



国家科学技术学术著作出版基金资助



Web 智能与科学

Web Intelligence and Web Science

1

Web 智能进展

Advances in Web Intelligence

钟 宁 刘际明 姚一豫



国家科学技术学术著作出版基金资助



Web 智能与科学

Web Intelligence and Web Science

1

WEB ZHINENG JINZHAN

Web 智能进展

Advances in Web Intelligence



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

图书在版编目(CIP)数据

Web 智能进展 / 钟宁, 刘际明, 姚一豫编著. —北京: 高等教育出版社, 2011.4

(Web 智能与科学)

ISBN 978-7-04-031778-7

I. ①W… II. ①钟… ②刘… ③姚… III. ①计算机网络-程序设计 IV. ①TP393.09

中国版本图书馆 CIP 数据核字 (2011) 第 027337 号

策划编辑 刘 英 责任编辑 张海波 封面设计 张 楠
版式设计 马敬茹 责任校对 王效珍 责任印制 朱学忠

出版发行	高等教育出版社	咨询电话	400-810-0598
社 址	北京市西城区德外大街 4 号	网 址	http:// www.hep.edu.cn
邮政编码	100120		http:// www.hep.com.cn
印 刷	涿州市星河印刷有限公司	网上订购	http:// www.landracom.com
开 本	787×1092 1/16		http:// www.landracom.com.cn
印 张	18.75	版 次	2011 年 4 月第 1 版
字 数	340 000	印 次	2011 年 4 月第 1 次印刷
购书热线	010-58581118	定 价	59.00 元

本书如有缺页、倒页、脱页等质量问题, 请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 31778-00

《Web 智能与科学》丛书编审委员会

主 编

钟 宁 北京工业大学国际 WIC 研究院, 日本前桥工业大学

刘际明 北京工业大学国际 WIC 研究院, 香港浸会大学

委 员 (按姓氏拼音顺序)

高 阳 南京大学

过敏意 上海交通大学

胡 斌 兰州大学

黄本雄 华中科技大学

黄智生 北京工业大学国际 WIC 研究院
荷兰阿姆斯特丹自由大学

金国庆 香港中文大学

寇 纲 电子科技大学

李娟子 清华大学

马建华 日本法政大学

漆桂林 东南大学

史忠植 中国科学院计算技术研究所

王飞跃 中国科学院自动化研究所

王国胤 重庆邮电大学

吴信东 美国佛蒙特大学
合肥工业大学

姚一豫 北京工业大学国际 WIC 研究院
加拿大里贾纳大学

张彦春 澳大利亚维多利亚大学

支志雄 清华大学

Philip S. Yu 美国伊利诺伊大学芝加哥分校

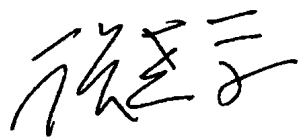
总 序

20 世纪 90 年代以来,人类社会经历了一场新的科技革命,科技进步日新月异,国际竞争日趋激烈。在这场竞争中,信息技术对全球社会经济的发展与进步起了巨大的推动作用,并极大地促进了世界经济结构的变革。网络化、智能化已成为当今世界科技革命的一个重要特征。在《国家中长期科学技术发展规划纲要(2006—2020)》中,已把与网络化、智能化相关的信息科学与技术纳入其中。因此,需尽快与国际前沿接轨并达到国际领先水平,以提高我国在该领域的国际竞争力。

在这个时代,Internet 和 Web 是发展最快的网络形式之一,也是最活跃的研究领域之一,网络智能化研究已经成为当今国际人工智能研究领域的一个新趋势。以钟宁教授、刘际明教授等为代表的一批学者顺应这一国际研究新趋势,率先对 Web 智能科学领域进行了系统的研究,获得了一系列的研究成果,在国际 Web 智能科学研究领域取得领先地位。但是,国内对 Web 智能科学的研究起步相对较晚,研究成果较少,还没有引起国内人工智能研究领域及相关研究领域科研人员的广泛关注和足够重视。

在高等教育出版社的支持下,钟宁、刘际明提议并任主编,建立了“Web 智能与科学”系列丛书,以解决国内 Web 智能科学研究领域中最新资源短缺的状况。该系列丛书主要向国内读者介绍最新的 Web 智能科学领域的学术动态、最新的研究成果和学术交流等情况。

在此,衷心希望通过该系列丛书的出版,为国内研究者构建一个 Web 智能科学研究领域交流互动的窗口,为国内该领域的研究提供最新的信息和学术资源。希望该系列丛书的出版成为我国 Web 智能科学研究进一步发展的新起点,进而带动该领域的水平提升,为把我国建设成自主创新的科技强国做出应有贡献。



2010 年 12 月

前 言

自从在 2000 年召开的 IEEE 国际计算机软件与应用学术会议 (COMPSAC 2000) 上提出 Web 智能概念以来, Web 智能已经成为一个研究热点, 被认为是一个崭新的、非常有前途的研究领域。从 Web 智能的提出到现在已经有十余年了, 十余年间 Web 智能的研究已经引起了研究者广泛的关注。Web 智能国际会议从 2001 年至今已经成功举办九届, 从 2003 年开始到现在, 每年举办一次, 每次 WI-IAT 联合会议均可收到来自 40 多个国家的 600 多篇论文, 已成为国际学术组织 IEEE 和 ACM 系列会议。在国内由中国人工智能学会粗糙集与软计算专业委员会和中国计算机学会人工智能与模式识别专业委员会主办的 Web 智能学术研讨会 (CWI) 已经成功举办四届。

作为第一个以研究 Web 智能理论及其应用为核心任务的国际化团队, 我们汇总了近几年的一些成果展示给读者, 以便进一步推广 Web 智能及其相关研究领域在国内的发展。本书重点阐述本团队近几年在 Web 智能理论和应用方面的成果, 并以问题解决为核心, 从面向自治计算和粒计算两方面阐述 Web 智能理论基础, 从 Web 知识表示与推理和 Web 挖掘方面阐述相关的技术, 从基于面向自治计算的分布式网络免疫策略、基于粒计算的 Web 知识检索与推理、基于多层网格的电子商务门户网站、面向 Web 的问答系统以及可操作电子邮件等方面阐述 Web 智能的应用实例。本书的主要内容以钟宁、刘际明、姚一豫三人公开发表的国际期刊、会议论文、专著以及在国际、国内学术会议上做的特邀报告和讲座资料为核心基础, 以北京工业大学国际 WIC 研究院的博士毕业论文为具体应用实例, 兼顾国际上有关 Web 知识表示与推理及 Web 挖掘的相关知识与技术。我们由衷感谢参与各章编撰整理工作的各位老师和同学: 黄佳进 (第 1 章, 第 8 章), 高超 (第 2 章, 第 6 章), 陈光 (第 3 章), 孙静宇 (第 4 章), 杨剑 (第 5 章), 曾毅 (第 7 章), 高明霞 (第 9 章), 李文斌 (第 10 章)。作为团队集体研究的成果, 各位老师和同学付出了大量时间和精力, 没有大家的努力, 也就不会有该书的出版。同时, 感谢高超、王凯明、王辉、廖先旭、赵碌林在统稿时所做出的贡献; 感谢周善玉为相关章节所做的工作; 感谢王劭博、马云飞为相关章节做出的翻译工作; 感谢荷兰阿姆斯特丹自由大学高级研究员黄智生博士对本书相关研究提出的建议。特别感谢高等教育出版社刘英编辑在该书构思、撰写和出版过程中提供的建议和无私帮助。

最后,感谢国家自然科学基金面上项目、北京市自然科学基金面上项目对书中相关研究的支持。为了进一步提高本书的质量,我们也热忱欢迎读者把对本书的意见和建议反馈给我们。如有不当之处,请批评指正。

钟宁 刘际明 姚一豫

2010年10月

目 录

第 1 章 Web 智能介绍	1
1.1 从不同角度理解 Web 智能	1
1.1.1 智慧 Web 角度	2
1.1.2 物联网角度	4
1.1.3 功能和技术角度	7
1.1.4 Web 信息系统角度	7
1.1.5 问题解决角度	9
1.1.6 人机交互角度	11
1.1.7 人类智能研究角度	12
1.2 本书组织结构	13
1.2.1 面向自治计算与 Web 智能	15
1.2.2 粒计算与 Web 智能	15
1.2.3 Web 知识表示与推理	16
1.2.4 Web 挖掘	16
1.2.5 基于多层网格的电子商务门户网站	17
1.2.6 面向 Web 的问答系统	17
1.2.7 可操作电子邮件系统	18
1.3 小结	18
参考文献	19

第一部分 Web 智能理论

第 2 章 面向自治计算	27
2.1 面向自治的计算简介	27
2.1.1 面向自治计算思想的由来	27
2.1.2 面向自治计算的主要机制和强调重点	28
2.1.3 面向自治计算和其他概念的比较	29
2.1.4 面向自治计算方法在 Web 智能中的应用	31
2.2 面向自治计算的应用实例	33
2.2.1 复杂难问题求解	33
2.2.2 复杂系统建模	40

2.3 小结	43
参考文献	44
第3章 粒计算	50
3.1 粒计算概述	50
3.1.1 粒计算基本思想	50
3.1.2 粒计算的普适性	57
3.1.3 粒计算研究内容及目标	59
3.2 粒结构	62
3.2.1 粒	62
3.2.2 层	63
3.2.3 分层结构	64
3.2.4 粒结构	66
3.3 粒计算三元论模型	66
3.3.1 模型的组成与结构	67
3.3.2 结构化思维——粒计算的哲学	69
3.3.3 结构化问题求解——粒计算的方法论	70
3.3.4 结构化信息处理——粒计算的计算模式	73
3.4 小结	78
参考文献	79
第4章 Web 知识表示与推理	91
4.1 Web 知识表示	91
4.1.1 本体	91
4.1.2 资源描述框架	94
4.1.3 Web 本体语言	101
4.1.4 问题解决标记语言	106
4.2 Web 推理	115
4.2.1 基于 OWL 的推理	116
4.2.2 Web 案例推理技术	121
4.2.3 Web 规模推理	122
4.2.4 Web 推理典型应用	123
4.3 小结	126
参考文献	126
第5章 Web 挖掘	129
5.1 数据挖掘	129
5.1.1 关联规则	130
5.1.2 文本挖掘	132

5.1.3 图挖掘	135
5.1.4 特异性指向挖掘	137
5.2 Web 的特点	140
5.3 Web 挖掘	141
5.3.1 Web 内容挖掘	142
5.3.2 Web 结构挖掘	143
5.3.3 Web 日志挖掘	144
5.3.4 组合挖掘	145
5.3.5 Web 挖掘相关技术	146
5.4 Web 耕作	147
5.5 小结	149
参考文献	149

第二部分 Web 智能应用

第 6 章 基于面向自治计算的免疫网络	157
6.1 网络免疫策略介绍	157
6.2 分布式网络免疫问题陈述	159
6.3 基于 AOC 的分布式免疫策略设计	161
6.3.1 计算体与环境定义	162
6.3.2 局部规则与评估函数设计	163
6.3.3 基于 AOC 的分布式搜索策略主函数	167
6.4 实验	168
6.4.1 搜索效率分析	169
6.4.2 免疫效率分析	178
6.5 小结	179
参考文献	179
第 7 章 基于粒计算的 Web 检索与推理	182
7.1 Web 检索支持系统的粒计算框架	182
7.1.1 信息检索支持系统的粒计算框架	183
7.1.2 基于粒结构的知识检索	187
7.2 从粒度的视角融合 Web 规模的搜索与推理	193
7.2.1 起始点策略	194
7.2.2 多层完备度策略	197
7.2.3 多层特定度策略	201
7.2.4 多视角策略	203
7.3 小结	205
参考文献	206

第 8 章 基于多层网格的电子商务门户网站	210
8.1 研究背景和相关研究	210
8.2 三层网格体系结构	211
8.3 市场值函数挖掘	213
8.4 基于用户行为的目标营销数据挖掘的框架	217
8.4.1 特异性挖掘 Agent	218
8.4.2 市场值函数挖掘 Agent	219
8.4.3 知识网格应用	219
8.4.4 实验评价	220
8.5 小结	224
参考文献	224
第 9 章 面向 Web 的问答系统	227
9.1 问答系统介绍	227
9.2 面向 Web 的问答系统	231
9.2.1 开域问答系统	232
9.2.2 受限域知识问答系统	238
9.2.3 开域问答系统和受限域知识问答系统的比较	252
9.3 小结	253
参考文献	254
第 10 章 可操作电子邮件系统	257
10.1 万维电子邮件网	257
10.2 WVEW 的访问协议	258
10.2.1 可操作电子邮件的语言栈	258
10.2.2 在 WVEW 中实现智能应用	260
10.2.3 与可操作电子邮件相关的研究问题	261
10.2.4 可操作电子邮件的一种实现方案	261
10.3 可操作电子邮件应用初探	266
10.3.1 异步 P2P 应用	266
10.3.2 邮客	270
10.3.3 可操作电子邮件应用的一个说明性演示程序	270
10.4 个人辅助应用	273
10.4.1 总体框架	274
10.4.2 基于本体的电子邮件归档	275
10.4.3 基于时间窗口技术的用户行为学习	276
10.4.4 基于多智能体的邮件过滤及实验结果	278
10.5 小结	281
参考文献	282

第1章 Web 智能介绍

在以计算机为基础的研究领域内,人工智能得到了广泛的研究。很多学者研究各种可计算的模型,并且开发出各种基于知识的模型和原型系统。这些模型和系统已经被用于实现推理、学习和问题求解。随着 Internet 技术的发展,Web 已经成为人们工作和生活的重要工具和平台。如何使人工智能技术能够适应 Web 这一开放的、巨大的、分布的、动态的多信息源环境,是研究者面临的新的挑战。

为了研究如何在 Web 环境中实现智能应用,出现了 Web 智能(Web Intelligence, WI)这一新的研究方向。Web 智能可以定义为系统地研究基于 Web 的人工智能与高级信息技术相结合的原理和技术,其最终目标是实现智慧 Web(Wisdom Web)^[1-9]。Web 智能是信息技术在 Web 时代最为关键和紧急的研究领域之一,体现了计算机网络与人工智能两个领域在高层次上的结合和创新。

自从 2000 年 Web 智能^[1]概念在 IEEE 国际计算机软件与应用学术会议(COMPSAC 2000)被首次提出以来,Web 智能已经成为一个研究热点,被公认为是一个崭新的、非常有前途的研究领域^[10]。本书主要内容以三位编著者公开发表或出版的国际期刊与会议论文、专著以及在国际、国内学术会议上做的特邀报告和讲座资料为核心内容,以近期北京工业大学国际 WIC 研究院在 Web 智能方面取得的研究进展为具体应用实例,兼顾国际上有关 Web 知识表示与推理及 Web 挖掘的相关知识与技术。全书重点阐述本团队近几年在 Web 智能理论和应用方面的成果,并以问题解决为核心,从面向自治计算和粒计算两方面阐述 Web 智能理论基础,从 Web 知识表示与推理和 Web 挖掘方面阐述相关的技术,从基于面向自治计算的分布式网络免疫策略、基于粒计算的 Web 知识检索与推理、基于多层网络的电子商务门户网站、面向 Web 的问答系统以及可操作电子邮件等方面阐述 Web 智能的应用实例。

本章首先从 Web 的发展趋势(智慧 Web 和物联网)、Web 智能实现的主要技术、Web 信息系统、问题解决、人机交互和人类智能等几方面对 Web 智能进行介绍。

1.1 从不同角度理解 Web 智能

计算机网络的发展对人们的生活、学习和工作等产生了深刻的影响。按照信

息科学和知识管理对数据、信息、知识和智慧 (Wisdom) 不同层次的定义, 可以将 Web 分成数据 Web、信息 Web、知识 Web 和智慧 Web^[9]。目前, 与 Web 相关的技术可以处理 Web 上的数据、信息并在一定程度上处理知识。为了使 Web 更好地为人类服务, 将 Web 转变成具有人类智能水平的智慧 Web 是新一代 Web 发展的重要方向。通过智慧 Web, 人们不仅可以从 Web 上浏览和搜索信息, 而且可以通过 Web 获得学习、工作、生活和娱乐的智慧。

研究 Web 上的问题将涉及语义 Web (Semantic Web)^[11]和智慧 Web^[6]这两个 Web 的重要发展趋势。语义 Web 是国际学术组织 W3C 提出的研究方向, 它是对现有 Web 的扩展, 注重解决目前 Web 上语义缺乏的问题。在语义 Web 上, 信息被赋予定义良好的含义, 更便于计算机之间以及计算机和人之间的交互。语义 Web 提供了一个通用的框架, 允许跨越不同的应用程序、企业和团体共享数据。智慧 Web 则可看作是按照信息科学和知识管理领域中数据、信息、知识和智慧之间的层次关系来定义的。目前的 Web 技术仍局限于处理数据、信息和在一定程度上处理知识, 由此可以将智慧 Web 看作 Web 发展的重要趋势之一。基于智慧 Web 的计算机制不仅能提供对知识和经验的共享, 而且还提供基于自组织资源的可持续知识创造以及科学和社会的发展。

当前, Web 科学、语义技术和 Web 智能是 Web 研究领域的三个主要理论体系。Web 科学和语义技术更是直接派生于语义 Web 理论体系。Web 的发明人 Tim Berners-Lee 等 2006 年在《Science》上发表了“Creating a science of the Web”^[12]一文, 呼吁建立一个崭新的学科——Web 科学。与传统计算机科学比较, Web 科学的观察和研究视角更广泛, 是多学科交叉和融汇的产物^[13]。Web 智能研究更强调通过多学科集成与交叉创新来开发智慧 Web 的理论基础与应用。

1.1.1 智慧 Web 角度

文献[6]对智慧 Web 进行了描述: 智慧 Web, 可以简单地理解为在合适的时间 (Right Time), 将合适的内容 (Right Content) 和合适的服务 (Right Service) 提供给合适的对象 (Right Object)。为了实现这样一个智慧 Web, 图 1-1 构建了一个从现实世界到基于 Web 的虚拟世界的映射。Web 智能就是要构建这样一个虚拟世界。在这个虚拟世界中, 电子/移动 (Electronic/Mobile, e/m)-商务、电子/移动-科学、电子/移动-学习、电子/移动-政府、电子/移动-金融、电子/移动-医疗、电子/移动-健康、电子/移动-社区等在基于 Web 的信息系统中得以实现。同时, 在网格、云计算和 Web 服务等平台下, 收集各种实际电子应用系统产生的数据, 并对这些数据进行管理、分析和利用, 产生新的知识。根据已有的知识和新产生的知识为用户提供个性化的服务。

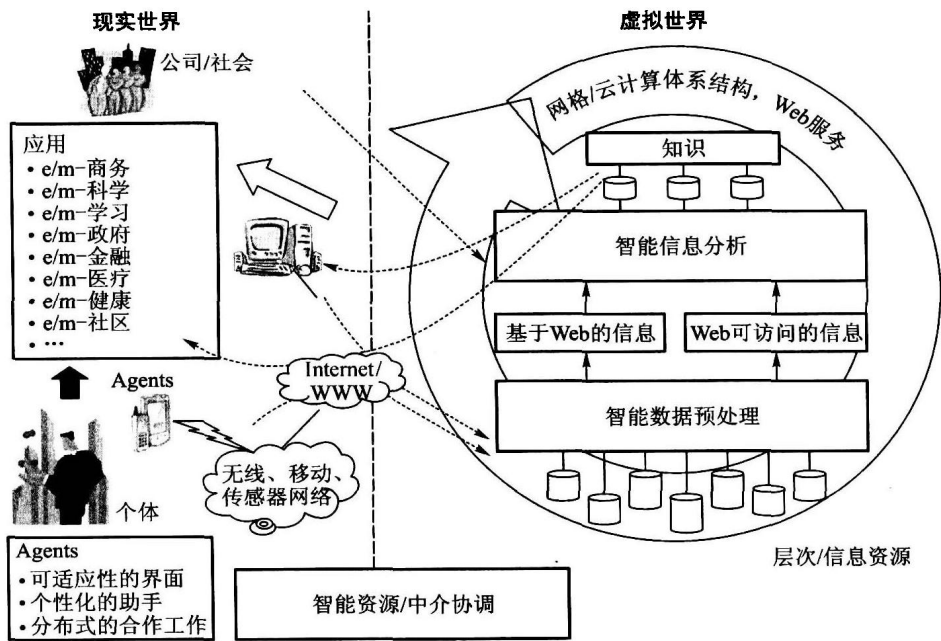


图 1-1 现实世界和虚拟世界

文献[6]指出了在研究智慧 Web 过程中可能遇到的一些挑战和各种研究课题之间的关系,如图 1-2 所示。箭头方向表明了一个研究课题的解决如何来支持另一个研究课题的解决。在每个方框内,标题部分表示有待解决的技术问题,括号里的内容则表示相应的关键点。

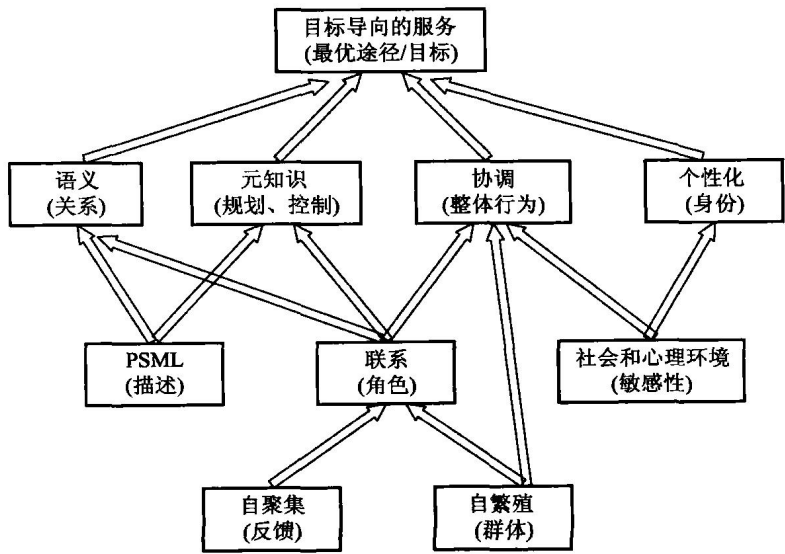


图 1-2 智慧 Web 研究路线图

一般来说, 智慧 Web 依赖于一个所谓智慧 Agent 的社会, 这些智慧 Agent 能够自主地进行发现、监测和管理网络化的资源。根据它们所处理的资源类型的不同, 这些 Agent 将表现出不同的功能。这些资源包括有形资源, 如数据库和计算设备; 也包括无形资源, 如长时间日积月累而产生的大量的经验、信息以及大规模的社会网络。在智慧 Agent 的生存周期内, 通过收集和整合那些在与外界环境相互作用中接收的反馈信息来进行功能的自聚集, 从而提升它们的专业性和技能。在这个过程中, 将会产生一些继承自父类的专业知识和技能且对特定任务具有更加专业化特性的子类 Agent。同时, 那些不成功的 Agent 将随着生存周期的结束而消亡。这样, 这些智慧 Agent 将能够动态地适应群体不断繁殖和变化的环境。

自聚集和自繁殖的过程将会进一步促使智慧 Agent 与一些特定类型的任务之间的联系得以加强, 同时也能强化那些在特定环境下将要展开的相互关系(为了保持成功)。这样, 智慧 Web 中 Agent 角色这个概念就出现了。

在这种联系过程中, 一方面 Agent 将会逐渐发展并且自组织它们的世界模型(也就是建立语义本体), 这些模型用来描述从环境中观察到的各种性质之间的相互关系和属性; 另一方面将为实现行为规划和控制的元知识以及为了从个体中产生目标导向的行为(群体层次上)而进行相应的协调。在该过程中, 使用问题解决标记语言(Problem Solver Markup Language, PSML)(包括它的形式体系以及相关关联的控制操作)可有效地描述和操纵(如更新升级)语义和元知识。

社会和心理方面的环境作为另一个因素使得 Agent 的行为对其所处的环境的感知更具敏感性, 而人类在交流过程中以及在不同既定社会条件下解决难题时也经常会表现出这种敏感性。与此同时, 这种敏感性将会促进个性化智慧 Agent 和服务的发展, 即: 将在社会和心理环境中建立 Agent 身份这一概念。通过基于本体关系的语义世界模型、基于策略规划和控制的解决问题方法、为实现目标导向的全局行为的群体协调以及拥有社会和心理身份的个性化 Agent, 智慧 Web 将能够构建和传递有求必应的、目标导向的服务, 即通过最优途径达到最优目标。

1.1.2 物联网角度

在 HTTP、HTML/XML 等 Web 核心技术的支持下, 一个全新的网络 Cyber 空间(Cyber Space)正在形成, 它在人的社会世界(Social World)和物的物质世界(Physical World)之间构建了一个飞速发展的 Cyber 世界(Cyber World), 或称之为信息世界(Information World)^[14]。

近年来, 新的信息、通信技术不断推动计算机世界向社会世界和物质世界扩展。一方面, Web 2.0^[15,16]、Web 3.0^[17,18]、网格计算(Grid Computing)^[19-21]、服务计算(Service Computing)、云计算(Cloud Computing)^[22,23]等基于互联网

的新应用的出现,使得网络数字空间不再是一个简单的信息检索服务系统,而成为一个全球化的交流和协作平台。各种虚拟的社区、社团、组织在这个空间中大量涌现,计算机世界正不断向社会世界扩展;另一方面,嵌入式技术、射频自动识别(RFID)技术、无线数据通信技术、普适计算等技术正推动着物联网的形成。进一步,传感器 Web 将分布在物质世界中的传感器有效地连接起来,实时地、持续地收集多样化的数据。借助传感器 Web,这些数据被汇入 Web,实现了计算机世界向物质世界的扩展^[24-29]。

当前,Web 2.0、Web 3.0、云计算、Smart World^[30]、Smart Planet^[31]、Green/Eco Computing^[32]等各种基于互联网/物联网的应用的研究已得到普遍关注。可以预见,在不久的将来,基于计算机互连的计算机世界将与人的社会世界和物的物质世界逐渐融合。一个涵盖了人、机、物的超世界(Hyper World)^[33]将逐渐形成,并对人类社会以及每个社会成员的生产、生活方式产生深远的影响^[34]。为应对随之而来的挑战和机遇,需要多领域专家之间的紧密合作。

超世界带来的关键问题是如何利用信息技术的网络化、普适化、智能化实现人-机-物的有机融合与和谐共存共生。自然界的基础是物质,一个有效的物质循环系统保证了自然界中各种事物的和谐发展。而超世界的基础是数据,要实现超世界中人-机-物的和谐发展,就需要一个高效的循环系统来实现“物-数据-信息-知识-智慧-服务-人-物”的数据循环。该系统必须能及时、准确地捕捉超世界中不断变化的人和物的情况(数据),从中抽取信息、知识,“智慧”地理解用户的真实需求,将主动的、透明的、安全的、可信的服务提供给用户。

虽然目前传感器 Web 有着多样化的定义,但是我们认为它可以被用来描述将各种传感器连接到 Web 后形成的一个横跨互联网和物联网的统的数据/资源存储,以及共享服务、提供交流和协作的平台。它是构建在传感器网络(Senor Network)上的一个应用系统,是超世界的数据载体和流通渠道。

作为超世界的数据载体和流通渠道,传感器 Web 要在超世界中建设高效的数据循环系统,就必须进行传感器 Web 的“智能化”,在智慧传感器 Web 中实现“物-数据-信息-知识-智慧-服务-人-物”的有效循环^[56]。其循环过程可简略地描述如下。

(1) 物-数据:通过传感器 Web,多样化的物的数据,包括来自传感器和检测设备(如 MRI、EEG、CT)的物的实时数据、各种可通过 Web 访问(Web Accessible)的物的历史性数据、各种 Web 数据(由人输入 Web 的数据和 Web 上人的活动产生的数据)都被收集到分布式的综合数据中心。

(2) 数据-信息:在数据中心,将来自于 Web 的、全局的历史数据和来自于传感器 Web 的、局部的实时数据进行收集、清理、集成和存储后,借助数据挖掘、案例构建等数据处理和组织方法,生成面向不同应用的多层次、多粒度的数据信息,并有效地表达和存储这些信息。

(3) 信息-知识:在数据中心,采用知识工程的各种手段,从信息中提炼出有价值的知识,并进行合理的描述和存储。

(4) 知识-智慧: 基于抽取的知识, 充分利用已有的信息和数据, 针对智慧 Web 所面临的各种问题, 开发智慧 Web 的关键技术。

(5) 智慧-服务: 使用智慧 Web 的关键技术构建透明服务 (Transparent Service) 平台, 提供以用户为中心的安全的、可靠的服务。

(6) 服务-人: 通过多样化的传感器和行为者 (Actuator), 将服务平台的各种服务提供给人。

(7) 人-物: 在接受服务的过程中, 人的行为会不断影响周围的事物, 进而带来物 (包括传感器 Web 本身) 的数据的变化。

上述过程是围绕着综合数据中心和透明服务平台^[34-38]而进行的, 它们是超世界数据循环系统的核心。所谓“综合数据中心”是指该数据中心实现了数据来源、层次和内容三个层面上的综合。在数据来源上, 该数据中心不仅收集来自于 Web 的数据, 还收集来自于多样化传感器的数据; 在层次上, 该数据中心不仅包括数据层, 还包括信息层和知识层; 在内容上, 面向不同的应用, 该数据中心包含了脑数据中心、电子健康数据中心、电子医疗数据中心等多个子中心。所谓“透明服务平台”是指该平台将提供以用户为中心的“透明服务”, 即用户根据自己的需要获取安全的、可信的服务, 而不受限于服务所在的平台、网络以及接受服务的客户端类型。通过传感器 Web, 图 1-1 中的实现可以扩展为图 1-3。

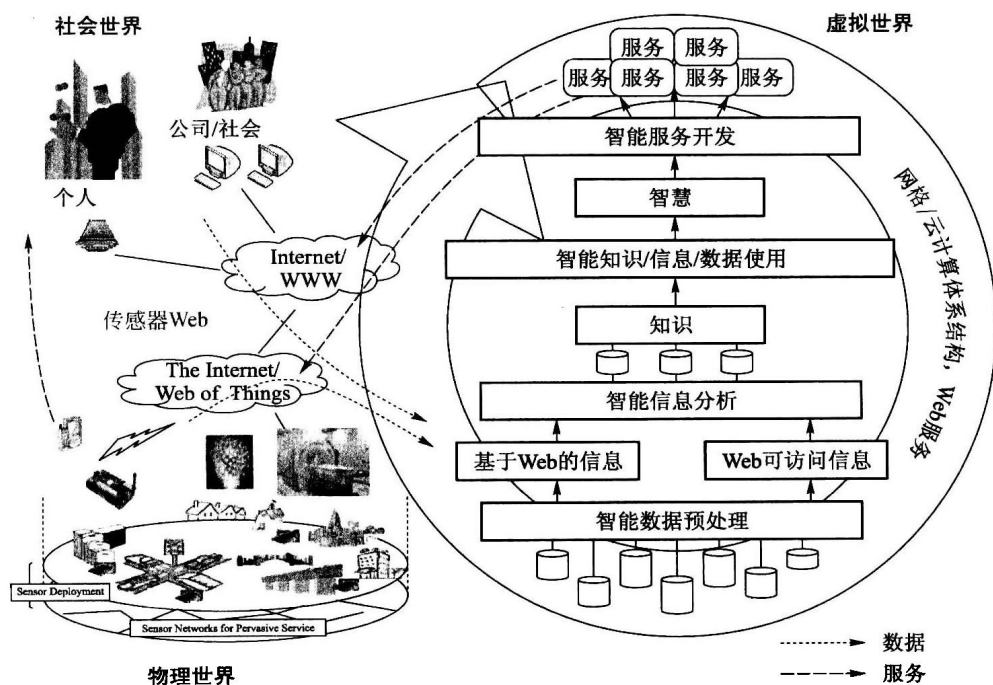


图 1-3 社会世界—物理世界—虚拟世界