  
王万良 浙江工业大学教授

## 参编（按姓氏笔画排序）

- 于 剑 北京交通大学教授  
山世光 中国科学院计算技术研究所研究员  
马少平 清华大学教授  
王万良 浙江工业大学教授  
朱 军 清华大学教授  
孙富春 清华大学教授  
李涓子 清华大学教授  
陈小平 中国科学技术大学教授  
周 明 微软亚洲研究院常务副院长  
高 阳 南京大学教授  
陶建华 中国科学院自动化研究所研究员

## 前言

人工智能再一次成为社会各界关注的焦点，甚至要把我们所处的时代用“智能”来命名，这距人工智能这一概念首次提出已经过去了六十多年。在这期间，我们经历了人工智能发展的两落三起，我国智能科学与技术专业作为154个本科特设专业之一，也已经有了15年的积累。2018年，针对研究生教育的人工智能研究院、针对本科生教育的人工智能学院，在全国各地如雨后春笋般地成立，都在倒逼面向非计算机类的大学本科教育。“本科不牢，地动山摇”，要培养所有在校大学生们的认知力和创造力，为大学生普及人工智能知识，如此紧迫地需要一本《人工智能导论》作为教材，这是始料不及的。

人类走过了农耕社会、工业社会、信息社会，已经进入智能社会，进入在动力工具基础上发展智力工具的新阶段。在农耕社会和工业社会，人类的生产主要基于物质和能量的动力工具，并得到了极大的发展；今天，劳动工具转向了基于数据、信息、知识、价值和智能的智力工具，人口红利、劳动力红利不那么灵了，智能的红利来了。于是，教育也就从传授知识、发明工具、认识和改造客观世界，拓展到人脑自身如何认知、如何再塑造的新阶段。这样一来，人工智能对教育的挑战就不单是一个学科、一个专业的问题，而是培养人们终身学习能力的挑战了。

创新驱动，智能担当，教育先行。我国《新一代人工智能发展规划》明确了我国人工智能发展三步走的战略发展目标。当形形色色的各种机器人成为人类认知自然与社会、扩展智力、走向智慧生活的重要伴侣的时候，当机器人和智能系统无处不在地改变着人类的生产活动、经济活动和社会生活的时候，中国要在2030年成为世界主要人工智能创新中心。为此，中国科协策划并组织编制以人工智能、云计算、大数据等为代表的新一代信息技术系列丛书，成立了新一代信息技术系列丛书编制委员会，聘请梅宏院士为编委会主任，李培根院士、李德毅院士、李伯虎院士、张尧学院士、李骏院士、谭铁牛院士、赵春江院士为编委会委员，统筹丛书编

制工作。《人工智能导论》是该系列丛书之一。

在移动互联网、大数据、超级计算、传感器、脑认知机理等新理论新技术以及经济社会发展强烈需求的共同驱动下，人工智能自身的发展，无论是深度学习、机器感知和模式识别、自然语言处理和理解、知识工程、机器人和智能系统等诸多方面，都把直接面对现实问题作为人工智能的切入点和落脚点，并正在引发链式反应，全方位推动经济社会各领域从数字化、网络化向智能化发展，人工智能正以润物无声的柔软改变着整个世界。从参与编撰《人工智能导论》一书的那一刻，在中国科协的领导下，众多中国科学院院士、中国工程院院士、全国顶尖高校的院长和知名教授、人工智能企业等一线的知识工程师们一起，以创新的方式和务实的精神，力求将人工智能的发展脉络、技术理论、产业成果以翔实的形态展现于人前。除了必要的知识点与宽泛的知识图谱，本书还深入浅出地介绍了有关智能搜索技术、机器学习、神经网络、计算机视觉、语言智能、机器人等在内的不同领域的应用实践成果。

本书邀请陆汝钤院士、桂卫华院士、谭铁牛院士、吴朝晖院士和戴琼海院士担任顾问专家，他们对本书的学术观点、技术方向以及内容组织都提供了极具价值的意见和建议。在此，对各位专家表示深深的敬意和感谢。

本书的编写汇集了多位专家学者的智慧。本书第1章和第2章由于剑编写，第3章、第6章和第8章由王万良编写，第4章由李涓子编写，第5章和第9章由马少平编写，第7章由朱军编写，第10章由山世光编写，第11章由周明编写，第12章由陶建华编写，第13章由陈小平编写，第14章由高阳编写，第15章由孙富春编写。全书由于剑统稿。在编写过程中，整个写作团队克服困难、团结协作、砥砺前行，体现了良好的奉献精神、协作精神和服务精神。本书编写过程中，侯磊参与了第4章编写，苏航参与了第7章编写，刘树杰、段楠和吴侯参与了第11章编写，易江燕、刘斌、郑艺斌和黄健参与了第12章编写，吉建民和吴锋参与了第13章编写，刘华平、刘春芳和方斌参与了第15章编写，在此表示感谢。

中国科协领导多次协调，确保丛书编制和推广工作顺利进行。中国科协学会学术部对丛书的撰写、出版、推广全过程提供了大力支持与具体指导。中国科协智能制造学会联合体承担了丛书的前期调研、组织协调和推广宣传工作。中国人工智能学会具体承担了本书编写的组织工作。中国科学技术出版社承担了本书的编辑校对和出版印刷工作。借此机会一并表示感谢。

当前，国内外一批人工智能领域的企业以及创业独角兽正在成长，获得了广泛的关注和认可。植入商业终端的智能语音交互、采用人脸识别等技术的支付应用、外科手术机器人等正更多地惠及千家万户，汇集多种人工智能技术的“新零售”也

逐渐成为我国科技发展的一张张名片。未来，人工智能还将进一步驱动制造业、教育业、医疗以及金融等行业的深刻变革。毫无疑问，人工智能已经成为经济发展的新引擎、社会发展的加速器。

凡是过往，皆为序章。自1956年人工智能诞生以来，大批杰出的科学家做出了卓越贡献，华人和中华文化在其中发挥了举足轻重的作用，成为中国人工智能崛起的重要原因。今天，以《人工智能导论》的出版为契机，希望能启发更多有志于投身人工智能领域的科学家、工程师、投资者、企业家和广大爱好者，和我们一起共同投身于人工智能发展的伟大建设事业中！衷心希望《人工智能导论》的出版，不仅能够为国家抢占全球人工智能制高点的战略需要贡献力量，也为全社会的人才培养、建立终身学习模式、提高人类自身智能素质添砖加瓦。

《人工智能导论》主编李德毅和编写组全体成员

2018年8月

# | 目 录 |

前言	李德毅
<b>第一章 绪论</b>	<b>001</b>
1.1 人工智能的起源和定义	001
1.2 人工智能的流派	003
1.3 人工智能的进展和发展趋势	009
<b>第二章 概念表示</b>	<b>012</b>
2.1 经典概念理论	012
2.2 数理逻辑	013
2.3 集合论	017
2.4 概念的现代表示理论	020
<b>第三章 知识表示</b>	<b>023</b>
3.1 知识与知识表示的概念	023
3.2 产生式表示法	026
3.3 框架表示法	029
3.4 状态空间表示法	032
3.5 本章小结	036
<b>第四章 知识图谱</b>	<b>038</b>
4.1 知识图谱	038
4.2 本体知识表示	042



4.3	万维网知识表示 .....	043
4.4	知识图谱的现状与发展 .....	049
4.5	知识图谱的应用示例 .....	055
4.6	本章小结 .....	056
<b>第五章</b>	<b>搜索技术 .....</b>	<b>057</b>
5.1	图搜索策略 .....	058
5.2	盲目搜索 .....	059
5.3	启发式搜索 .....	063
5.4	博弈搜索 .....	067
5.5	本章小结 .....	070
<b>第六章</b>	<b>群智能算法 .....</b>	<b>072</b>
6.1	群智能算法产生的背景 .....	072
6.2	遗传算法 .....	073
6.3	粒子群优化算法及其应用 .....	083
6.4	蚁群算法 .....	088
6.5	本章小结 .....	092
<b>第七章</b>	<b>机器学习 .....</b>	<b>094</b>
7.1	机器学习的发展 .....	095
7.2	监督学习 .....	096
7.3	无监督学习 .....	101
7.4	弱监督学习 .....	104
7.5	讨论 .....	108
7.6	本章小结 .....	109
<b>第八章</b>	<b>人工神经网络与深度学习 .....</b>	<b>111</b>
8.1	神经网络的发展历史 .....	111
8.2	神经元与神经网络 .....	112
8.3	BP 神经网络及其学习算法 .....	115
8.4	卷积神经网络 .....	119

8.5	生成对抗网络	126
8.6	深度学习的应用	128
8.7	本章小结	131
<b>第九章</b>	<b>专家系统</b>	<b>133</b>
9.1	专家系统概述	133
9.2	推理方法	135
9.3	一个简单的专家系统	136
9.4	非确定性推理	142
9.5	专家系统工具	146
9.6	专家系统的应用	147
9.7	专家系统的局限性	148
9.8	本章小结	148
<b>第十章</b>	<b>计算机视觉</b>	<b>150</b>
10.1	计算机视觉概述	150
10.2	数字图像的类型及机内表示	152
10.3	常用计算机视觉模型和关键技术	153
10.4	应用实例：人脸识别技术	162
10.5	本章小结	164
<b>第十一章</b>	<b>自然语言处理</b>	<b>166</b>
11.1	自然语言处理概述	166
11.2	机器翻译	172
11.3	自然语言人机交互	177
11.4	智能问答	189
11.5	本章小结	193
<b>第十二章</b>	<b>语音处理</b>	<b>195</b>
12.1	语音的基本概念	195
12.2	语音识别	196
12.3	语音合成	202

12.4	语音增强 .....	206
12.5	语音转换 .....	208
12.6	情感语音 .....	210
12.7	本章小结 .....	214
<b>第十三章</b>	<b>规划 .....</b>	<b>216</b>
13.1	基本概念 .....	216
13.2	经典规划 .....	217
13.3	概率规划 .....	221
13.4	典型应用 .....	227
<b>第十四章</b>	<b>多智能体系统 .....</b>	<b>229</b>
14.1	智能体 .....	229
14.2	智能体的具体结构 .....	234
14.3	多智能体协商 .....	236
14.4	多智能体学习 .....	242
14.5	本章小结 .....	244
<b>第十五章</b>	<b>智能机器人 .....</b>	<b>246</b>
15.1	概述 .....	247
15.2	人工智能技术在机器人中的应用 .....	249
15.3	智能机器人发展展望 .....	259
<b>参考文献</b>	<b>.....</b>	<b>262</b>

# 第一章 绪论

---

自 1946 年第一台计算机诞生，人们一直希望计算机能够具有更加强大的功能。进入 21 世纪，由于计算能力的提高和大数据的积累，人们发现人工智能（Artificial Intelligence, AI）可以使计算机更加智能。这不仅可以创造一些新行业，也可以给传统行业赋能，从而导致了人工智能的新一轮热潮。

在国家政策层面，2017 年 7 月 20 日国务院印发了《新一代人工智能发展规划》，在 2018 年的中国政府工作报告中更明确提出“加强新一代人工智能研发应用”。国外的科技发达国家，如法国、德国、美国、日本等也出台了相关扶持政策。

在产业界，许多信息技术企业都相继涉足人工智能领域。例如，以做硬件著称于世的 IBM 公司已经开始转型做人工智能了，许多的互联网企业如百度、谷歌、微软等更是全面转型人工智能。如今，许多创业公司更是以人工智能为主攻方向。

在实际产品开发方面，人工智能技术也得到了广泛应用，如寒武纪 1H8 等 AI 芯片、百度 Apollo 计划开放自动驾驶平台、手机的指纹识别与人脸识别产品等。

由于人工智能技术和产品日新月异，为了规范人工智能技术的合理利用和研发，人工智能伦理也被提上了议事日程。2017 年 1 月举行的 Beneficial AI 会议提出了 23 条阿西洛马人工智能原则（Asilomar AI Principles），指出人工智能和脑机接口技术必须尊重和保护人的隐私、身份认同、能动性和平等性。

对人工智能有所了解和研究，是新时代对当代大学生提出的新要求。本章将分节论述人工智能的起源和定义、人工智能的流派、人工智能的进展和发展趋势。

## 1.1 人工智能的起源和定义

如果不做很远的追溯，现代人工智能的起源就非常明确。现代人工智能的起

源公认是 1956 年的达特茅斯会议。达特茅斯会议主要参加者有 10 人，分别是麦卡锡、明斯基、香农、罗切斯特、纽厄尔、西蒙、撒缪尔、伯恩斯坦、摩尔、所罗门诺夫，其中前四位是发起人。达特茅斯会议的最主要成就是使人工智能成了一个独立的研究学科。人工智能的英文名称是“Artificial Intelligence”，有文献可考的记录是出自 1956 年的达特茅斯会议。在此之前，即使有相关的名词术语，也不是大家对人工智能学科的命名共识。

人工智能如何定义呢？严格来说，历史上有很多人工智能的定义，这些定义对于人们理解人工智能都起过作用，甚至是很大的作用。比如，达特茅斯会议的发起建议书中对于人工智能的预期目标的设想是“制造一台机器，该机器可以模拟学习或者智能的所有方面，只要这些方面可以精确描述”（Every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it）。该预期目标也曾经被当作人工智能的定义使用，对人工智能的发展起到了举足轻重的作用。

时至今日，还没有一个被大家一致认同的精确的人工智能定义。

但目前最常见的 AI 定义有两个：一个是明斯基提出的，即“人工智能是一门科学，是使机器做那些人需要通过智能来做的事情”；另一个更专业一些的定义是尼尔森给出的，即“人工智能是关于知识的科学”，所谓“知识的科学”就是研究知识的表示、知识的获取和知识的运用。

在这两个定义中，专业人士更偏向于第二个定义。原因很简单，因为第一个定义中涉及两个未明确定义的概念，一个是人，一个是智能。什么是人？什么是智能？到现在依然是很难清楚回答的问题。相比之下，第二个定义只涉及一个未明确定义的概念，就是知识。在人、智能、知识这三个概念当中，知识被研究的应该会比较彻底的。同时，人和智能的定义也与知识紧密相关，而且知识是智能的基础。如果没有任何知识，很难发现什么是智能。

所以一般来说，AI 的研究是以知识的表示、知识的获取和知识的应用为归依。虽然不同的学科致力于发现不同领域的知识，但应承认所有的学科都是以发现知识为目标的。比如，数学研究数学领域的知识、物理研究物理领域的知识，等等。而人工智能希望发现可以不受领域限制、适用于任何领域的知识，包括知识表示、知识获取以及知识应用的一般规律、算法和实现方式等。因此，相对于其他学科，AI 具有普适性、迁移性和渗透性。一般来说，将人工智能的知识应用于某一特定领域，即所谓的“AI+ 某一学科”，就可以形成一个新的学科，如生物信息学、计算历史学、计算广告学、计算社会学等。因此，掌握人工智能知识已经不仅仅是对人工智能研究者的要求，也是时代的要求。

## 1.2 人工智能的流派

根据前面的论述，我们知道要理解人工智能就要研究如何在一般的意义上定义知识。可惜的是，准确定义知识也是一个十分复杂的事情。严格来说，人们最早使用的知识定义是柏拉图在《泰阿泰德篇》中给出的，即“被证实的、真的和被相信的陈述”，简称知识的 JTB 条件。

然而，这个延续了两千多年的定义在 1963 年被哲学家盖梯尔否定了。盖梯尔提出了一个著名的悖论（简称“盖梯尔悖论”），该悖论说明柏拉图给出的知识定义存在严重缺陷。虽然后来人们给出了很多知识的替代定义，但直到现在仍然没有定论。

但关于知识，至少有一点是明确的，那就是知识的基本单位是概念。精通掌握任何一门知识，必须从这门知识的基本概念开始学习。而知识自身也是一个概念。因此，如何定义一个概念，对于人工智能具有非常重要的意义。给出一个定义看似简单，实际上是非常难的，因为经常会涉及自指的性质。一旦涉及自指，就会出现非常多的问题，很多的语义悖论都出于概念自指。关于这方面的深入讨论，有兴趣的同学可以读一读侯世达的《哥德尔、艾舍尔、巴赫：集异璧之大成》，该书对于概念自指有一些非常深入浅出的例子。

知识本身也是一个概念这件事情非同寻常。据此，人工智能的问题就变成了如下三个问题——如何定义（或者表示）一个概念、如何学习一个概念、如何应用一个概念。因此对概念进行深入研究就非常必要了。

那么，如何定义一个概念呢？简单起见，这里先讨论最为简单的经典概念。经典概念的定义由三部分组成：第一部分是概念的符号表示，即概念的名称，说明这个概念叫什么，简称概念名；第二部分是概念的内涵表示，由命题来表示，命题就是能判断真假的陈述句；第三部分是概念的外延表示，由经典集合来表示，用来说明与概念对应的实际对象是哪些。

举一个常见的经典概念的例子——素数。其概念名在汉语中为“素数”，在英语中称为 prime number；其内涵表示是一个命题，即只能够被 1 和自身整除的自然数；其外延表示是一个经典集合，就是  $\{1, 2, 3, 5, 7, 11, 13, 17, \dots\}$ 。

概念有什么作用呢？或者说概念定义的各个组成部分有什么作用呢？很容易发现，经典概念定义的三部分各有其作用，且彼此不能互相代替。具体来说，概念有三个作用或功能，要掌握一个概念，必须清楚其三个功能。

第一个功能是概念的指物功能，即指向客观世界的对象，表示客观世界的对象

的可观测性。对象的可观测性是指对象对于人或者仪器的知觉感知特性，不依赖于人的主观感受。举一个《阿Q正传》里的例子：那赵家的狗，何以看我两眼呢？句子中“赵家的狗”应该是指现实世界当中的一条真正的狗。但概念的指物功能有时不一定能够实现，有些概念其设想存在的对象在现实世界并不存在，例如“鬼”。

第二个功能是指心功能，即指向人心智世界里的对象，代表心智世界里的对象表示。鲁迅有一篇著名的文章《论丧家的资本家的乏走狗》，显然，这个“狗”不是现实世界的狗，只是他心智世界中的狗，即心里的狗（在客观世界，梁实秋先生显然无论如何不是狗）。概念的指心功能一定存在。如果对于某一个人，一个概念的指心功能没有实现，则该词对于该人不可见，简单地说，该人不理解该概念。

最后一个功能是指名功能，即指向认知世界或者符号世界表示对象的符号名称，这些符号名称组成各种语言。最著名的例子是乔姆斯基的“colorless green ideas sleep furiously”，这句话翻译过来是“无色的绿色思想在狂怒地休息”。这句话没有什么意思，但是完全符合语法，纯粹是在语义符号世界里，即仅仅指向符号世界而已。当然也有另外，“鸳鸯两字怎生书”指的就是“鸳鸯”这两个字组成的名字。一般情形下，概念的指名功能依赖于不同的语言系统或者符号系统，由人类所创造，属于认知世界。同一个概念在不同的符号系统里，概念名不一定相同，如汉语称“雨”，英语称“rain”。

根据波普尔的三个世界理论，认知世界、物理世界与心理世界虽然相关，但各不相同。因此，一个概念的三个功能虽然彼此相关，也各不相同。更重要的是，人类文明发展至今，这三个功能不断发展，彼此都越来越复杂，但概念的三个功能并没有改变。

在现实生活中，如果你要了解一个概念，就需要知道这个概念的三个功能：要知道概念的名字，也要知道概念所指的对象（可能是物理世界），更要在自己的心智世界里具有该概念的形象（或者图像）。如果只有一个，那是不行的。

举一个简单的例子。清华大学计算机系的马少平教授曾经讲过一个很有趣的故事：有一天马老师出去开会，自己一个人在一张桌子上吃饭，有人就过来问他是哪个单位的，马老师回答说是清华大学的。那人听了很高兴，就接着问清华大学哪个系的，马老师说是计算机系的。那人说：“我认识清华大学计算机系的一个老师，不知您认不认识？”马老师回应：“我在清华大学计算机系待了30年，你说的老师我应该认识”，那人很骄傲地一抬头说：“我认识马少平。”

这个人认识马少平老师吗？显然，不能说他认识，因为他不知道跟他说话的就是马少平。所以说，掌握一个概念需要三指都对才行。如果只能指名、不能指物，还是不能说理解了相应的概念。

知道了概念的三个功能之后，就可以理解人工智能的三个流派以及各流派之间的关系。人工智能也是一个概念，而要使一个概念成为现实，自然要实现概念的三个功能。人工智能的三个流派关注于如何才能让机器具有人工智能，并根据概念的不同功能给出了不同的研究路线。专注于实现 AI 指名功能的人工智能流派称为符号主义；专注于实现 AI 知心功能的人工智能流派称为连接主义；专注于实现 AI 指物功能的人工智能流派称为行为主义。

### 1.2.1 符号主义

符号主义的代表人物是 Simon 与 Newell，他们提出了物理符号系统假设，即只要在符号计算上实现了相应的功能，那么在现实世界就实现了对应的功能，这是智能的充分必要条件。因此，符号主义认为，只要在机器上是正确的，现实世界就是正确的。说得更通俗一点，指名对了，指物自然正确。

在哲学上，关于物理符号系统假设也有一个著名的思想实验——图灵测试。图灵测试要解决的问题就是如何判断一台机器是否具有智能。

图灵测试的思想实验如下：一个房间里有一台计算机和一个人，计算机和人分别通过各自的打印机与外面联系。外面的人通过打印机向屋里的计算机和人提问，屋里的计算机和人分别作答，计算机尽量模仿人。所有回答都是通过打印机用语言描述出来的。如果屋外的人判断不出哪个回答是人、哪个回答是计算机，就可以判定这台计算机具有智能。图灵测试示意图见图 1.1。

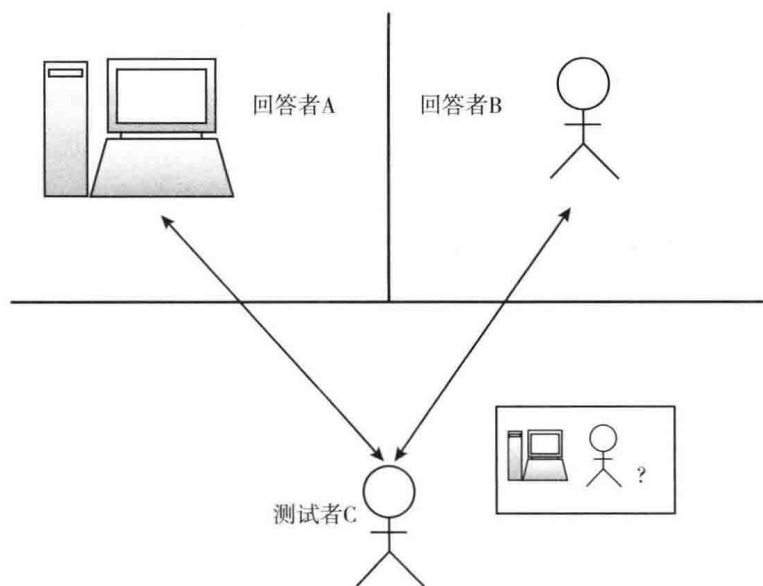


图 1.1 图灵测试



显然，上述测试都是在符号层面进行的，是一个符号测试方式。这个测试方式具有十分重要的意义。因为截至目前，并没有人给出一个为大家所一致认可的智能的内涵定义，而如何判定是否具有智能就面临很大困难。有了图灵测试，我们就可以将研究智能的重点放在智能的外在功能性表现上，使智能在工程上看似乎是可以实现和判断的。

图灵测试将智能的表现完全限定在指名功能里。但马少平教授的故事已经说明，只在指名功能里实现了概念的功能，并不能说明一定实现了概念的指物功能。实际上，根据指名与指物的不同，哲学家 Searle 专门设计了一个思想实验用来批判图灵测试，这就是著名的中文屋实验。

中文屋实验设计如下：一个人住在一个房间里，他懂英文。但是在这个房间里有一个构造好的计算机程序，这个计算机程序可以根据中文输入回答任意中文问题，同时这个房间有一个窗口可以递出和递入纸条。通过这个窗口递入中文问题，屋里的这个人按照这个计算机程序输出相应的中文答案。由于其应对无误，显然屋外的人会认为其精通中文，但实际上屋里的人对中文一无所知。中文屋实验示意图见图 1.2。

中文屋实验明确说明，即使符号主义成功了，这全是符号的计算跟现实世界也不一定搭界，即完全实现指名功能也不见得具有智能。这是哲学上对符号主义的一个正式批评，明确指出了按照符号主义实现的人工智能不等同于人的智能。

虽然如此，符号主义在人工智能研究中依然扮演了重要角色，其早期工作的主要成就体现在机器证明和知识表示上。在机器证明方面，早期 Simon 与 Newell 做出了重要的贡献，王浩、吴文俊等华人也得出了很重要的结果。机器证明以后，符号主义最重要的成就是专家系统和知识工程，最著名的学者就是 Feigenbaum。如果认为沿着这一条路就可以实现全部智能，显然存在问题。日本第五代智能机就是沿着知识工程这条路走的，其后来的失败在现在看来是完全合乎逻辑的。

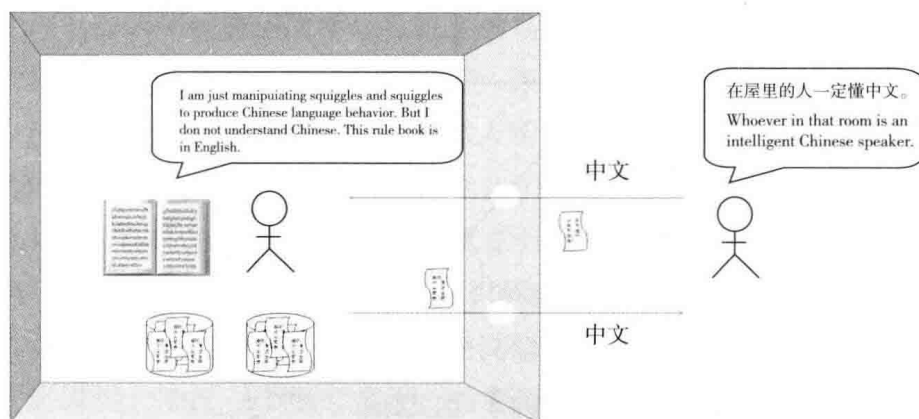


图 1.2 中文屋实验