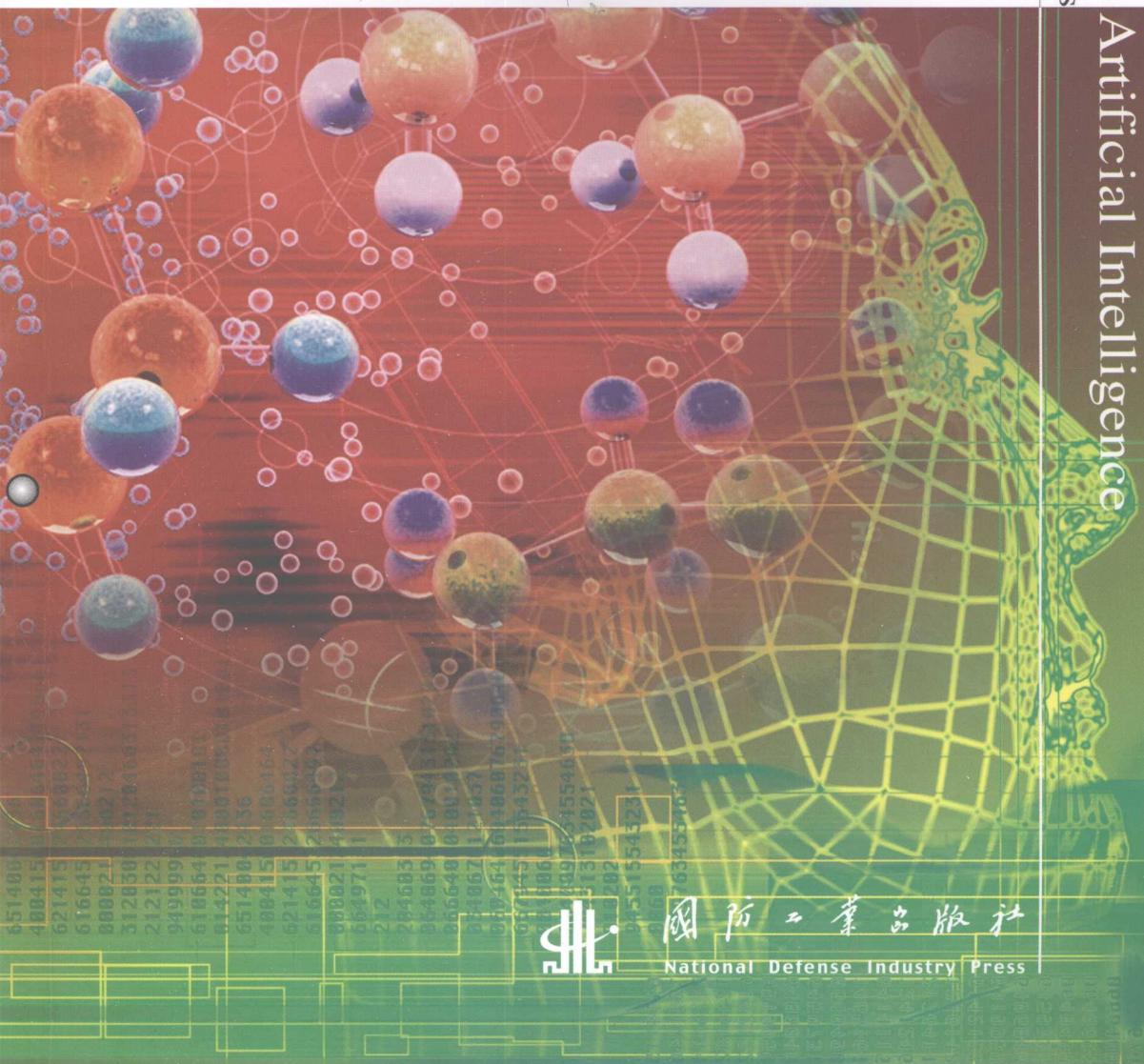


刘峡壁 ◎ 编著

# 人工智能 基础论

## —方法与系统

Introduction to Artificial Intelligence  
Approaches and Systems



国防工业出版社

National Defense Industry Press

# 人工智能导论——方法与系统

## Introduction to Artificial Intelligence

### Approaches and Systems

刘峡壁 编著

国防工业出版社

北京 · 100084

图书在版编目(CIP)数据

人工智能导论：方法与系统/刘峡壁编著. —北京：国防工业出版社，2008.8

ISBN 978 - 7 - 118 - 05810 - 9

I. 人... II. 刘... III. 人工智能 IV. TP18

中国版本图书馆 CIP 数据核字(2008)第 094452 号

\*

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100044)

腾飞印务有限公司印刷

新华书店经售

\*

开本 710×960 1/16 印张 22 1/2 字数 398 千字

2008 年 8 月第 1 版第 1 次印刷 印数 1—4000 册 定价 36.00 元

---

(本书如有印装错误,我社负责调换)

国防书店: (010)68428422

发行邮购: (010)68414474

发行传真: (010)68411535

发行业务: (010)68472764

## 前言

探索生物智能的奥秘并予以非生物性的复制是极富乐趣之事。在这一领域，人类已取得许多令人振奋的成就，其中一些正在影响和改变着人们的生活，而另外一些以及以后还会不断涌现的新思想、新方法和新技术将在更大的范围内和更深的程度上影响和改变我们的生活，甚至影响我们对世界和人生的看法。

作为延伸人类智力而非体力的机器，计算机可谓人工智能的雏形。随后人工智能长期沿着符号主义和连接主义的道路平行发展。但最近二十年来，人工智能的哲学基础似乎正在变得更加丰富。行为智能、进化计算、群智能、量子计算、DNA 计算等激动人心的新进展不仅拓宽了人们实现人工智能的视野，而且充分证实了人工智能是目前人类科学大厦中最为活跃、最富有魅力的组成部分之一。

本书旨在向读者介绍人工智能的核心知识与最新进展，使读者建立起对于人工智能的总体认识，为以后进入人工智能各分支的研究和应用奠定基础。人工智能是一个庞杂的学科体系，从概念上讲，一切为复制生物智能而做出的努力都可纳入其中。如何能够系统、全面、简洁地描述人工智能的全貌而不显得凌乱，绝非易事。编者在为本科生讲授“人工智能基础”课程时，就感到涵盖自己所欲讲授全部内容的教材或书籍尚不多见，需要查阅许多不同资料才能获取相关信息，于是萌生撰写此书的念头，并一路坚持下来，直至此书面世。

为了详尽地阐述人工智能的核心知识，必须有一条主线将这些知识串联起来。本书所确定的主线是从实现人工智能的角度，将有关知识划分为哲学基础和工程实践两大块。哲学基础是实现人工智能的不同哲学思想和在相应思想指导下的具体方法；工程实践则是有关方法在实际问题中的应用和集成，以及方法实现所需要的软硬件条件。

本书第二章至第八章为实现人工智能的哲学基础部分，其中将目前实现人工智能的不同思想归纳为符号主义、连接主义、学习主义、行为主义、进化主义和群体主义六种。

符号主义是以符号处理为核心的人工智能实现途径，关键问题是搜索、知识表示和推理，相应内容被概括为基于搜索的问题求解和基于知识的推理两大主

题,分别在第二章和第三章中讲述。具体内容包括问题表示与求解方法、盲目图搜索技术与启发式图搜索技术、计算机博弈搜索技术、知识表示方法、确定性推理方法和不确定性推理方法。

连接主义是以神经系统结构模拟为基础的人工智能实现途径,试图通过结构的模拟达到功能的模拟。目前,连接主义主要是指人工神经网络,其基本问题在于网络结构和学习算法。不同的网络结构和学习算法对应不同的神经网络模型。本书第四章阐述神经网络基本问题,并分为感知器、霍普菲尔德网络和自组织特征映射网三大类型对感知器与多层感知器、误差反向传播网络(BP网络)、自适应线性神经元、径向基函数网络、霍普菲尔德网络、玻耳兹曼机、自组织特征映射网等经典神经网络模型展开讨论。

学习主义是赋予机器以类似于人的学习能力,通过自我学习提高机器的智能水平。虽然严格地讲,目前还不存在独立的学习主义,但机器学习对于人工智能是重要的,在其他各条实现途径中几乎都是不可缺少的组成部分,没有学习能力的智能系统难以称其为真正的智能系统。本书第五章介绍机械式学习、指导式学习、类比学习、归纳学习和解释学习等学习策略,并重点阐述监督归纳学习和非监督归纳学习的一些主流方法,包括决策树、朴素贝叶斯分类器、贝叶斯信念网、划分聚类方法和层次聚类方法。

行为主义是对生物行为能力的模拟,强调在与环境的交互中表现并逐步提高智能系统的能力,其基本特点是“感知—行为”模式、现场式智能和渐进的智能。这些将在第六章进行讨论。智能体具有现场性、反应性和自主性的行为主义的基本特征。强化学习则是在环境状态与系统行为之间建立最优映射关系的重要学习方法。因此,第六章还将介绍智能体的概念、体系结构以及强化学习方法。

进化主义是对生物进化机制的模拟,试图通过群体的进化来解决人工智能中核心的搜索(优化)问题。目前,进化主义主要指遗传算法、进化策略和进化规划这三种进化计算方法。本书第七章介绍三种进化计算方法的基本原理和特点。

群体主义通过个体之间的协作,解决单个个体难以有效解决的问题或者提高个体解决问题的效率。目前,群体主义主要包括多智能体系统和以蚁群优化算法、粒子群优化算法为代表的群智能优化算法。这些是本书第八章的主要内容。

本书第九章至第十一章为工程实践部分,包括三个方面的内容。  
① 介绍具有代表性和广泛应用性的人工智能系统。这里,仅在第九章阐述专家系统的概念、体系结构、主要问题和开发方法。  
② 在第十章介绍实现人工智能的软件条

件,指两种人工智能程序设计语言:LISP 语言和 PROLOG 语言。③ 在第十一章介绍未来实现人工智能的硬件基础,指三种下一代智能计算机:光计算机、量子计算机和生物计算机。

除了主要内容以外,本书还有两项辅助性内容希望读者能予以关注。第一项是在每章最后提供的深入学习资源,包括书籍、论文和网络资源。作为教材,本书在叙述上力求系统、全面和简洁,因此着力介绍有关概念、方法和系统的核芯内容,不涉及过多技术细节和高级话题。这些书籍、论文和网络资源则为有兴趣或需要深入了解有关知识的读者提供了获取相应信息的地图,希望读者能够借助这些资源在每一章阐述的基础知识之上,更加广泛、深入、透彻地理解和思考有关问题。第二项是在附录 1 中提供的汉英—英汉术语对照与索引。希望读者通过该附录,在阅读完本书后,建立起人工智能的英文概念体系,从而具备查询和阅读人工智能领域英文文献的初步能力。这是读者了解学科前沿和参与国际交流所不可缺少的。

感谢“北京理工大学优秀青年教师资助计划”对出版本书的资助。感谢国防工业出版社在本书编写和出版过程中给与的大力支持和帮助。

感谢我的妻子一直以来在生活上对我无微不至的照顾和关爱,感谢我的儿子常常带给我欢笑,感谢我和我妻子的父母始终在背后默默支持我们。没有这些,很难想象本书能够完成。谨以此书献给他们。

人工智能内容广泛,发展迅猛,加之作者学术素养和水平均有限,因此虽经仔细推敲,但书中疏漏和错误之处可能仍然难以避免。如能收到读者对本书提出的建设性的批评、意见和建议,我将感到十分高兴和感激。接收读者反馈的电子邮件地址是 liuxiabi@ bit. edu. cn。

刘峡壁

2008 年 5 月

于北京理工大学

# 目 录

<b>第一章 绪论</b>	1
1.1 智能	1
1.2 人工智能	3
1.2.1 如何衡量机器是否具有智能	3
1.2.2 人工智能的研究目标	4
1.2.3 人工智能的研究和应用领域	5
1.3 人工智能发展简史	10
1.3.1 孕育期(1956年以前)	10
1.3.2 形成期(1956年—1970年)	12
1.3.3 知识工程时代(1970年至20世纪80年代初)	13
1.3.4 发展期(20世纪80年代初至今)	15
1.4 人工智能的实现途径	16
1.4.1 符号主义	16
1.4.2 连接主义	18
1.4.3 学习主义	18
1.4.4 行为主义	19
1.4.5 进化主义	20
1.4.6 群体主义	21
1.5 本书的内容与组织	21
小结	22
深入学习资源	23
习题	24
<b>第二章 搜索与问题求解</b>	25
2.1 搜索概述	25

2.2	问题求解 .....	27
2.2.1	状态空间 .....	27
2.2.2	与或图 .....	29
2.3	图搜索 .....	33
2.3.1	图搜索算法的一般结构 .....	33
2.3.2	盲目图搜索 .....	34
2.3.3	启发式图搜索 .....	36
2.4	博弈搜索 .....	43
2.4.1	博弈树 .....	44
2.4.2	极大极小搜索 .....	46
2.4.3	$\alpha - \beta$ 剪枝 .....	50
	小结 .....	52
	深入学习资源 .....	54
	习题 .....	54
<b>第三章</b>	<b>知识与推理</b> .....	<b>57</b>
3.1	推理概述 .....	57
3.1.1	推理方式 .....	57
3.1.2	推理控制策略 .....	58
3.2	知识及其表示 .....	60
3.2.1	知识 .....	60
3.2.2	知识的表示 .....	62
3.3	知识表示方法 .....	63
3.3.1	一阶谓词逻辑表示法 .....	63
3.3.2	产生式表示法 .....	70
3.3.3	其他知识表示方法 .....	73
3.4	经典逻辑推理的逻辑基础 .....	81
3.4.1	推理规则 .....	81
3.4.2	范式 .....	83
3.4.3	置换与合一 .....	84
3.4.4	自然演绎推理 .....	85
3.5	归结演绎推理 .....	87
3.5.1	子句集及其转化方法 .....	87
3.5.2	归结原理 .....	88
3.5.3	命题逻辑的归结反演 .....	90
3.5.4	谓词逻辑的归结反演 .....	91

3.5.5 归结策略	93
3.6 不确定性推理基本问题	94
3.7 基于概率论的不确定性推理	96
3.7.1 概率论基础	96
3.7.2 可信度方法	98
3.7.3 主观贝叶斯方法	102
3.8 基于证据理论的不确定性推理	106
3.8.1 证据理论基础	106
3.8.2 推理方法	109
3.9 模糊推理	112
3.9.1 模糊集基础	113
3.9.2 模糊推理模型	116
小结	120
深入学习资源	122
习题	122

## 连接主义

第四章 人工神经网络	127
4.1 人工神经元模型	128
4.1.1 生物神经元	128
4.1.2 人工神经元	130
4.1.3 人工神经元互连模型	133
4.2 人工神经网络的拓扑结构	134
4.3 人工神经网络的学习方式	136
4.3.1 赫伯型学习	136
4.3.2 误差修正型学习	137
4.3.3 竞争型学习	138
4.3.4 随机型学习	139
4.4 多层感知器及其学习算法	140
4.4.1 感知器与多层感知器	140
4.4.2 误差反向传播网络(BP 网络)	143
4.4.3 自适应线性神经元网络	147
4.4.4 径向基函数网络	148

4.5	霍普菲尔德网络与玻耳兹曼机	151
4.5.1	霍普菲尔德网络	152
4.5.2	联想记忆	154
4.5.3	优化计算	155
4.5.4	玻耳兹曼机	158
4.6	自组织特征映射网	159
小结		162
深入学习资源		164
习题		164

## 学 习 主 义

第五章	机器学习	166
5.1	什么是机器学习?	166
5.2	机器学习的主要策略	168
5.2.1	机械式学习	168
5.2.2	指导式学习	169
5.2.3	类比学习	172
5.2.4	解释学习	172
5.2.5	归纳学习	172
5.3	类比学习方法	173
5.3.1	属性类比学习	174
5.3.2	转换类比学习	175
5.4	基于解释的学习	176
5.4.1	学习模型	176
5.4.2	基于解释的学习方法	178
5.5	规则集的归纳学习	181
5.5.1	示例空间	182
5.5.2	归纳过程	182
5.5.3	规则空间	183
5.5.4	验证过程	184
5.6	决策树学习	184
5.6.1	决策树	184
5.6.2	基本决策树生成算法(ID3 算法)	184

5.6.3 决策树剪枝	188
<b>5.7 贝叶斯学习</b>	<b>190</b>
5.7.1 贝叶斯法则	190
5.7.2 朴素贝叶斯分类器	192
5.7.3 贝叶斯信念网	195
<b>5.8 聚类分析</b>	<b>199</b>
5.8.1 数据相似性度量	200
5.8.2 划分聚类方法	202
5.8.3 层次聚类方法	206
小结	207
深入学习资源	209
习题	210

## 行为主义

<b>第六章 行为智能</b>	<b>214</b>
<b>6.1 没有表示和推理的智能</b>	<b>215</b>
6.1.1 感知—行为模式	215
6.1.2 渐进的智能	216
<b>6.2 智能体</b>	<b>217</b>
6.2.1 智能体的特点	218
6.2.2 智能体的体系结构	218
6.2.3 智能体的学习	220
<b>6.3 强化学习</b>	<b>220</b>
6.3.1 强化学习任务与模型	221
6.3.2 Q—学习算法	223
6.3.3 Dyna—Q	225
6.3.4 强化学习的应用	225
小结	226
深入学习资源	227
习题	227

## 进化主义

第七章 进化计算	229
7.1 生物进化对计算的启示	230
7.1.1 生物进化的启示	230
7.1.2 不同的进化计算观点	231
7.2 进化计算的一般框架和共同特点	232
7.2.1 进化计算的一般框架	232
7.2.2 进化计算的共同特点	233
7.3 遗传算法	233
7.3.1 遗传算法的起源与发展	233
7.3.2 遗传算法的基本思想	234
7.3.3 模式理论	239
7.3.4 遗传算法的改进与变形	244
7.4 进化规划	244
7.4.1 进化规划的起源与发展	244
7.4.2 进化规划的主要特点	245
7.4.3 进化规划中的遗传算子	245
7.4.4 进化规划的算法流程	246
7.5 进化策略	247
7.5.1 进化策略的起源与发展	247
7.5.2 进化策略的主要特点	247
7.5.3 进化策略的不同形式	247
7.5.4 进化策略中的遗传操作	248
7.5.5 进化策略的算法流程	250
小结	251
深入学习资源	252
习题	252

## 群体主义

第八章 群智能	253
---------	-----

8.1	多智能体系统 .....	253
8.1.1	智能体之间的通信机制 .....	253
8.1.2	智能体之间的协调策略 .....	255
8.1.3	智能体之间的协作策略 .....	255
8.1.4	多智能体强化学习 .....	256
8.2	蚁群优化算法 .....	257
8.2.1	蚁群觅食行为对计算的启示 .....	257
8.2.2	蚁群优化算法的基本原理 .....	258
8.2.3	蚁群优化算法的改进 .....	261
8.3	粒子群优化算法 .....	263
8.3.1	鸟群飞行方式对计算的启示 .....	263
8.3.2	粒子群优化算法的基本原理 .....	263
8.3.3	粒子群优化算法中的有关参数 .....	264
	小结 .....	266
	深入学习资源 .....	267
	习题 .....	267

## 智能系统

<b>第九章</b>	<b>专家系统 .....</b>	<b>269</b>
9.1	什么是专家系统 .....	269
9.1.1	基本概念 .....	269
9.1.2	专家系统类型 .....	270
9.2	专家系统的体系结构 .....	273
9.3	知识获取 .....	275
9.3.1	知识获取的任务 .....	276
9.3.2	知识获取的方法 .....	276
9.4	专家系统的开发与评价 .....	278
9.4.1	开发专家系统的前提 .....	278
9.4.2	专家系统的生命期 .....	279
9.4.3	专家系统的评价 .....	281
9.5	专家系统开发工具与环境 .....	282
9.5.1	程序设计语言 .....	283
9.5.2	知识工程语言 .....	283

9.5.3 辅助型工具	285
9.5.4 支持工具	286
9.5.5 专家系统开发环境	286
9.6 专家系统的发展趋势	288
小结	289
深入学习资源	290
习题	290
<b>第十章 人工智能程序设计语言</b>	<b>292</b>
10.1 LISP 语言	292
10.1.1 主要特点	292
10.1.2 表	293
10.1.3 函数	294
10.1.4 LISP 程序设计方法	297
10.1.5 LISP 编程举例	299
10.2 PROLOG	301
10.2.1 主要特点	301
10.2.2 语法与数据结构	301
10.2.3 程序设计原理	302
10.2.4 编程举例	305
小结	305
深入学习资源	306
习题	306
<b>第十一章 智能计算机</b>	<b>307</b>
11.1 光计算机	308
11.1.1 光计算机的基本原理	308
11.1.2 光计算机的基本问题	309
11.2 量子计算机	310
11.2.1 量子计算的基本原理	310
11.2.2 量子计算方法	312
11.2.3 量子计算的主要问题	314
11.3 生物计算机	314
11.3.1 生物开关	315
11.3.2 DNA 计算机	315
小结	318

深入学习资源 .....	319
习题 .....	319
附录 1 汉英—英汉术语对照与索引 .....	320
附录 2 汉英—英汉人名对照与索引 .....	334
参考文献 .....	338

# 第一章 緒論

人工智能,译自英文“Artificial Intelligence”,简称AI,是一门关于理解人类智能内在机制,并在机器上予以实现的科学。自古以来,人类就有关于智能机器的设想,希望机器能辅助人类甚至代替人类完成一些需要通过自我思维才能完成的任务。这些设想部分体现在各种科幻电影中,因此人工智能学者鲁塞尔·贝尔勒形象地说:人工智能就是“使真正的机器表现得像科幻电影中的那些机器一样”<sup>[1]</sup>。在科技日益发达,信息量迅速膨胀,同时人们想象力日渐丰富、实现力不断增强的今天,人类对智能机器的需求更为迫切,人工智能技术的发展对社会的进步具有重要的意义,因此人工智能技术与能源技术、空间技术被并列为三大尖端技术<sup>[2]</sup>。但相对于人类在能源技术、空间技术领域业已取得的巨大成就,人类在人工智能技术领域不过刚刚起步,有不少关于人工智能的根本性问题等待人们去探索。

本章阐述人工智能的基本概念、研究目标、研究和应用领域;简要介绍人工智能的发展简史;简要讨论当前的各种人工智能实现途径,以使读者了解人工智能的基本知识,作为进一步学习的基础。同时,本章最后给出了本书的内容及其组织方法,以方便读者阅读本书。

## 1.1 智能

从工程实现的角度来看,人工智能就是人造的智能。智能与人工智能的关系,犹如皮革与人造皮革的关系。不了解皮革的构成和特性,制造不出相应的人造皮革。同样,不清楚什么是智能,难以真正理解和实现人工智能。

直觉地说,智能是人们认识和改造客观世界的综合能力,是人类区别于其他事物的本质特征。但要给“智能”一个科学的明确的定义,却是非常困难的问题,正所谓“有多少专家,就会有多少种关于智能的说法”(R. J. 斯腾伯格)<sup>[3]</sup>。解决这一问题以及由此带来的真正人工智能的实现,需要人们彻底揭示人脑及其思维能力的奥秘,但目前还没有找到揭示这一奥秘的可能途径。人们只在一点上达成一定的共识,即智能不是单个能力,而是多种能力的综合,通常包括:感

知能力、行为能力、问题求解能力、推理能力、学习与适应能力、社交能力和创造力。

### 1. 感知能力

感知能力是指人们通过视觉、听觉、嗅觉、味觉、触觉等感觉器官感知外部世界的能力。它是人类最基本的生理、心理现象，是人类获取外界信息的基本途径。

### 2. 行为能力

行为能力是指人们对外界或内部刺激做出反应的能力。反应一般通过语言、表情、肢体、动作等手段来表现。如果说感知部分是人体的信息输入机构，那么行为部分就是人体的信息输出机构。人类对感知到的外界信息，通常有两种不同的处理方式：一种是对简单或紧急情况，似乎不经大脑思索，直接做出反应，如人们骑自行车遇到紧急情况时的急刹车动作；另一种是对复杂情况，通常要经过大脑的思维，然后才能做出反应，如人们在进行重大决策时的“三思而后行”。

### 3. 推理能力

推理能力是指人们根据当前掌握的信息，得出适当结论的能力，可分为演绎推理、归纳推理、类比推理等。当面对不确定、不完全以及随时间变化的信息时，人们也能进行相应的推理。

### 4. 问题求解能力

人们在生活中总是面对各种各样的问题，如何达到机场，如何解决一道数学证明题等。人们具有解决自身所面临问题的能力。如果将问题的答案看作目标，则问题求解过程就是一种规划过程。人们能够制定出达到相应目标的计划，能够在不同的计划之间做出选择，当面对环境中不可预知的变化时还能够调整自己的计划。在此过程中，人们还表现出自治能力。人们能够自主确定目标，自主确定相应的行动计划，而无需得到其他实体明确的指令。

### 5. 学习与自适应能力

学习是一个具有特定目的的知识获取过程。学习与自适应是人类的一种本能。人们通过学习，可以增加知识、提高能力、对自己的思想和行为进行修正以适应变化中的环境。尽管不同人在学习方法、学习效果等方面有较大差异，但学习却是每个人都具有一种基本能力。

### 6. 社交能力

人类具有与他人交往的能力，能够根据共同的目标组成社会团体，能够相互协作以解决共同面临的问题，能够形成类似的思维方式、生活习惯习俗等，从而表现出相应的文化。而文化背景、社会交往等因素又会对人们的智能行为产生影响。